

***Final Draft***  
**of the original manuscript:**

Wilson, R.; Anchukaitis, K.; Briffa, K.R.; Buentgen, U.; Cook, E.; D'Arrigo, R.;  
Davi, N.; Esper, J.; Frank, D.; Gunnarson, B.; Hegerl, G.; Helama, S.;  
Klesse, S.; Krusik, P.J.; Linderholm, H.W.; Myglan, V.; Osborn, T.J.;  
Rydval, M.; Schneider, L.; Schurer, A.; Wiles, G.; Zhang, P.; Zorita, E.:

**Last millennium northern hemisphere summer temperatures  
from tree rings: Part I: The long term context**

In: Quaternary Science Reviews (2016) Elsevier

DOI: 10.1016/j.quascirev.2015.12.005

## **Last millennium Northern Hemisphere summer temperatures from tree rings: Part I: the long term context**

Rob Wilson, Kevin Anchukaitis, Keith R. Briffa, Ulf Büntgen, Ed Cook, Rosanne D' Arrigo, Nicole Davi, Jan Esper, Dave Frank, Björn Gunnarson, Gabi Hegerl, Samuli Helama, Stefan Klesse, Paul J. Krusic, Hans W. Linderholm, Vladimir Myglan, Timothy J. Osborn, Miloš Rydval, Lea Schneider, Andrew Schurer, Greg Wiles, Peng Zhang and Eduardo Zorita.

### **Abstract**

Large-scale millennial length Northern Hemisphere (NH) temperature reconstructions have been progressively improved over the last 20 years as new datasets have been developed. This paper, and its companion (Part II, Anchukaitis et al. *in prep*), details the latest tree-ring (TR) based NH land air temperature reconstruction from a temporal and spatial perspective. This work is the first product of a consortium called N-TREND (Northern Hemisphere **T**ree-Ring **N**etwork **D**evelopment) which brings together dendroclimatologists to identify a collective strategy for improving large-scale summer temperature reconstructions. The new reconstruction, N-TREND2015, utilises 54 records, a significant expansion compared with previous TR studies, and yields an improved reconstruction with stronger statistical calibration metrics. N-TREND2015 is relatively insensitive to the compositing method and spatial weighting used and validation metrics indicate that the new record portrays reasonable coherence with large scale summer temperatures and is robust at all time-scales from 918 to 2004 where at least 3 TR records exist from each major continental mass. N-TREND2015 indicates a longer and warmer medieval period (~900 – 1170) than portrayed by previous TR NH reconstructions and by the CMIP5 model ensemble, but with better overall agreement between records for the last 600 years. Future dendroclimatic projects should focus on developing new long records from data-sparse regions such as North America and eastern Eurasia as well as ensuring the measurement of parameters related to latewood density to complement ring-width records which can improve local based calibration substantially.

### **Keywords**

Tree-rings; Northern Hemisphere; Last Millennium; Summer Temperatures; Reconstruction; CMIP5 Models

### **Introduction**

Over the past two decades, substantial effort has been directed towards reconstructing late Holocene Northern Hemisphere temperature trends and variability from high resolution palaeoclimate proxy data (Frank et al. 2010, Masson-Delmotte et al., 2013). While some studies have focused entirely on the use of tree rings (Esper et al. 2002; D'Arrigo et al. 2006; Frank et al. 2007a; Wilson et al. 2007; Schneider et al. 2015; Stoffel et al. 2015), most have also included different types of proxy archives and are thus considered multi-proxy reconstructions (Jones et al. 1998; Mann et al. 1999, 2009; Moberg et al. 2005, Hegerl et al. 2007, Wahl and Amman 2007; Ljungqvist 2010). The majority of these reconstructions have

generated a single hemispheric wide mean temperature composite, although a few noteworthy spatial field reconstructions have been produced (Briffa et al. 2002b; Mann et al. 1998, 2009; Ljungqvist et al. 2012; Tingley and Huybers 2013). Such analyses are crucial not only to place recent warming in a longer term spatiotemporal context, but also allow, through comparative analysis with global climate models, improved understanding of the high-to-low frequency forcing mechanisms of long-term and recent climate changes using formal detection and attribution (Hegerl et al. 2006; Schurer et al. 2014).

Figure 5.7 in the most recent IPCC Working Group 1 report (section 5.3.5; Masson-Delmotte et al., 2013) synthesises the majority of the published Common Era reconstructions for the Northern Hemisphere (NH). Focusing on the last millennium, the NH ensemble indicates warm conditions around the end of the 10<sup>th</sup> century, a cooling trend over the subsequent 500 years, a prolonged cool period from ~1450-1850 (often referred to as the Little Ice Age) and steep warming from the middle of the 19<sup>th</sup> century to present. Although the basic centennial “shape” of these NH reconstructions is in general agreement, there are notable differences in the timing and amplitude of some cold and warm periods (e.g. Esper et al. 2004). This uncertainty hampers detection and attribution studies and makes it difficult to constrain modelled scenarios of future temperature change (Edwards et al. 2007).

We hypothesise that these differences partly reflect the use of multi-proxy data in the current ensemble of NH reconstructions and the quality of many constituent proxy series. Although the multi-proxy approach may appear ideal - incorporating and subsuming the independent advantages and disadvantages of the various proxy types - in practice the biases inherent to each proxy archive potentially compound the uncertainty in the final large-scale composites, which are further exacerbated by the influence of the varying statistical methods used to combine the proxy records (Smerdon et al. 2011). For example, many of the proxies utilized in Common Era reconstructions actually reflect varying seasonal signals (e.g. summer vs. annual) and some do not even represent local temperature well (Mann et al. 1998, 2009). Combining such heterogeneous data may produce a network possessing a reasonable correlation with an annual instrumental target, but that also contains spatio-temporal proxy biases that add to the uncertainty in the final reconstruction. Furthermore, non-annual proxies, lacking precise annual resolution and dating, must be smoothed to decade or longer time-scales reducing (1) the ability of these data to accurately record short-term climatic response to large volcanic events, and (2) the degrees of freedom – which places limitations on the ability to perform robust calibration and validation analyses to assess reconstruction fidelity, stability and skill.

In this paper (and its part II companion, Anchukaitis et al. *in prep*), we focus entirely on temperature-sensitive tree-ring (TR) records to both reduce and better characterize many of the uncertainties detailed above. Tree-ring archives are annually resolved (facilitating robust validation), precisely dated (Stokes and Smiley 1968) and the interpretation of their measurements is supported by a wealth of ecological and biological process-based knowledge of how climate variability influences ring formation (Fritts 1976, Korner 2003, Vaganov et al. 2006; Deslauriers et al. 2007, Vaganov et al. 2011, Rossi et al. 2012, Rossi et al. 2013, Rossi et al. 2014, Palacio et al. 2014). Temperature-sensitive TR archives are derived from many high elevation/latitude environments where growth is predominantly limited by summer temperatures, permitting the development of a well replicated, annually resolved network of summer temperature proxies for the mid/high latitudes of the Northern Hemisphere (Briffa et al. 2001). Furthermore, as the summer season is most sensitive to short term volcanically forced cooling, a tree-ring only reconstruction increases the potential for more robust assessment of volcanic forcing in attribution studies in comparison to large-scale mixed season multi-proxy composites (Schneider et al. 2015; Stoffel et al. 2015).

Herein, we introduce a new and substantially updated network of temperature-sensitive tree-ring records for the mid-to-high latitudes of the Northern Hemisphere. These data are used to reconstruct both a single mean hemispheric series (Part I) as well as explicit spatial temperature fields (Part II, Anchukaitis et al. *in prep*) for the last millennium. Careful comparison is made between the new reconstruction and two previous TR based NH index reconstructions; D'Arrigo et al. (2006 – hereafter DWJ06 – a predominantly ring-width (RW) based composite) and Schneider et al. (2015 – hereafter SCH2015 – a maximum latewood density (MXD) based composite) to explore the strengths and limitations of RW and MXD. This work is the product of a tree-ring community consortium called N-TREND (**N**orthern Hemisphere **T**ree-Ring **N**etwork **D**evelopment). The aim of N-TREND is to bring together dendroclimatologists to identify a collective strategy for improving large-scale summer temperature reconstructions from new and existing tree-ring archives. These two papers, the first product of the N-TREND consortium, aim to not only improve our knowledge of Northern Hemisphere temperature changes, but to also emphasise the strengths and limitations of TR archives and guide future dendroclimatic research.

## Tree-Ring Data

Many thousands of tree-ring site chronologies are archived within the International Tree-Ring Data Bank but the majority are either relatively short (< 200-years in length), have not been updated to the recent post-2000 period and/or do not reflect summer temperatures (St. George 2014; St George and Ault 2014). For N-TREND, rather than statistically screening all extant TR chronologies for a significant local temperature signal, we utilise mostly published TR temperature reconstructions (or chronologies used in published reconstructions) that start prior to 1750. This strategy explicitly incorporates the expert judgment the original authors used to derive the most robust reconstruction possible from the available data at that particular location. The data-set used herein is defined as version number N-TREND2015, recognizing the fact that as new local/regional tree-ring temperature reconstructions are developed, the N-TREND database will expand and be further refined. N-TREND2015 includes 54 TR records located between 40°N and 75°N (Table 1 and Figure 1). TR data south of 40°N are not included as trees tend to exhibit an increasingly complex sensitivity to multiple climate influences including precipitation (Fritts 1976, St. George 2014, St George and Ault 2014) which will bias the temporal and spectral properties of a reconstruction (Osborn and Briffa 2000; Franke et al. 2013). The N-TREND data-set includes single site, regional and large-scale gridded reconstructions. Specific information on each TR series is detailed in Appendix Table A1.

Most of the TR records have been processed using Regional Curve Standardisation (RCS - Briffa et al., 1996; Esper et al., 2003 – Table 1) which improves the likelihood of retaining secular trends beyond the mean sample length of the chronologies (Cook et al. 1995). Eighteen records are based entirely on maximum latewood density (MXD), and 11 utilise only ring-width (RW) measurements. A further 25 records are defined as “mixed” and represent a heterogeneous group of local/regional/gridpoint reconstructions that have incorporated combinations of RW, MXD and Blue Intensity (BI - defined in Appendix Table A1 – McCarroll et al. 2002; Campbell et al. 2011; Björklund et al. 2014; Rydval et al. 2014) ring parameters aggregated using a variety of multivariate methodologies (Appendix Table A1). Post-1950  $r^2$  values of the individual records regressed on their local gridded (CRUTS3.2 - Harris et al. 2014), empirically derived, optimal temperature season (Table 1) range from 0.12 to 0.71 (Figure 1). Of particular note are the overall weaker results observed for those sites derived using only RW data. When MXD (or BI) is used, local/regional correlations are almost always superior. The common period for all 54 records is 1710-1988; 38 records extend back to 1500 while 23 records extend back to at least 1000. Substantially more TR

records are available for Eurasia (39) compared to North America (15) with a distinct imbalance of data at 1000 (20 vs. 3 – Figure 2a).

## Reconstruction Methodology

A similar iterative nesting method (Meko 1997; Cook et al. 2002), as utilised in D'Arrigo et al (2006) and Wilson et al. (2007), was used to develop the N-TREND2015 NH temperature reconstruction. This approach involves first normalising the TR data over a common period (1750-1950), averaging the series to derive a mean series and iteratively removing the shorter series to allow the extension of the reconstruction back (as well as forward) in time. Each nest is then scaled to have the same mean and variance as the most replicated nest (hereafter referred to as NEST 1) and the relevant time-series sections of each nest spliced together to derive the full-length reconstruction. For each nest, separate average time series were first generated for 4 longitude quadrats (Figure 1). These continental scale time series were then averaged (after again normalising to 1750-1950) to produce the final large-scale hemispheric mean to ensure it is not biased to data rich regions in any one continent. 37 backward nests and 17 forward nests were calculated to produce the full reconstruction length from 750 – 2011. Other methodological options (i.e. different spatial weighting procedures and total least squares regression) were also tested to assess the sensitivity of the final NH series to varying compositing and reconstruction approaches (see Appendix C).

Full period calibration was performed over the 1880-1988 period targeting the latitudinal band 40-75 °N. We utilised herein the land-only CRUTEM4 data-set (Jones et al. 2012) although similar coherence results are obtained using different gridded temperature products (Appendix Figure B1a). Calibration experiments using NEST 1 identified mean May-August (MJJ) temperatures as the optimal large-scale instrumental target. Separate calibration and validation analyses were performed on each nested mean series allowing assessment of temporal changes in reconstruction quality. For independent validation of the TR based temperature estimates, calibration/validation was performed using a traditional split period (1880-1934 and 1935-1988) approach as well as calibrating on the combined 1880-1915/1953-1988 periods and verifying over 1916-1952. As large-scale NH instrumental temperatures show a steep increasing trend since the 19<sup>th</sup> century, the reduced variance of the simple regression based estimates make it difficult to predict the steep trend outside a short calibration period, particularly when calibration  $r^2$  values are modest. The latter calibration/validation approach allows a theoretical interpolation based assessment of the temperature estimates rather than the extrapolations of the simple early/late split period approach.

Evaluation of the reconstruction was performed using the square of the Pearson's correlation coefficient ( $r^2$ ), the Reduction of Error (RE) statistic, and the Coefficient of Efficiency (CE) statistic (Cook and Kairiukstis 1990; Cook et al. 1994). To test the robustness of the decadal to long-term signal in the reconstructed nested series, the regression model residuals (from the full period calibration) were assessed for linear trends, and autocorrelation using the Durbin-Watson statistic (see Wilson et al. 2006). A poor composite proxy fit at decadal time-scales will result in significant residual autocorrelation, while biases in centennial scale trend estimates would result in residual linear trends. The root mean square error (RMSE) was used to estimate reconstruction uncertainty. To derive the final reconstruction, NEST 1 was scaled to the instrumental data over the 1880-1988 period and the RMSE values calculated using the residual differences between the actual values and scaled estimates (referred to as  $RMSE_{scl}$ ). All other nested series were scaled to NEST 1 over the 1710-1988 period, and the  $RMSE_{scl}$  value, calculated for each nest, provides an estimate of the change in calibration based uncertainty through time.

Although many of the individual reconstructions (Table 1) have published uncertainty estimates, some do not. Even for those that have presented reconstruction uncertainty, they are mostly related to calibration uncertainty, and conventionally, few tree-ring temperature reconstructions have attempted to quantify uncertainty related to both detrending choices and decreasing sample size back in time as the numbers of trees decrease (Esper et al. 2007; Matsuksky and Helama 2014; Wilson et al. 2014). This is problematic as the actual uncertainty of such large-scale reconstructions will be greater than that simply estimated from the nested mean composite calibration based  $RMSE_{sc1}$  values. This extra uncertainty will increase back in time but there is no single way to quantify it, particularly without fundamental knowledge of the raw data used for individual tree-ring chronologies. To produce a conservative estimate of this additional error that we acknowledge may exist in our data, a theoretical worst case scenario is proposed wherein an individual input reconstruction is degraded until it has zero correlation with its local gridded data. The maximal amount of error arising in this scenario is modelled by fitting a horizontal straight-line to the instrumental data and calculating the grid specific RMSE. The grand mean RMSE for all 54 grid locations comes to 1.00 °C ( $\pm 0.23$  standard deviations). This value was used to derive an additional uncertainty estimate (referred to as  $RMSE_{add}$ ) that is added to the  $RMSE_{sc1}$  values but is scaled by dividing the  $RMSE_{add}$  by the square root of N (number of records) as input series replication changes through time (see Appendix Figure B2).

To compute the error for the smoothed reconstruction, we follow a method first outlined by Briffa et al. (2002a - modified by Gouirand et al. 2008) and applied by Divine et al. (2011), but modified for use with RMSE rather than the standard error of the regression estimate:

$$(RMSE_{adj})^2 = \frac{1}{r^2} \left( \frac{(RMSE_{tot})^2}{n} \right)$$

where  $RMSE_{tot} = RMSE_{sc1}$  and  $RMSE_{add}$  (see Appendix Figure B2), n is the window length in years of the smoothing function (a 20-year cubic smoothing spline (Cook and Peters 1981)) and  $r^2$  is the explained variance for each smoothed nest calibration.

## Results and Discussion

### ***Calibration and Validation of the N-TREND2015 reconstruction***

The N-TREND2015 reconstruction and associated quality metrics are shown in Figure 2 (B-E – calibration/validation results detailed in Appendix Table B1). NEST 1, representing the period 1710-1988, with an unfiltered calibration  $r^2$  of 0.41, does not explain the greatest amount of instrumental variance in the full record. NEST 11 (relevant for the period 1510-1520) explains the most temperature variance at 47%. This appears to occur because more recent nests contain shorter and noisier TR records that, at the hemisphere scale considered here, are somewhat detrimental to the overall skill of the reconstruction. However, as the N-TREND2015 network is also used for spatial field reconstruction (see Part II, Anchukaitis et al. *in prep*), removing the weaker shorter series would be deleterious for those locations where only short records exist. Overall, unfiltered calibration  $r^2$  values are  $> 0.40$  from 1200 – 1988 with values dropping below 0.30 before 918 and after 2004 (Figure 2F). These are better than those obtained either by DWJ06 (RCS version) which produced unfiltered  $r^2$  values  $> 0.30$  from 1556 – 1978 and were never  $> 0.40$  or by SCH2015, which also never exceeded 0.40 and demonstrated calibration  $r^2$  values  $> 0.30$  only from 1363 – 1976. When the data are smoothed with a 20-year cubic smoothing spline, N-TREND2015 calibration  $r^2$  values substantially increase, as expected, and are  $> 0.80$  from 910 – 1989, only dropping below 0.70

before 760 and after 2010 (Figure 2F). DWJ06 and SCH2015 explained  $> 0.70$  of the smoothed temperature variance between 1340 and 1995 and 881 and 1976, respectively. Neither record attains  $r^2$  values above 0.80 when the data are smoothed. Spectral coherence analysis (Appendix Figure B1b) clearly shows the stronger coherence at decadal and longer time-scales, but also highlights the improved high-mid frequency coherence in the N-TREND2015 data compared to DWJ06. SCH2015 explains a similar amount of mid-frequency variance but shows dampened coherence at longer time-scales.

No significant autocorrelation is noted in the N-TREND2015 regression model residuals for any nest except after 2006 where the proxy network is quite sparse and the calibration is already weak (unfiltered  $r^2 \leq 0.20$ ). No residual linear trends are found for any nest spanning the period 760 – 2010 (Figure 2G). With the relatively strong smoothed  $r^2$  values (Figure 2F) and generally robust residuals for most nests, the decadal and longer scale fidelity in the N-TREND2015 reconstruction is a significant improvement over previous studies back to the 10<sup>th</sup> century. However, prior to 1880, there is a distinct divergence (Figure 2D+E) between the instrumental data (warmer) and N-TREND2015 reconstruction (cooler) potentially reflecting the previously discussed “warm bias” in early instrumental summer temperature series (Parker 1994; Frank et al. 2007b; Böhm et al. 2010). Also, there is a hint of under-prediction of temperature estimates in the recent two decades which is likely mainly related to the decreasing number of TR records through this period (Figure 2A).

Validation metrics are presented in Figures 2H-J (individual period validation results detailed in Appendix Table B1). Results are rather variable depending on the validation period used. Validation  $r^2$  values are greater than zero for all nests, but are strongest for the late 1935-1988 period and weakest for the mid 1916-1952 period. Maximum  $r^2$  values are attained from about the 12<sup>th</sup> century to the recent period. While RE values are positive for all nests, CE values are positive only since the late 14<sup>th</sup> century. CE values first go below zero before 1373 using the 1935-1988 validation period when the Quebec MXD (QUEx) series drops out from the nesting process. CE is a particularly stringent validation metric and care is needed when using this metric alone to assess reconstruction quality (Cook et al. 1994). One further driver of the poor CE values is related to the reconstruction’s warmer values centred around 1950 versus the CRUTEM instrumental data showing peak mid-20<sup>th</sup> century warming in the 1940s (Figure 2E). This slight misfit appears to be related to both the spatial representativeness of the TR network and the fact that mean MJJA temperature is not the optimal season for some TR records (Table 1). Appendix Figure B3 compares an instrumental based reconstruction, which utilises the 0.5 latitude x 0.5 longitude CRUTS3.20 (Harris et al. 2014) gridded optimal calibrated seasonal instrumental series at the locations of the 54 TR series against the MJJA CRUTEM4 (40-75°N – Jones et al. 2012) hemisphere series. The instrumental based composite explains 78% of the CRUTEM4 variance (95% when smoothed), but places the mid-20<sup>th</sup> century warming peak in the 1950s and under-predicts the recent period similar to that observed for the N-TREND2015 reconstruction (Figure 2E). These observations suggest that despite the increased number of TR proxy records, substantially more than 54 instrumental records are needed to attain a complete representation of the large-scale decadal to high frequency climate signals (see also discussion in Diaz 1996, Jones and Briffa 1996 and Anchukaitis et al. *in prep*).

Compared to previous millennial length TR NH reconstructions (Esper et al. 2002; D’Arrigo et al. 2006; Frank et al. 2007a; Schneider et al. 2015) the N-TREND2015 reconstruction shows almost no late 20<sup>th</sup> century divergence (D’Arrigo et al. 2008 – see also discussion in Wilson et al. 2007) from the instrumental data (Figure 2D+E) and is statistically more robust for a longer period of time. In this respect, N-TREND2015 therefore represents a substantial update of these earlier studies. Fidelity weakens, as expected, when replication is low, especially over North America. Balancing calibration/validation metrics and site replication (Figure 1), we therefore define the period 918-2004, represented by at least 3 TR records for

each major continental mass, as a compromise period for “robust” comparison with other records although the comparative analyses below use mostly the full period.

### ***Comparison to other NH reconstructions***

N-TREND2015 shows a modest increase in temperatures from 750 to ~850 followed by a prolonged warm period until the end of the 11<sup>th</sup> century, a brief cooler 50-year period and then a short period of peak medieval warmth in the 1160s (Figure 3). 1161-1170 is the 3<sup>rd</sup> warmest decade in the reconstruction followed by 1946-1955 (2<sup>nd</sup>) and 1994-2003 (1<sup>st</sup> – see Table 2). It should be noted that these three decades cannot be statistically distinguished when uncertainty estimates are taken into account. Following 2003, 1168 is the 2<sup>nd</sup> warmest year, although caution is advised regarding the inter-annual fidelity of the reconstruction (Appendix Figure B1B). The period 1200-1800 is characterised by substantial multi-decadal variability with a range of about 1° Celsius between warm peaks and cold troughs. The two coldest decades, 1812-1821 and 1832-1841 (Table 2), are followed by a near continuous warming to present.

Figure 3 compares N-TREND2015 with DWJ06 (D’Arrigo et al. 2006) and SCH2015 (Schneider et al. 2015). Prior to 2015, DWJ06, using 19 predominantly RW records, was the most comprehensive TR specific millennial length NH reconstruction. SCH2015 is a new NH summer temperature reconstruction based on 16 MXD records. N-TREND2015 is not independent from these two studies. 3 records (ICE, TAY and YAK) were used from DWJ06 and 5 from SCH2015 (QUEx, TYR, POLx, MAN, and ALT). N-TREND2015 also shares common data with SCH2015 for some sites (e.g. ICE, EFmean, ALPS), but has used the original published versions rather than the SCH2015 processed versions. DWJ06 and SCH2015 can be seen as two end members representing quite different TR parameters - both reflecting predominantly summer temperatures but each expressing different spectral properties and potential biases. RW generally demonstrates strong biological persistence and higher autocorrelation (Helama et al. 2009b; Esper et al. 2015) while MXD has lower temporal autocorrelation. Although MXD has been shown to capture secular scale summer temperature trends (Luckman and Wilson 2005; Esper et al. 2012) it may also be influenced by light availability (Stine and Huybers 2014). Figure 1 clearly shows the superiority of MXD (and BI) based TR measurements for modelling inter-annual local scale summer temperatures. However, there is still uncertainty as to the potential spectral biases of using these different TR parameters to capture inter-annual, decadal and longer-scale climate.

DWJ06 shows similar warm and cool periods to N-TREND2015, particularly after 1400. D’Arrigo et al. (2006) advised caution when using the pre-1400 period of the reconstruction due to the low number of TR records (10 in 1300; 6 in 1000). N-TREND2015 should therefore represent a considerable improvement in this period due to the greater number of records used (32 in 1300; 23 in 1000). In fact, the N-TREND2015 reconstruction is substantially different before 1400 with it showing a more prolonged warming from the 9<sup>th</sup> to 12<sup>th</sup> centuries compared to the short-term warmth from 950-1100 in DWJ06. SCH2015 displays more modest warmth during the 850-1050 period with a peak at the end of 9<sup>th</sup> century and overall muted multi-decadal variability until 1300. From 1300-1570, SCH2015 shows warmer conditions than N-TREND2015 although overall, multi-decadal trends are similar. The decadal cooling of the 1450s is however more pronounced in N-TREND2015 than DWJ06 and SCH2015. The three records are quite similar after 1600 with similar short term cool periods coinciding with known volcanic events. Only N-TREND2015 expresses the post mid-20<sup>th</sup> century warming trend. N-TREND2015 correlates (918-1999 – unfiltered (filtered)) marginally stronger with DWJ06 ( $r = 0.65$  (0.73)) than with SCH2015 ( $r = 0.61$ (0.61)), but for running shorter 50-year periods, N-TREND2015 and SCH2015 are clearly more coherent (Figure 3 – lower panel). These correlations highlight the greater coincidence in multi-decadal and longer term signals between N-TREND2015 and

DWJ06, and the greater high-frequency coherence with SCH2015 – the latter’s inter-annual variability driven by the stronger high-frequency signal within the MXD data used.

Figure 4 compares the aforementioned TR composites (plus Frank et al. 2007a – FRK07) with recently published multi-proxy (MP) Northern Hemisphere temperature reconstructions (Amman and Wahl 2007; Hegerl et al. 2007; Ljungqvist 2010; Mann et al. 2008) and their mean. Overall, N-TREND2015 correlates strongest ( $r = 0.76$ ) with the MP mean series although only marginally better than DWJ06 ( $r = 0.72$ ) and substantially better than SCH2015 ( $r = 0.49$ ) and FRK07 ( $r = 0.48$ ). N-TREND2015 and the MP composite show common warm conditions from the mid-9<sup>th</sup> to the end of the 11<sup>th</sup> centuries although the MP series suggests a warmer medieval period until the mid-12<sup>th</sup> century. When divided into four longitudinal quadrants (Figure 1), this protracted period of medieval warming appears to be spatially focused over eastern North America (although in that quadrant the medieval period is expressed by only a single record from Quebec – Table 1) and Western Eurasia and extends to the end of the 12<sup>th</sup> century (Appendix C2) suggesting a common dynamical link with North Atlantic variability, the thermohaline circulation/meridional overturning circulation and purported prolonged positive North Atlantic Oscillation anomalies through this period (Trouet et al. 2009; Helama et al. 2009a; Cunningham et al. 2013). Although a persistently positive NAO during the medieval period has recently been challenged by Ortega et al. (2015), the sharp rise in temperatures around 1150, which remains largely positive for around three centuries in eastern North America and western Eurasia coincides with the timing of an inferred positive phase of the NAO during winter (Trouet et al. 2009). Focusing on summer, Loader et al. (2013) suggested that the period 900-1200 was dominated by a northerly shift of the NH storm tracks, yielding warm summers in the high latitudes of NW Europe. This warming is also reflected in an extended period of warm sea surface temperatures in the North Atlantic (Cunningham et al. 2013), associated with a positive phase of the Atlantic Multidecadal Oscillation (AMO) during the medieval period, as suggested by Oglesby et al. (2012), leading to a decrease in the mid-latitude – Arctic temperature gradient.

From ~1100 to ~1700, N-TREND shows a prolonged period of multi-decadal variability embedded within a general cooling trend. Common cool conditions are observed in all the proxy records until the mid-19<sup>th</sup> century after which temperatures rise until present. As noted above, SCH2015 expresses muted medieval warmth, shows warmer conditions during the late 14<sup>th</sup> and early 15<sup>th</sup> centuries, and does not capture recent warming well. FRK07 has a shorter period of warmth during the medieval period (similar to DWJ06), but is substantially cooler in the 13<sup>th</sup> century.

Overall, comparing the TR and MP reconstructions provides a broadly coherent picture of warm conditions around the end of the 10<sup>th</sup> century, a cooling trend for about 500 years, a prolonged cool period from ca. 1450-1850 and steep warming until present. However, there are subtle differences between the reconstructions that likely reflect (1) reconstruction composition (i.e. older studies used less data; including non-annual proxy archives with associated dating uncertainties), (2) varying spectral properties of different proxy types (including tree-ring RW and MXD in comparison with MP networks) and (3) the different reconstruction’s seasonal target (summer vs. annual). The issue of seasonality is particularly important for understanding the stochastic nature of climate. The current debate over the so-called “temperature hiatus” (Meehl et al. 2011; Cohen et al. 2012; Kosaka and Xie 2013; Trenberth and Fasullo 2013; England et al. 2014; Karl et al. 2015) exemplifies the consequence of mixing proxy records with varying seasonal sensitivities. The “hiatus” is observed in large scale winter temperatures and no slowdown is noted in NH summer temperatures (Cohen et al. 2012). Differences in timing and duration of warm and cold periods between TR and MP reconstructions could well be related to the same issue – highlighting the need for the palaeoclimate community to carefully evaluate the seasonal signal embedded in their proxies.

### ***Proxy/model Comparison – long term context***

A palaeoclimate aspiration is that independently derived reconstructions and models of past climate should express similar trends (where externally forced) and variance structure through time. Agreement in hindcast modelling ability will provide some confidence in scenario forecasts of the future. In comparison to the challenging regional model-data assessments (Zorita et al. 2010, Deser et al. 2012), hemispheric scale comparison is generally more straightforward: heterogeneous local/regional internally-generated variability is often dampened through averaging, allowing the influence of external forcing factors such as volcanic events and anthropogenic emissions to be addressed (Fernández-Donado et al. 2013, but see Anchukaitis et al. 2012). The strong coherence between the North American and Eurasian TR composites (Appendix Figure C3) indicates the dominance of large scale forcing on such proxy records.

Figure 5 compares the TR NH reconstructions with 10 CMIP5 coupled ocean-atmosphere global climate models (See Appendix Table A2 for information) for land only MJJA temperatures over the same 40-75°N spatial domain. The 20-year low-pass smoothed time-series are shown as anomalies with respect to 1400-1850. Overall, N-TREND2015 sits well within the range of the model ensemble back to ~1200 (Figure 5A). Very good agreement is observed from ~1500 to the early 19<sup>th</sup> century, although this partially reflects the common reference period for anomaly calculations. Prior to ~1500, decadal variance of the reconstruction is generally greater than the models with the 1430s and 1160s temperature estimates well above the model range. Warming from the 19<sup>th</sup> to 20<sup>th</sup> century is in general good agreement with most models although N-TREND2015 appears to show more prolonged cooler temperatures in the 19<sup>th</sup> century before the start of warming.

In two important periods of climate transition, the cooling from the medieval period (900-1250) to Little Ice Age (LIA - 1450-1850), and the post LIA warming to present, the TR reconstructions and model estimations differ. During the medieval to LIA transition (Figure 5C), the amplitude of temperature change from warm to cool, as expressed by N-TREND2015, is greater than most models and other TR reconstructions. The only model that agrees with this reconstruction is CCSM4. Most model results produce a cooler medieval period relative to the 1450-1850 mean, more in-line with DWJ06 and SCH2015. However, for the LIA to 20<sup>th</sup> century transition (Figure 5D) N-TREND2015 agrees well with the models - except with CCSM4 and HADCM3 - which is likely due to differences in these models' response to anthropogenic aerosols (see Meehl et al. 2012 and Schurer et al. 2014). Overall, N-TREND2015, DWJ06 and SCH2015 share reasonable multi-decadal coherence ( $r = 0.58, 0.44$  and  $0.64$ , respectively) with the multi-model mean from 1251-1850 CE (inset table in Figure 5B) but correlations are substantially weaker for 850-1250 CE.

Comparing the spectral properties (Figure 5E) of the three TR reconstructions to the 10 CMIP5 models allows a general assessment of potential frequency biases. SCH2015 (using only MXD data) contains more high frequency variance compared to DWJ06 (using mostly RW data) with N-TREND2015 sitting in the middle. Both N-TREND2015 and SCH2015 sit within the spectral power range of the models at these high frequencies despite the fact that they show only modest coherence at high frequencies with NH instrumental temperatures (Appendix Figure B2). Only by the addition of more long MXD datasets from key areas currently lacking such high resolution archives (e.g. much of North America, eastern Eurasia) could the large-scale high frequency coherence be improved. At mid and lower frequencies, the spectral density of all three TR reconstructions sits within the model range.

Most attribution studies, and many proxy/model comparative studies (Hegerl et al. 2007; Mann et al. 2012), have focused on the post 1400 period, when relatively ample and well dated proxy data are available. Prior to that period there is a distinct mismatch between the proxies and models during the medieval period (Fernández-Donado et al. 2013; Schurer et al. 2014). Although the response to forcing is

detectable in some reconstructions, there is a tendency for less prominent forcing at this time which makes detection more difficult (Schurer et al., 2013). Since N-TREND2015 contains more data during the medieval period than all previous TR reconstructions, it should provide more robust estimates of medieval summer season temperature. The relatively simple comparison shown in Figure 5 suggests medieval NH temperatures were likely warmer than those simulated by the CMIP5 models, although a detailed evaluation of this observation is beyond the scope of the current work. In a recent appraisal of proxy and CMIP5 model output over the Last Glacial Maximum, mid-Holocene, and the recent period, a substantial disagreement between palaeoclimate and modelled estimates was reported, although the models were marginally better at representing modern late Holocene climate (Harrison et al. 2015). However, underlying this comparison is the quality of the proxy data used and although N-TREND2015 represents an improvement on previous studies, more data from key regions such as North America and eastern Eurasia are needed.

### ***Proxy/model Comparison – volcanic forcing***

Coherency analysis (Appendix Figure B1) shows that large-scale TR composite reconstructions are generally poor at representing mean hemisphere-scale inter-annual variability, mainly because of the relatively small numbers of TR proxy series used for the hemispheric composites (see discussion in Diaz 1996 and Jones and Briffa 1996) and the biological memory effects in RW data (Esper et al. 2015). However, after major volcanic events, short term cooling can be much more spatially homogenous, allowing even relatively small networks of TR records to potentially capture the amplitude of cooling after such events (Hegerl et al. 2003; Schneider et al. 2015; Stoffel et al. 2015, but see Anchukaitis et al. 2012). Though it has long been known that TR records can reflect volcanically forced cool summer conditions (Briffa et al. 1998), it is less well known that MXD is the superior TR parameter for assessing such short term climate perturbations (Jones et al. 1995; Frank et al. 2007b; Anchukaitis et al. 2012; D'Arrigo et al. 2013; Esper et al. 2015). We briefly return to this issue by comparing the response of N-TREND2015, DWJ06, SCH2015 and the CMIP5 models to selected volcanic events. Esper et al. (2013) showed that the choice of candidate volcanic events can greatly influence the observed estimated amount of cooling experienced on average over large regions. Herein, we use three different lists of significant volcanic events (Table 3 – combined information from Gao et al. (2008) and Crowley and Unterman (2013); Esper et al. 2013; Sigl et al. 2013) to assess the mean response of each TR NH series and CMIP5 model to large volcanic events using superposed epoch analysis (SEA). It should be noted that the CMIP5 models were forced with either the Gao et al. (2008) or Crowley and Unterman (2013) volcanic forcing series (Appendix Table A2).

Overall the mean TR proxy response to Esper events (minimal cooling in year T+1) is substantially muted compared to those of Gao/Crowley and Sigl (Figure 6). A muted response may occur if some event dates are in error, conversely, an amplified response can occur if event dates are related, in a circular manner, to tree-ring evidence (Esper et al. 2013). The latter condition may in fact be applicable to some ice core records (e.g. the Gao/Crowley), which would partly explain the stronger response observed using such lists. We concede the amplification effect may be limited to just a few events, such as the 1450s eruption(s), which are not firmly dated by any evidence (Plummer et al. 2012) other than tree-rings (Gao et al. 2006). This, however, is not the case for Sigl et al. (2013). Using the Gao/Crowley and Sigl lists, N-TREND2015 and SCH2015 express a maximum cooling in year T+1 of ~0.4-0.5 °C while DWJ06 does not reach peak cooling until years T+5 or T+6 and with a more modest cooling of ~0.3 °C. This lagged and muted response of the RW dominated DWJ06 reconstruction has been detailed elsewhere and is mainly related to autocorrelated biological memory in RW data (Krakauer and Randerson, 2003; Frank et al., 2007b;

Anchukaitis et al. 2012; Esper et al. 2015). As N-TREND2015 includes RW as well as MXD data, there is potential for this biological persistence bias to influence the record's ability to express true large-scale temperature response (and recovery) to volcanic events. The lower panel of Figure 6 compares the mean response of the three reconstructions to the 10 CMIP5 simulations. For the Gao/Crowley and Sigl events, the models on average show marginally greater cooling ( $\sim 0.6\text{--}0.7\text{ }^{\circ}\text{C}$ ) in year T+1 than N-TREND2015 and SCH2015, but similar amounts of cooling for the Esper ( $\sim 0.2\text{ }^{\circ}\text{C}$ ) events. Most importantly, the rate of recovery after volcanic cooling is slower for N-TREND2015 than SCH2015, with the latter record agreeing well with the models.

Dividing the N-TREND2015 data-set into its three TR parameter types RW, MXD and MIX (Figure 7) and producing a parameter specific hemispheric mean composite from each, highlights the behaviours mentioned above, although the MIX data displays a slower recovery than RW when using the Gao/Crowley and Sigl event lists. Overall, the response of each parameter composite appears specific to different eruptions (Table 3, Figure 7). In 1258 (1257, Samalas eruption – Table 3), the MXD response is immediate, RW data is strongest in year T+1 and in the MIX composite at year T+2. The Kuwae eruptions of 1452 (not identified as a significant eruption event – Table 3) and 1458 are clearly captured by MXD data in 1453, and although cooling starts in 1453, peak cold conditions in the MIX series does not occur until 1459. All three parameter types start decreasing after the unknown 1809 event, but only MXD and MIX contain a defined cool peak in 1816. In 1883 (Krakatau), MXD and MIX respond similarly, with cold conditions in 1884, while in 1912 (Katmai), the response is immediate in MXD and RW, but lagged by 1 year in MIX. By just looking at the response to individual volcanic eruptions, it is difficult to ascertain whether there are true lagged biases in the MIX composite compared to MXD. A more detailed analysis is needed to examine the spatiotemporal response of the individual TR parameters, which is beyond the scope of this paper but is partly addressed in Anchukaitis et al. (*in prep*).

In a recently published new TR reconstruction of NH JJA mean temperatures, Stoffel et al. (2015) utilised 30 regional records (using both RW and MXD) to assess large-scale volcanic response expressed in their composite record by comparison to a modified version of the IPSL-CM5A-LR model (Schmidt et al. 2012). Their model version simulated less cooling and they stated that with respect to volcanic forcing their work had reconciled previously reported dissimilarities between climate simulations and TR based reconstructions. Although the “missing ring” hypothesis (Mann et al. 2012) is now thoroughly rejected (Anchukaitis et al. 2012; D'Arrigo et al. 2013; Esper et al. 2015), Stoffel et al. (2015) and the results detailed herein, indicate that there are still substantial uncertainties that need to be addressed before we fully understand the temporal and spatial dynamical response of the climate system to volcanic forcing as well as the response of specific TR parameters to volcanically forced cool summers.

## Conclusion

Over the last 20 years, palaeoclimatology has augmented and improved the available proxy data that can be used to produce large-scale, high resolution, reconstructions of the Earth's climate for the late Holocene. This improvement is clearly shown in synthesis figures of NH temperatures in the IPCC reports, which have moved from a single qualitative representation (FAR – Houghton et al. 1990), to a quantitative ensemble (AR4/5 – Masson-Delmotte et al., 2013 - see discussion in Frank et al. 2010). It is quite understandable that when few data existed, there was no real option but to amalgamate all possible records together, ignoring the local/regional seasonal signal they represented, or their quality and individual biases. However, substantial investment and progress has been made over the past decade and many high quality records now exist from different proxy archives for multiple locations around the NH

(Jones et al. 2009; PAGES2 Consortium 2013). This data richness now allows the palaeoclimate community to focus more specifically on the unique properties of different proxy archives allowing selection for optimising seasonal and temporal fidelity.

This paper (and its part II companion, Anchukaitis et al. *in prep*), represents a tree-ring community initiative to improve the Northern Hemisphere-scale proxy network (temporally and spatially), and reconstruct NH summer temperatures back to the 8<sup>th</sup> century AD. Following recommendations detailed in Wilson et al. (2007), we have focused on using published tree-ring reconstructions which reflect a statistically and biologically reasonable amount of local/regional summer temperature variance, and have utilised a substantially increased data-base of records compared to previous studies (D'Arrigo et al. 2006; Frank et al. 2007a; Wilson et al. 2007; Schneider et al. 2015; Stoffel et al. 2015). Consequently, the increased replication in the N-TREND2015 reconstruction alone represents a substantial improvement over previous TR NH millennial reconstructions and the final record appears reasonably insensitive to the statistical methods used (Appendix Figure C1). The N-TREND2015 reconstruction expresses reasonable fidelity, at all time-scales, from 918-2004 (Figure 2) and reasonably captures late 20<sup>th</sup> century warming. Nevertheless, N-TREND2015 is far from perfect!

How well should a TR reconstruction calibrate locally before it can be considered for inclusion in such a large-scale data-set? Wilson et al. (2007) specifically used a minimum **correlation** of 0.4 (only 16% explained variance) against local gridded temperature data, but this value is still rather low, and for realistic reconstructions of local, regional, and hemisphere-scale temperatures, greater fidelity should be required, especially when the number of input series is modest. Of the records used herein, local scale calibration  $r^2$  values range from ~10-70% explained variance (Figure 1) with all RW based records explaining less than 40% variance. Considering as well, the well-known lagged high frequency biases of RW data due to biological persistence (Krakauer and Randerson, 2003; Frank et al., 2007b; Anchukaitis et al. 2012; Esper et al. 2015), it could be argued that RW derived reconstructions should not be deemed robust estimates of local temperatures without the inclusion of (or replacement with) MXD or BI data. This is a rather contentious statement but it is more defensible if the study focus is on addressing past climate response to volcanic forcing. At the very least, whether RW is used alone or in combination with MXD/BI data, local calibration needs to express “reasonable” fidelity and multiple statistical metrics exist to assess reconstruction quality (e.g.  $r^2$ , RE, CE and residual analysis – Cook et al. 1994; Wilson et al. 2006; Macias-Fauria et al. 2012). As a minimum, local climate calibration for any single TR parameter or combination of multiple TR parameters must be statistically significant with no significant long-term trends within the calibration residuals. Any TR record expressing local based divergence (D'Arrigo et al. 2008) should not be considered in a large composite database unless the cause of the divergence can be truly identified as unique to the recent period (Wilson et al. 2007). This situation can potentially be further confused if the quality and/or homogeneity of the instrumental data cannot be assumed to be consistent back in time (Parker 1994; Frank et al. 2007b). Correlations between some TR records and pre-1950 gridded data weaken considerably, especially for locations in central Asia (Table 1 and Cook et al. 2012), which make defining an acceptable calibration threshold difficult.

Although the number of TR records used in N-TREND2015 represents a substantial increase over previous TR NH composites (D'Arrigo et al. 2006; Frank et al. 2007a; Wilson et al. 2007; Schneider et al. 2015; Stoffel et al. 2015), there are still vast regions of the NH that are poorly represented (see Figure 1 and discussion in Anchukaitis et al. (*in prep*)). Only 3 millennial length temperature-sensitive TR series exist for North America and only one of those incorporates MXD. Very few long TR records have been developed for eastern Eurasia and no millennial length MXD series exist between 90°E and 170°E. Where forests grow in non-temperature limiting environments (e.g. mid latitude eastern North America, Siberian Steppes, low

elevation central Europe), it may never be possible to develop temperature-sensitive chronologies using traditional TR parameters, but stable isotopes may help in some regions (Young et al. 2012). Nevertheless, there are substantial swaths of high latitude boreal forest and high altitude environments where tree-growth is predominantly limited by summer temperatures, and a strategically focussed sample collection in these regions (including sub-fossil material to extend living datasets) is needed.

N-TREND2015 represents the latest iteration of tree-ring based NH reconstructions. It will not be the last, but we hope that the creation of the N-TREND consortium will facilitate further expansion of the NH tree-ring network. A strategic plan is needed to update existing records to the present and develop new long millennial records from regions where little or no data exist. It is also paramount that MXD or BI records are concurrently developed with RW data to produce locally, as well as regionally, robust temperature reconstructions. As the network expands, further assessment of the unique properties of RW, MXD and BI can be performed that should lead towards even further, substantial refinement of large temporal and spatial scale summer temperature reconstructions.

## **Acknowledgement**

The N-TREND consortium is not specifically funded, but many individuals acknowledge relevant funds; RW: UK Natural Environment Research Council (NERC - NE/K003097/1) and Leverhulme Trust project (F/00 268/BG); KJA: National Science Foundation - NSF AGS-[1501856](#) and NSF AGS-[1501834](#); KB and TO: NERC (NE/G018863/1); UB: Czech project 'Building up a multidisciplinary scientific team focused on drought' No. CZ.1.07/2.3.00/20.0248; RD: National Science Foundation Paleoclimatic Perspectives of Climatic Change; SH: Academy of Finland; GH and AS: ERC advanced grant TITAN (EC-320691), NCAS and GH specifically with a as Royal Society Wolfson Research Merit Award; HL: The Swedish Science Council (VR); 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 reconstructions can be found at: <https://ntrenddendro.wordpress.com/>. Lamont-Doherty Earth Observatory contribution No. XXXX.

## **References**

Ammann, C., Joos, F., Schimel, D., Otto-Bliesner, B., & Tomas, R., 2007. Solar influence on climate during the past millennium: results from transient simulations with the NCAR Climate System Model. *Proc. Natl. Acad. Sci. U.S.A.*, 104(10), 3713–8.

Anchukaitis, K., D'Arrigo, R., Andreu-Hayles, L., Frank, D., Verstege, A., Buckley, B., Curtis, A., Jacoby, G., Cook, E., 2013. Tree-ring reconstructed summer temperatures from northwestern North America during the last nine centuries. *Journal of Climate* 26, 10, 3001–3012, doi: 10.1175/JCLI-D-11-00139.1

Anchukaitis, K., Breitenmoser, P., Briffa, K., Buchwal, A., Büntgen, U., Cook, E., D'Arrigo, R., Esper, J., Evans, M., Frank, D., Grudd, H., Gunnarson, B., Hughes, M., Kirilyanov, A., Körner, C., Krusic, P., Luckman, B., Melvin, T., Salzer, M., Shashkin, A., Timmreck, C., Vaganov, E., Wilson, R., 2012. Tree rings and volcanic cooling. *Nature Geoscience* 5, 836– 837, doi:10.1038/ngeo1645

Anchukaitis, K., Wilson, R., Briffa, K., Büntgen, U., Cook, E., D'Arrigo, R., Davi, N., Esper, J., Frank, D., Gunnarson, B., Hegerl, G., Helama, S., Klesse, S., Krusic, P., Linderholm, H., Myglan, V., Osborn, T., Rydval, M., Schneider, L., Schurer, A., Wiles, G., Zhang, P. and Zorita, E. Last millennium Northern Hemisphere summer temperatures from tree rings: Part I: the spatial context. *In Preparation. To be submitted to Quaternary Science Reviews.*

Biondi, F., Perkins, D., Cayan, D., Hughes, M. 1999. July temperature during the second millennium reconstructed from Idaho tree rings. *Geophysical Research Letter* 26, 1445-1448.

Björklund, J., Gunnarson, B., Seftigen, K., Esper, J., Linderholm, H., 2014. Blue intensity and density from northern Fennoscandian tree rings, exploring the potential to improve summer temperature reconstructions with earlywood information. *Climate of the Past* 10, 877-885, doi: 10.5194/cp-10-877-2014

Böhm, R., Jones, P., Hiebl, J., Frank, D., Brunetti, M., Maugeri, M., 2010. The early instrumental warm-bias: a solution for long central European temperature series 1760 – 2007. *Climatic Change* 101(1-2), 41-67.

Briffa, K., Jones, P., Schweingruber, F., Karlen, W., Shiyatov, G., 1996. Tree ring variables as proxy climate indicators: problems with low-frequency signals, in: Jones, P.D., Bradley, R.S., Jouzel, J. (Eds.), *Climatic Variations and Forcing Mechanisms of the Last 2000 Years*. Springer-Verlag, Berlin, pp. 9-41.

Briffa, K., Jones, P., Schweingruber, F., Osborn, T., 1998. Influence of volcanic eruptions on Northern Hemisphere summer temperature over the past 600 years. *Nature* 393, 450–455, doi:10.1038/30943

Briffa, K., Osborn, T., Schweingruber, F., Harris, I., Jones, P., Shiyatov, S., Vaganov, E. 2001. Low frequency temperature variations from a northern tree-ring density network. *Journal of Geophysical Research* 106(D3), 2929–2941.

Briffa, K., Osborn, T., Schweingruber, F., Jones, P., Shiyatov, S., Vaganov, E., 2002a. Tree-ring width and density data around the Northern Hemisphere: Part 1, Local and regional climate signals. *The Holocene* 12(6), 737–757.

Briffa, K., Osborn, T., Schweingruber, F., Jones, P., Shiyatov, S., Vaganov, E., 2002b. Tree-ring width and density data around the Northern Hemisphere: Part 2, spatio-temporal variability and associated climate patterns. *The Holocene* 12(6), 759-789.

Briffa, K., Melvin, T., Osborn, T., Hantemirov, R., Kirilyanov, A., Mazepa, V., Shiyatov, S., Esper, J. 2013. Reassessing the evidence for tree-growth and inferred temperature change during the Common Era in Yamalia, northwest Siberia. *Quaternary Science Reviews* 72, 83-107. doi: 10.1016/j.quascirev.2013.04.008

Büntgen, U., Frank, D.C., Nievergelt, D., Esper, J., 2006. Summer temperature variations in the European Alps, AD 755-2004. *Journal of Climate* 19(21), 5606-5623.

Büntgen, U., Kyncl, T., Ginzler, C., Jaks, D.S., Esper, J., Tegel, W., Heussner, K.U., Kyncl, J., 2013. Filling the Eastern European gap in millennium-long temperature reconstructions. *Proceedings of the National Academy of Science, USA* 110, 1773-1778.

- Campbell, R., McCarroll, D., Robertson, I., Loader, N.J., Grudd, H., Gunnarson, B.E., 2011. Blue intensity in *Pinus sylvestris* tree rings: A manual for a new palaeoclimate proxy. *Tree-Ring Research* 67(2), 127–134.
- Cohen, J., Furtado, J., Barlow, M., Alexeev, V., Cherry, J. 2012. Asymmetric seasonal temperature trends. *Geophysical Research Letters* 39(4), doi: 10.1029/2011GL050582
- Cook, E. , Peters, K., 1981. The smoothing spline: A new approach to standardizing forest interior tree-ring width series for dendroclimatic studies. *Tree-Ring Bulletin* 41, 45–53.
- Cook, E., Kairiukstis, L. 1990. *Methods of Dendrochronology*, Springer, New York.
- Cook, E., Briffa, K., Jones, P. 1994. Spatial regression methods in dendroclimatology: a review and comparison of two techniques. *International Journal of Climatology* 14(4), 379-402.
- Cook, E., Briffa, K., Meko, D., Graybill, D., Funkhouser, G., 1995. The 'segment length curse' in long tree-ring chronology development for palaeoclimatic studies. *The Holocene* 5, 229–237.
- Cook, E., D'Arrigo, R., Mann, M., 2002. A well-verified, multiproxy reconstruction of the winter North Atlantic Oscillation index since AD 1400. *Journal of Climate* 15, 1754–1764.
- Cook, E., Krusic, P., Anchukaitis, K., Buckley, B., Nakatsuka, T., Sano, M., PAGES Asia2k Members, 2012. Tree-ring reconstructed boreal summer temperature anomalies for temperate East Asia since 800 C.E.. *Climate Dynamics* 41(11-12), 2957-2972, doi: 10.1007/s00382-012-1611-x
- Crowley, T. 2000. Causes of Climate Change Over the Past 1000 Years. *Science*, 289(5477), 270–277. <http://doi.org/10.1126/science.289.5477.270>
- Crowley, T.J., Unterman, M.B., 2013. Technical details concerning development of a 1200 yr proxy index for global volcanism. *Earth System Science Data* 5, 187–197, doi:10.5194/essd-5-187-2013
- Cunningham, L., Austin, W., Knudsen, K., Eiriksson, J., Scourse, J., Wanamaker Jr., A., Butler, P., Cage, A., Richter, T., Husum, K., Hald, M., Andersson, C., Zorita, E., Linderholm, H.W., Gunnarson, B.E., Sicre, M.-A., Sejruo, H.P., Jiang, H., Wilson, R., 2013. Reconstructions of surface ocean conditions from the northeast Atlantic and Nordic Seas during the last millennium. *The Holocene* 23, 921–935, doi:10.1177/0959683613479677
- D'Arrigo, R., Buckley, B., Kaplan, S., Woollett, J., 2003. Interannual to multidecadal modes of Labrador climate variability inferred from tree rings. *Climate Dynamics* 20(2-3), 219–228.
- D'Arrigo, R., Mashig, E., Frank, D., Jacoby, G., Wilson, R., 2004. Reconstructed Warm Season Temperatures For Nome, Seward Peninsula, Alaska. *Geophysical Research Letters* 31(9), doi: 10.1029/2004GL019756
- D'Arrigo, R., Wilson, R., Jacoby, G., 2006. On the long-term context for late 20th century warming. *Journal of Geophysical Research* (1984-2012) 111(D3), doi: 10.1029/2005JD006352

- D'Arrigo, R., Wilson, R., Liepert, B., Cherubini, P., 2008. On the 'Divergence Problem' in Northern Forests: A Review of the Tree-Ring Evidence and Possible Causes. *Global and Planetary Change* 60, 289–305, doi: 10.1016/j.gloplacha.2007.03.004
- D'Arrigo, R., Wilson, R., Anchukaitis, K., 2013. Volcanic Cooling Signal in Tree-Ring Temperature Records for the Past Millennium. *Journal of Geophysical Research: Atmospheres* 118(16), 9000-9010, doi: 10.1002/jgrd.50692
- D'Arrigo, R., Wilson, R., Wiles, G., Anchukaitis, K., Solomina, O., Davi, N., Deser, C., Dolgova, E., 2014. Tree-ring reconstructed temperature index for coastal northern Japan: implications for western North Pacific variability. *International Journal of Climatology*, doi: 10.1002/joc.4230
- Davi, N., Jacoby, G., Wiles, G., 2003. Boreal temperature variability inferred from maximum latewood density and tree-ring width data, Wrangell Mountain region, Alaska. *Quaternary Research* 60, 252–262.
- Davi, N., D'Arrigo, R., Jacoby, G., Cook, E.R., Anchukaitis, K., Nachin, B., Rao, M.P., Leland, C., 2015. A long-term context (931–2005 C.E.) for rapid warming over Central Asia. *Quaternary Science Reviews* 121, 89-97, doi: 10.1016/j.quascirev.2015.05.020
- Deser, C., Knutti, R., Solomon, S., Phillips, A.S., 2012. Communication of the role of natural variability in future North American climate. *Nature Climate Change* 2, 775–779.
- Deslauriers, A., Anfodillo, T., Rossi, S., Carraro, V., 2007. Using simple causal modeling to understand how water and temperature affect daily stem radial variation in trees. *Tree Physiology* 27(8), 1125–1136.
- Diaz, H. 1996. Temperature changes on long time and large spatial scales: inferences from instrumental and proxy records, in: Jones, P.D., Bradley, R.S., Jouzel, J. (Eds.), *Climate variations and forcing mechanisms of the last 2,000 years*. Springer Berlin Heidelberg, pp 585–601.
- Divine, D., Isaksson, E., Martma, T., Meijer, H.A.J., Moore, J., Pohjola, V., van de Wal, R.S.W., Godtliebsen, F., 2011. Thousand years of winter surface air temperature variations in Svalbard and northern Norway reconstructed from ice core data. *Polar Research* 30, doi: 10.3402/polar.v30i0.7379
- Dorado-Linan, I., Büntgen, U., Gonzalez-Rouco, F., Zorita, E., Montavez, J., Gomez-Navarro, J., Brunet, M., Heinrich, I., Helle, G., Gutierrez, E., 2012. Estimating 750 years of temperature variations and uncertainties in the Pyrenees by tree-ring reconstructions and climate simulations. *Climate of the Past* 8, 919-933.
- Edwards, T.L., Crucifix, M., Harrison, S.P., 2007. Using the past to constrain the future: how the palaeorecord can improve estimates of global warming. *Progress in Physical Geography* 31 (5), 481-500.
- England, M., McGregor, S., Spence, P., Meehl, G., Timmermann, A., Cai, W., Gupta, A., McPhaden, M., Purich, A., Santoso, A., 2014. Recent intensification of wind-driven circulation in the Pacific and the ongoing warming hiatus. *Nature Climate Change* 4, 222–227, doi: 10.1038/nclimate2106
- Esper, J., Cook, E.R., Schweingruber, F., 2002. Low-frequency signals in long tree-ring chronologies for reconstructing past temperature variability. *Science* 295, 2250-2253.

- Esper, J., Cook, E., Peters, K., Krusic, P., Schweingruber, F., 2003. Detecting low-frequency tree-ring trends by the RCS method. *Tree-Ring Research* 59, 81–98.
- Esper, J., Frank, D., Wilson, R., 2004. Climate Reconstructions: Low-Frequency Ambition and High-Frequency Ratification. *Eos, Transactions American Geophysical Union* 85(12), 113-120.
- Esper, J., Frank, D.C., Büntgen, U., Verstege, A., Luterbacher, J., Xoplaki, E., 2007. Long-term drought severity variations in Morocco. *Geophysical Research Letters* 34, doi: 10.1029/2007GL030844
- Esper, J., Schneider, L., Krusic, P.J., Luterbacher, J., Büntgen, U., Timonen, M., Sirocko, F., Zorita, E., 2013. European summer temperature response to annually dated volcanic eruptions over the past nine centuries. *Bulletin of Volcanology* 75(7), 1-14, doi: 10.1007/s00445-013-0736-z
- Esper, J., Dühorn, E., Krusic, P., Timonen, M., Büntgen, U., 2014. Northern European summer temperature variations over the Common Era from integrated tree-ring density records. *Journal of Quaternary Science* 29, 487-494.
- Esper, J., Schneider, L., Smerdon, J., Schöne, B., Büntgen, U., 2015. Signals and memory in tree-ring width and density data. *Dendrochronologia*, 35, 62-70.
- Fernández-Donado, L., González-Rouco, J.F., Raible, C.C., Ammann, C.M., Barriopedro, D., García-Bustamante, E., Jungclauss, J.H., Lorenz, S.J., Luterbacher, J., Phipps, S.J., Servonnat, J., Swingedouw, D., Tett, S.F.B., Wagner, S., Yiou, P., Zorita, E., 2013. Large-scale temperature response to external forcing in simulations and reconstructions of the last millennium. *Climate of the Past* 9, 393–421, doi: 10.5194/cp-9-393-2013
- Frank, D., Esper, J., Cook, E. 2007a. Adjustment for proxy number and coherence in a large-scale temperature reconstruction. *Geophysical Research Letters* 34, doi:10.1029/2007GL030571
- Frank, D., Büntgen, U., Böhm, R., Maugeri, M., Esper, J., 2007b. Warmer early instrumental measurements versus colder reconstructed temperatures: shooting at a moving target. *Quaternary Science Reviews* 26(25), 3298-3310
- Frank, D., Esper, J., Zorita, E., Wilson, R., 2010. A noodle, hockey stick, and spaghetti plate: a perspective on high resolution paleoclimatology. *Wiley Interdisciplinary Reviews Climate Change* 1(4), 507-516, doi: 10.1002/wcc.53
- Franke, J., Frank, D., Raible, C., Esper, J., 2013. Spectral biases in tree-ring climate proxies. *Nature Climate Change* 3, 360-364.
- Fritts, H.C., 1976. *Tree Rings and Climate*. Academic Press, San Diego, Calif..
- Gao, C., A. Robock, S. Self, J. B. Witter, J. P. Steffenson, H. B. Clausen, M.-L. Siggaard-Andersen, S. Johnsen, P. A. Mayewski, and C. Ammann. 2006. The 1452 or 1453 A.D. Kuwae eruption signal derived from multiple ice core records: Greatest volcanic sulfate event of the past 700 years, *Journal of Geophysical Research*, 111, D12107, doi:10.1029/2005JD006710

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, doi: 10.1029/2008JD010239

Guo, Z., and Zhou, T., 2013. Why does FGOALS-gl reproduce a weak Medieval Warm Period but a reasonable Little Ice Age and 20th century warming? *Adv. Atmos. Sci.*, 30(6), 1758–1770. <http://doi.org/10.1007/s00376-013-2227-8>

Gennaretti, F., Arseneault, D., Nicault, A., Perreault, L., Bégin, Y., 2014. Volcano-induced regime shifts in millennial tree-ring chronologies from Northeastern North America. *Proceedings of the National Academy of Sciences, USA* 111, 10077–10082.

Gouirand, I., Linderholm, H., Moberg, A., Wohlfarth, B., 2008. On the spatio-temporal characteristics of Fennoscandian tree-ring based summer temperature reconstructions. *Theoretical and Applied Climatology* 91, doi: 10.1007/s00704-007-0311-7

Hansen, J., Ruedy, R., Sato, M., Lo, K., 2010. Global surface temperature change. *Reviews of Geophysics* 48(4), RG4004, doi:10.1029/2010RG000345

Harris, I., Jones, P., Osborne, T., Lister, D., 2014. Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset. *International Journal of Climatology* 34(3), 623–642, doi: 10.1002/joc.3711

Harrison, S., Bartlein, P., Izumi, K., Li, G., Annan, J., Hargreaves, J., Braconnot, P. & Kageyama, M. (2015). Evaluation of CMIP5 palaeo-simulations to improve climate projections. *Nature Climate Change*, 5(8), 735–743.

Hegerl G., Crowley T., Baum S., Yul Kim K, Hyde W., 2003. Detection of volcanic, solar and greenhouse gas signals in paleo-reconstructions of Northern Hemispheric temperature. *Geophysical Research Letters* 30:461–464

Hegerl, G., Crowley, T., Hyde, W., Frame, D., 2006. Climate sensitivity constrained by temperature reconstructions over the past seven centuries. *Nature* 440, 1029–1032, doi: 10.1038/Nature 04679

Hegerl, G., Crowley, T., Allen, M., Hyde, W., Pollack, H., Smerdon, J., Zorita, E., 2007. Detection of human influence on a new, validated 1500-year temperature reconstruction. *Journal of Climate* 20(4), 650–666.

Helama, S., Timonen, M., Holopainen, J., Ogurtsov, M., Mielikäinen, K., Eronen, M., Lindholm, M., Meriläinen, J., 2009a. Summer temperature variations in Lapland during the Medieval Warm Period and the Little Ice Age relative to natural instability of thermohaline circulation on multi-decadal and multi-centennial scales. *Journal of Quaternary Science* 24, 450–456.

Helama, S., Makarenko, N.G., Karimova, L.M., Kruglun, O.A., Timonen, M., Holopainen, J., Meriläinen, J., Eronen, M., 2009b. Dendroclimatic transfer functions revisited: Little Ice Age and Medieval Warm Period summer temperatures reconstructed using artificial neural networks and linear algorithms. *Annales Geophysicae* 27 (3), 1097–1111, doi: 10.5194/angeo-27-1097-2009

Helama, S., Vartiainen, M., Holopainen, J., Mäkelä, H.M., Kolström, T., Meriläinen, J., 2014. A palaeotemperature record for the Finnish Lakeland based on microdensitometric variations in tree rings. *Geochronometria* 41(3), 265-277.

Hughes, M., Vaganov, E., Shiyatov, S., Touchan, R., Funkhouser, G., 1999. Twentieth-century summer warmth in northern Yakutia in a 600 year context. *The Holocene* 9, 603– 608.

Hughes, M., 2011. Dendroclimatology in High-Resolution Paleoclimatology, in: Hughes, M.K., Swetnam, T.W., Diaz, H.F. (Eds.), *Dendroclimatology*. Springer, Netherlands, pp. 17-34.

Jacoby, G., Lovelius, N., Shumilov, O., Raspopov, O., Karbainov, J., Frank, D., 2000. Long-term temperature trends and tree growth in the Taymir region of northern Siberia. *Quaternary Research* 53(3), 312– 318.

Jones, P., Briffa, K., Schweingruber, F., 1995. Tree-ring evidence of the widespread effects of explosive volcanic eruptions. *Geophysical Research Letters* 22, 1333–1336.

Jones, P., Briffa, K., 1996. What can the instrumental record tell us about longer timescale paleoclimatic reconstructions? in: Jones, P.D., Bradley, R.S., Jouzel, J. (Eds.). *Climatic Variations and Forcing Mechanisms of the Last 2000 Years*. Springer, Berlin, pp. 625–644.

Jones, P., Briffa, K., Barnett, T., Tett, S., 1998. High-resolution paleoclimatic records for the last millennium: Interpretation, integration and comparison with general circulation model control-run temperatures. *The Holocene*, 8(4), 455– 471.

Jones, P., Briffa, K., Osborn, T., Lough, J., van Ommen, T., Vinther, B., Luterbacher, J., Wahl, E., Zwiers, F., Mann, M., Schmidt, G., Ammann, C., Buckley, B., Cobb, K., Esper, J., Goosse, H., Graham, N., Jansen, E., Kiefer, T., Kull, C., Küttel, M., MosleyThompson, E., Overpeck, J., Riedwyl, N., Schulz, M., Tudhope, A., Villalba, R., Wanner, H., Wolff, E., Xoplaki, E., 2009. High-resolution palaeoclimatology of the last millennium: A review of current status and future prospects. *The Holocene*, 19, 3–49.

Jones, P., Lister, D., Osborn, T., Harpham, C., Salmon, M., Morice, C.P., 2012. Hemispheric and large-scale land surface air temperature variations: an extensive revision and an update to 2010. *Journal of Geophysical Research* 117, D05127, doi: 10.1029/2011JD017139

Jungclaus, J., Lohmann, K., and Zanchettin, D., 2014. Enhanced 20th-century heat transfer to the Arctic simulated in the context of climate variations over the last millennium. *Climate of the Past*, 10(6), 2201–2213. <http://doi.org/10.5194/cp-10-2201-2014>

Kaplan, J., Krumhardt, K., Ellis, E., Ruddiman, W., Lemmen, C., and Goldewijk, K., 2010. Holocene carbon emissions as a result of anthropogenic land cover change. *Holocene*, 21(5), 775–791. <http://doi.org/10.1177/0959683610386983>

Karl, T., Arguez, A., Huang, B., Lawrimore, J., McMahon, J., Menne, M., Peterson, T., Vose, R., Zhang, H., 2015. Possible artifacts of data biases in the recent global surface warming hiatus. *Science* 348(6242), 1469-1472.

- Klesse, S., Ziehmer, M., Trouet, V., Frank, D.C., 2014. Synoptic drivers of 400 years of summer temperature and precipitation variability on Mt. Olympus, Greece. *Climate Dynamics* 45(3-4), 807-824, doi: 10.1007/s00382-014-2313-3
- Korner, C., 2003. Carbon limitation in trees. *Journal of Ecology*, 91(1), 4–17.
- Kosaka, Y., Xie, S.P., 2013. Recent global warming hiatus tied to equatorial Pacific surface cooling. *Nature* 501, 403–407.
- Krakauer, N., Randerson, J., 2003. Do volcanic eruptions enhance or diminish net primary production? Evidence from tree rings. *Global Biogeochemical Cycles* 17(4), doi: 10.1029/2003GB002076
- Landrum, L., Otto-Bliesner, B. L., Wahl, E. R., Conley, A., Lawrence, P. , Rosenbloom, N., & Teng, H., 2013. Last Millennium Climate and Its Variability in CCSM4. *Journal of Climate*, 26(4), 1085–1111. <http://doi.org/10.1175/JCLI-D-11-00326.1>
- Lawrimore, J., Menne, M., Gleason, B., Williams, C., Wuertz, D., Vose, R., Rennie, J., 2011. An overview of the Global Historical Climatology Network monthly mean temperature data set, version 3. *Journal of Geophysical Research: Atmospheres* (1984-2012) 116, D19121, doi:10.1029/2011JD016187
- Linderholm, H., Björklund, J., Seftigen, K., Gunnarson, B., Fuentes, M., 2014. Fennoscandia revisited: a spatially improved tree-ring reconstruction of summer temperatures for the last 900 years. *Climate Dynamics* 45(3-4), 933-947, doi: 10.1007/s00382-014-2328-9
- Ljungqvist, F., 2010. A new reconstruction of temperature variability in the extra-tropical Northern Hemisphere during the last two millennia. *Geografiska Annaler: Series A, Physical Geography* 92(3), 339–351.
- Ljungqvist, F., Krusic, P., Brattström, G., Sundqvist, H., 2012. Northern Hemisphere temperature patterns in the last 12 centuries. *Climate of the Past* 8, 227-249.
- Loader, N., Young, G., Grudd, H., McCarroll, D., 2013. Stable carbon isotopes from Torneträsk, northern Sweden provide a millennial length reconstruction of summer sunshine and its relationship to Arctic circulation. *Quaternary Science Reviews* 62, 97-113.
- Luckman, B., Wilson, R., 2005. Summer temperature in the Canadian Rockies during the last millennium – a revised record. *Climate Dynamics* 24, 131-144, doi:10.1007/s00382-004-0511-0
- Macias-Fauria M., Grinsted, A., Helama, S., Holopainen, J., 2012. Persistence matters: Estimation of the statistical significance of paleoclimatic reconstruction statistics from autocorrelated time series. *Dendrochronologia* 30, 179-187.
- Mann, M., Bradley, R., Hughes, M., 1998. Global-scale temperature patterns and climate forcing over the past six centuries. *Nature* 392(6678), 779-787.
- Mann, M., Bradley, R., Hughes, M., 1999. Northern hemisphere temperatures during the past millennium: Inferences, uncertainties and limitations. *Geophysical Research Letters* 26, 759–762.

- Mann, M.E., Zhang, Z., Hughes, M.K., Bradley, R.S., Miller, S.K., Rutherford, S., Ni, F., 2008. Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia. *Proceedings of the National Academy of Sciences, USA* 105(36), doi: 10.1073/pnas.0805721105
- Mann, M., Zhang, Z., Rutherford, S., Bradley, R., Hughes, M., Shindell, D., Ammann, C., Faluvegi, G. and Ni, F. 2009. Global signatures and dynamical origins of the Little Ice Age and Medieval Climate Anomaly. *Science* 326 (5957): 1256-1260.
- Mann, M., Fuentes, J., Rutherford, S., 2012. Underestimation of volcanic cooling in tree-ring-based reconstructions of hemispheric temperatures. *Nature Geoscience*. 5(3), 202–205.
- Masson-Delmotte, V., Schulz, M., Abe-Ouchi, A., Beer, J., Ganopolski, A., González Rouco, J.F., Jansen, E., Lambeck, K., Luterbacher, J., Naish, T., Osborn, T., Otto-Bliesner, B., Quinn, T., Ramesh, R., Rojas, M., Shao, X., Timmermann, A., 2013. Information from Paleoclimate Archives, in: Stocker, T.F., Qin, D., Plattner, G.K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M. (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*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 383–464, doi:10.1017/CBO9781107415324.013
- Matskovsky, V., Helama, S., 2014. Testing long-term summer temperature reconstruction based on maximum density chronologies obtained by reanalysis of tree-ring data sets from northernmost Sweden and Finland. *Climate of the Past* 10(4), 1473-1487.
- McCarroll, D., Pettigrew, E., Luckman, A., Guibal, F., Edouard, J.L., 2002. Blue reflectance provides a surrogate for latewood density of high-latitude pine tree rings. *Arctic, Antarctic, and Alpine Research* 34, 450–453.
- McCarroll, D., Loader, N.J., Jalkanen, R., Gagen, M.H., Grudd, H., Gunnarson, B.E., Kirchhefer, A.J., Friedrich, M., Linderholm, H.W., Lindholm, M., Boettger, T. Los, S.O., Remmele, S., Kononov, Y.M., Yamazaki, Y.H., Young, G.H.F., Zorita, E., 2013. A 1200-year multiproxy record of tree growth and summer temperature at the northern pine forest limit of Europe. *The Holocene* 23(4), 471-484.
- Meehl, G., Arblaster, J., Fasullo, J., Hu, A., Trenberth, K.E., 2011. Model-based evidence of deep-ocean heat uptake during surface-temperature hiatus periods. *Nature Climate Change* 1, 360–364.
- Meehl, G., Washington, W., Arblaster, J., Hu, A., Teng, H., Tebaldi, C., Sanderson, B., Lamarque, J.-F., Conley, A., Strand, W.G., White III, J.B., 2012. Climate System Response to External Forcings and Climate Change Projections in CCSM4. *Journal of Climate* 25(11), 3661–3683, doi:http://dx.doi.org/10.1175/JCLI-D-11-00240.1
- Meko, D. 1997. Dendroclimatic reconstruction with time varying predictor subsets of tree indices. *Journal of Climate*, 10:687–696.
- Moberg, A., Sonechkin, D., Homgren, K., Datsenko, N., Karlen, W., 2005. Highly variable Northern Hemisphere temperatures reconstructed from low-and high-resolution proxy data. *Nature* 433, 613–617, doi: 10.1038/nature03265

Oglesby, R., Feng, S., Hu, Q., Rowe, C., 2012. The role of the Atlantic Multidecadal Oscillation on medieval drought in North America: Synthesizing results from proxy data and climate models. *Global and Planetary Change* 84-85: 56–65.

Ortega, P., Lehner, F., Swingedouw, D., Masson-Delmotte, V., Raible, C.C., Casado, M., Yiou, P., 2015. A model-tested North Atlantic Oscillation reconstruction for the past millennium. *Nature* 523, 71-74.

Osborn, T., Briffa, K., 2000. Revisiting timescale-dependent reconstruction of climate from tree-ring chronologies. *Dendrochronologia* 18, 9-25.

PAGES 2k Consortium, 2013. Continental-scale temperature variability during the past two millennia. *Nature Geoscience* 6, 339–346, doi:10.1038/ngeo1797

PAGES2k-PMIP3 group, 2015. Continental-scale temperature variability in PMIP3 simulations and PAGES 2k regional temperature reconstructions over the past millennium, *Climate of the Past Discussion* 11, 2483-2555, doi:10.5194/cpd-11-2483-2015

Palacio, S., Hoch, G., Sala, A., Körner, C., Millard, P., 2014. Does carbon storage limit tree growth?. *New Phytologist* 201(4), 1096-1100.

Parker, D., 1994. Effects of changing exposure of thermometers at land stations. *International Journal of Climatology* 14(1), 1–31.

Payette, S., 2007. Contrasted dynamics of northern Labrador tree lines caused by climate change and migrational lag. *Ecology* 88, 770– 780.

Plummer C, Curran M, van Ommen T, Rasmussen S, Moy A, Vance T, Clausen H, Vinther B and Mayewski P. 2012. An independently dated 2000-yr volcanic record from Law Dome, East Antarctica, including a new perspective on the dating of the 1450s CE eruption of Kuwae, Vanuatu. *Climate of the Past*, 8, 1929-1940, doi:10.5194/cp-8-1929-2012

Pongratz, J., Reick, C., Raddatz, T., and Claussen, M., 2008. A reconstruction of global agricultural areas and land cover for the last millennium. *Global and Biogeochemical Cycles*, 22(3), n/a–n/a. <http://doi.org/10.1029/2007GB003153>

Rohde, R., Muller, R., Jacobsen, R., Muller, E., Perlmutter, S., Rosenfeld, A., Wurtele, J., Groom, D., Wickham, C., 2013. A New Estimate of the Average Earth Surface Land Temperature Spanning 1753 to 2011. *Geoinformatics & Geostatistics: An Overview* 1:1, doi:10.4172/gigs.1000101

Rossi, S., Morin, H., Deslauriers, A., 2012. Causes and correlations in cambium phenology: towards an integrated framework of xylogenesis. *Journal of Experimental Botany* 63(5), 2117-2126.

Rossi, S., Anfodillo, T., Čufar, K., Cuny, H., Deslauriers, A., Fonti, P., Frank, D., Gričar, J., Gruber, A., King, G.M., Krause, C., Morin, H., Oberhuber, W., Prislán, P., Rathgeber, G., 2013. A meta-analysis of cambium phenology and growth: linear and non-linear patterns in conifers of the northern hemisphere. *Annals of Botany* 112(9), 1911-1920.

Rossi, S., Deslauriers, A., Lupi, C., Morin, H., 2014. Control over growth in cold climates, in: Tausz, M., Grulke, N. (Eds.), *Trees in a Changing Environment*. pp. 191-219. Springer, Netherlands, pp. 191-219.

Rydval, M., Larsson, L.Å., McGlynn, L., Gunnarson, B.E., Loader, N., Young, G., Wilson, R., 2014. Blue intensity for dendroclimatology: Should we have the blues? Experiments from Scotland. *Dendrochronologia* 32(3), 191-204.

Rydval, M., Gunnarson, B., Loader, N., Druckenbrod, D., Moreton, S. and Wilson, R., in preparation. Reconstructing 800 years of summer temperatures in Scotland.

Schneider, L., Smerdon, J., Büntgen, U., Wilson, R., Myglan, V., Kirilyanov, A., Esper, J., 2015. Revising midlatitude summer temperatures back to A.D. 600 based on a wood density network. *Geophysical Research Letters* 42(11), 4556-4562, doi: 10.1002/2015GL063956

Schmidt, G. , Jungclauss, J., Ammann, C. , Bard, E., Braconnot, P., Crowley, T., Vieira, L., 2012. Climate forcing reconstructions for use in PMIP simulations of the Last Millennium (v1.1). *Geoscientific Model Development*, 5(1), 185-191. doi:10.5194/gmd-5-185-2012

Schmidt, G., Kelley, M., Nazarenko, L., Ruedy, R., Russell, G. L., Aleinov, I., ... Zhang, J., 2014. Configuration and assessment of the GISS ModelE2 contributions to the CMIP5 archive. *Journal of Advances in Modeling Earth Systems*, 6(1), 141–184. <http://doi.org/10.1002/2013MS000265>

Schurer, A. , Hegerl, G. , Mann, M. , Tett, S. F, Phipps, S. , 2013. Separating Forced from Chaotic Climate Variability over the Past Millennium. *Journal of Climate*, 26(18), 6954–6973. <http://doi.org/10.1175/JCLI-D-12-00826.1>

Schurer, A., Tett, S., Hegerl, G, 2014. Small influence of solar variability on climate over the past millennium. *Nature Geoscience*. 7, 104–108, doi: 10.1038/ngeo2040

Sigl, M., McConnell, J., Layman, L., Maselli, O., McGwire, K., Pasteris, D., Dahl-Jensen, D., Steffensen, J., Vinther, B., Edwards, R., Mulvaney, R., Kipfstuhl, S., 2012. A new bipolar ice core record of volcanism from WAIS Divide and NEEM and implications for climate forcing of the last 2000 years. *Journal of Geophysical Research: Atmospheres* 118(3), 1151-1169, doi: 10.1029/2012JD018603

Smerdon, J., Kaplan, A., Zorita, E., González-Rouco, J., Evans, M., 2011. Spatial performance of four climate field reconstruction methods targeting the Common Era. *Geophysical Research Letters* 38, L11705, doi:10.1029/2011GL047372

St. George, S., 2014. An overview of tree-ring width records across the Northern Hemisphere. *Quaternary Science Reviews* 95, 132-150.

St. George, S., Ault, T., 2014. The imprint of climate within Northern Hemisphere trees. *Quaternary Science Reviews* 89, 1-4.

Stine, A., Huybers, P., 2014. Arctic tree rings as recorders of variations in light availability. *Nature Communications* 5, 1–8, doi:10.1038/ncomms4836

Steinhilber, F., Beer, J., & Fröhlich, C., 2009. Total solar irradiance during the Holocene. *Geophysical Research Letters*, 36(19), L19704. <http://doi.org/10.1029/2009GL040142>

Stoffel, M., Khodri, M., Corona, C., Guillet, S., Poulain, V., Bekki, S., Guiot, J., Luckman, B., Oppenheimer, C., Lebas, N., Beniston, M & Masson-Delmotte, V. (2015). Estimates of volcanic-induced cooling in the Northern Hemisphere over the past 1,500 years. *Nature Geoscience*. 8, 784–788 (2015) doi:10.1038/ngeo2526.

Stokes, M., Smiley, T., 1968. An introduction to tree-ring dating. University of Chicago Press, Chicago, Illinois, USA.

Szeicz, J., MacDonald, G., 1995. Dendroclimatic reconstruction of summer temperatures in northwestern Canada since A. D. 1638 based on age-dependent modeling, *Quaternary Research* 44(2), 257–266.

Tingley, M. , & Huybers, P., 2013. Recent temperature extremes at high northern latitudes unprecedented in the past 600 years. *Nature*, 496(7444), 201–205. doi:10.1038/nature11969

Tolwinski-Ward, S., Evans, M., Hughes, M., Anchukaitis, K., 2011. An efficient forward model of the climate controls on interannual variation in tree-ring width. *Climate Dynamics* 36(11-12), 2419-2439.

Trenberth, K., Fasullo, J., 2010. Simulation of present-day and twenty-first-century energy budgets of the Southern Oceans. *Journal of Climate* 23(2), 440–454, doi: 10.1175/2009jcli3152.1

Trouet, V., Esper, J., Graham, N., Baker, A., Scourse, J., Frank, D., 2009. Persistent positive North Atlantic Oscillation mode dominated the medieval climate anomaly. *Science* 324, 78–80.

Vaganov, E., Hughes, M., Shashkin, A., 2006. Growth dynamics of conifer tree rings images of past and future environments. Springer, Berlin, New York.

Vaganov, E., Anchukaitis, K., Evans, M., 2011. How well understood are the processes that create dendroclimatic records? A mechanistic model of climatic control on conifer tree- ring growth dynamics, in: Hughes, M.K., Swetnam, T.W., Diaz, H.F. (Eds.), *Dendroclimatology: Progress and Prospects, Developments in Paleocological Research*. Springer-Verlag, Netherlands, 11(2), pp. 37-75, doi:10.1007/978-1-4020-5725-0\_3

Vieira, L. , Solanki, S., Krivova, N. , & Usoskin, I., 2011. Evolution of the solar irradiance during the Holocene. *Astronomy & Astrophysics*, 531, A6. <http://doi.org/10.1051/0004-6361/201015843>

Wahl, E., Ammann, C., 2007. Robustness of the Mann, Bradley, Hughes reconstruction of Northern Hemisphere surface temperatures: examination of criticisms based on the nature and processing of climate evidence. *Climatic Change* 85(1-2), 33–69, doi: 10.1007/s10584-006-9105-7

Wiles, G., D'Arrigo, R., Barclay, D., Wilson, R., Jarvis, S., Vargo, L., Frank, D., 2014. Surface air temperature variability reconstructed with tree rings for the Gulf of Alaska over the past 1200 years. *The Holocene* 24(2), 198-208, doi: 10.1177/0959683613516815

Wilson, R., Tudhope, A., Brohan, P., Briffa, K., Osborn, T. and Tett, S. 2006. 250-years of reconstructed and modeled tropical temperatures. *Journal of Geophysical Research*. 111, C10007, doi:10.1029/2005JC003188

Wilson, R., D'Arrigo, R., Buckley, B., Büntgen, U., Esper, J., Frank, D., Luckman, B., Payette, S., Vose, R., Youngblut, D., 2007. A matter of divergence – tracking recent warming at hemispheric scales using tree-ring data. *Journal of Geophysical Research:– Atmospheres (1984-2012)* 112(D17), doi: 10.1029/2006JD008318

Wilson, R., Rao, R., Rydval, M., Wood, C., Larsson, L.Å., Luckman, B.H., 2014. Blue Intensity for Dendroclimatology: The BC Blues: A Case Study from British Columbia, Canada. *The Holocene* 24(11), 1428-1438, doi: 10.1177/0959683614544051

Wu, T., Li, W., Ji, J., Xin, X., Li, L., Wang, Z., Zhang, Y., Li, J., Zhang, F., and Wei, M., 2013. Global carbon budgets simulated by the Beijing Climate Center Climate System Model for the last century. *J. Geophys. Res.: Atm.*, 118(10), 4326–4347. <http://doi.org/10.1002/jgrd.5032>

Young, G., Bale, R., Loader, N., McCarroll, D., Nayling, N., Vousden, N., 2012. Central England temperature since AD 1850: the potential of stable carbon isotope in British oak trees to reconstruct past summer temperatures. *Journal of Quaternary Science* 27(6), 606–614.

Youngblut, D., Luckman, B., 2008. Maximum June-July temperatures in the southwest Yukon region over the last three hundred years reconstructed from tree-rings. *Dendrochronologia* 25(3), 153-166.

Zhang, P., Linderholm, H., Gunnarson, B., Björklund, J., Chen, D., 2015. 1200 years of warm-season temperature variability in central Fennoscandia inferred from tree-ring density. *Climate of the Past Discussion* 11, 489-519, doi: 10.5194/cpd-11-489-2015

## Last millennium Northern Hemisphere summer temperatures from tree rings: Part I: the long term context

Rob Wilson, Kevin Anchukaitis, Keith [R. Briffa](#), Ulf Büntgen, Ed Cook, Rosanne D' Arrigo, Nicole Davi, Jan Esper, Dave Frank, Björn Gunnarson, Gabi Hegerl, Samuli Helama, Stefan Klesse, Paul Krusic, Hans W. Linderholm, Vladimir Myglan, Timothy [J. Osborn](#), Milos Rydval, Lea Schneider, Andrew Schurer, Greg Wiles, Peng Zhang and Eduardo Zorita.

### Abstract

Large-scale millennial length Northern Hemisphere (NH) temperature reconstructions have [been progressively improved](#) over the last 20 years as new data-sets [are developed](#). This paper, and its [part II](#) companion ([Part II](#), Anchukaitis et al. [this volume in prep](#)), details the latest tree-ring (TR) based NH [land air temperature](#) reconstruction from a temporal and spatial perspective. This work is the first product of a consortium called N-TREND (Northern Hemisphere Tree-Ring Network Development) which [aims to bring together](#) dendroclimatologists to identify a collective strategy for improving large-scale summer temperature reconstructions. The new reconstruction, N-TREND2015, [utilises 54 records, a significant expansion compared with previous TR studies, and yields an improved reconstruction with stronger statistical calibration metrics](#)~~utilises 54 records and is a significant improvement on previous studies~~. N-TREND2015 is relatively insensitive to the compositing method and spatial weighting used and validation metrics indicate that the new record portrays reasonable coherence with large scale summer temperatures and is robust at all time-scales from 918 to 2004 where at least 3 TR records exist from each major continental mass. N-TREND2015 indicates a longer and warmer medieval period (~900 – 1170) than portrayed by previous TR NH reconstructions and [by](#) the CMIP5 model ensemble, but ~~all records are in~~ [with better overall agreement between records](#) for the last 600 years. Future dendroclimatic projects ~~need to~~ [should](#) focus on developing new long records from data-sparse regions such as North America and eastern Eurasia as well as [ensuring](#) the measurement of [parameters related to](#) latewood density ~~related parameters~~ to complement ring-width records which can improve local based calibration substantially.

### Keywords

Tree-rings; Northern Hemisphere; Summer Temperatures; Reconstruction; CMIP5 Models; [last millennium](#)

### Introduction

Over the past two decades, substantial effort has been directed towards reconstructing late Holocene Northern Hemisphere temperature trends and variability from high resolution paleoclimate proxy data (Frank et al. 2010, Masson-Delmotte et al., 2013). While some studies have focused entirely on the use of tree-rings (Esper et al. 2002; D'Arrigo et al. 2006; Frank et al. 2007a; Wilson et al. 2007; Schneider et al. 2015; Stoffel et al. 2015), most have also included different types of proxy archives and are

thus considered multi-proxy reconstructions (Jones et al. 1998; Mann et al. 1999, 2009; Moberg et al. 2005, Hegerl et al. 2007, Wahl and Amman 2007; Ljungqvist 2010). The majority of these reconstructions have generated a single hemispheric wide mean temperature composite, although a few noteworthy spatial field reconstructions have been produced (Briffa et al. 2002b; Mann et al. 1998, 2009; Ljungqvist et al. 2012; Tingley and Huybers 2013). Such analyses are crucial not only to place recent warming in a longer term spatiotemporal context, but also allow, through comparative analysis with global climate models, to improved understanding of the high-to-low frequency forcing mechanisms of long-term and recent climate changes using formal detection and attribution (Hegerl et al. 2006; Schurer et al. 2014).

Figure 5.7 in the most recent IPCC Working Group 1 report (section 5.3.5; Masson-Delmotte et al., 2013) synthesises the majority of the published Common Era reconstructions for the Northern Hemisphere (NH). Focusing on the last millennium, the NH ensemble indicates warm conditions around the end of the 10<sup>th</sup> century, a cooling trend over the subsequent 500 years, a prolonged cool period from ~1450-1850 (often referred to as the Little Ice Age) and steep warming from the middle of the 19<sup>th</sup> century to present. Although the basic centennial “shape” of these NH reconstructions is in general agreement, there are notable differences in the timing and amplitude of some cold and warm periods (e.g. Esper et al. 2004). This uncertainty hampers detection and attribution studies and makes it difficult to constrain modelled scenarios of future temperature change (Edwards et al. 2007).

We hypothesise that these differences partly reflect the use of multi-proxy data in the current ensemble of NH reconstructions and the quality of many constituent proxy series. Although the multi-proxy approach is a theoretical ideal - incorporating and subsuming the independent advantages and disadvantages of the various proxy types - in practice the biases inherent to each proxy archive may also potentially compound the uncertainty in the final large-scale composites, which are further exacerbated by the influence of the varying statistical methods used to combine the proxy records (Smerdon et al. 2011). For example, many of the proxies utilized in Common Era reconstructions actually reflect varying seasonal signals (e.g. summer vs. annual) and some may do not even represent local temperature well (Mann et al. 1998, 2009). Combining such heterogeneous data may produce a network possessing a reasonable correlation with an annual instrumental target, but may that also contains spatio-temporal proxy biases that add to the uncertainty in the final reconstruction. Furthermore, non-annual proxies, lacking precise annual resolution and dating, must be smoothed to decade or longer time-scales reducing (1) the ability of these data to accurately record short-term climatic response to large volcanic events, and (2) the degrees of freedom – which places limitations on the ability to perform robust calibration and validation analyses to assess reconstruction fidelity, stability, and skill.

In this paper (and its part II companion, Anchukaitis et al. *this volume in prep*), we focus entirely on temperature-sensitive tree-ring (TR) records to both reduce and better characterize many of the uncertainties detailed above. Tree-ring archives are annually resolved (facilitating robust validation), precisely dated (Stokes and Smiley 1968) and the interpretation of their measurements is supported by a wealth of ecological and biological process-based knowledge of how climate variability influences ring formation (Fritts 1976, Korner 2003, Vaganov et al. 2006; Deslauriers et al. 2007, Vaganov et al. 2011, Rossi et al. 2012, Rossi et al. 2013, Rossi et al. 2014, Palacio et al. 2014). Temperature-sensitive tree-ring TR archives are derived from many high elevation/latitude environments where growth is predominantly limited by summer temperatures, permitting the development of a well replicated, annually resolved, network of summer temperature proxies for the mid/high latitudes of the Northern Hemisphere (Briffa et al. 2001). Furthermore, as the summer season is most sensitive to short term volcanically forced cooling, a tree-ring only reconstruction increases the potential for more robust assessment of volcanic forcing in

attribution studies in comparison to ~~a~~ large-scale mixed season multi-proxy composites (SchneiderSchneider et al. 2015; Stoffel et al. 2015).

Herein, we introduce a new and substantially updated network of temperature-sensitive tree-ring records for the mid-to-high latitudes of the Northern Hemisphere. These data are used to reconstruct both a single mean hemispheric series (Part I) as well as explicit spatial temperature fields (Part II, Anchukaitis et al. ~~this volume in prep~~) for the last millennium. Careful comparison is made between the new reconstruction and two previous TR based NH index reconstructions; D'Arrigo et al. (2006 – hereafter DWJ06 – a predominantly ring-width (RW) based composite) and Schneider et al. (2015 – hereafter SCH2015 – a maximum latewood density (MXD) based composite) to explore ~~both~~ the strengths and limitations of RW and MXD. This work is the product of a tree-ring community consortium called N-TREND (Northern Hemisphere Tree-Ring Network Development). The aim of N-TREND is to bring together dendroclimatologists to identify a collective strategy for improving large-scale summer temperature reconstructions from new and existing tree-ring archives. These two papers, the first product of the N-TREND consortium, aim to not only improve our knowledge of Northern Hemisphere temperature changes, but to also emphasise the strengths and limitations of TR archives and guide future dendroclimatic research.

## Tree-Ring Data

Many thousands of tree-ring site chronologies are archived within the International Tree-Ring Data Bank but the majority are either relatively short (< 200-years in length), have not been updated to the recent post-2000 period and/or do not reflect summer temperatures (St. George 2014; St George and Ault 2014). For N-TREND, rather than statistically screening all extant TR chronologies for a significant local temperature signal, we utilise mostly published TR temperature reconstructions ~~(or chronologies used in published reconstructions)~~ that start prior to 1750. This strategy explicitly incorporates the ~~knowledge of expert judgment~~ the original authors ~~as they derived~~ used to derive the most robust reconstruction possible from the available data at that particular location. The data-set used herein is defined as version number N-TREND2015, recognizing the fact that as new local/regional tree-ring temperature reconstructions are developed, the N-TREND database will expand and be further refined. N-TREND2015 includes 54 TR records located between 40°N and 75°N (Table 1 and Figure 1). TR data south of 40°N are not included as trees ~~tend to exhibit an increasingly complex sensitivity to multiple climate influences including precipitation growing at lower latitudes can be sensitive to both temperature and precipitation~~ (Fritts 1976, St. George 2014, St George and Ault 2014) which will bias the temporal and spectral properties of a reconstruction (Osborn and Briffa 2000; Franke et al. 2013). The N-TREND data-set includes single site, regional and large-scale gridded reconstructions. Specific information on each TR series is detailed in Appendix Table A1.

Most of the TR records have been processed using Regional Curve Standardisation (RCS - Briffa et al., 1996; Esper et al., 2003 – Table 1) which improves the likelihood of retaining secular trends beyond the mean sample length of the chronologies (Cook et al. 1995). ~~18-Eighteen~~ records are based entirely on maximum latewood density (MXD), and 11 utilised ~~d~~ only ring-width (RW) measurements. A further 25 records are defined as “mixed” and represent a heterogeneous group of local/regional/gridpoint reconstructions that have incorporated combinations of RW, MXD and Blue Intensity (BI - defined in Appendix Table A1 – McCarroll et al. 2002; Campbell et al. 2011; Björklund et al. 2014; Rydval et al. 2014) ring parameters aggregated using a variety of multivariate methodologies (Appendix Table A1). Post-1950  $r^2$  values of the individual records regressed on their local gridded (CRUTS3.2 - Harris et al. 2014), empirically derived, optimal temperature season (Table 1) range from 0.12 to 0.71 (Figure 1). Of particular

note are the overall weaker results observed for those sites derived using only RW data. When MXD (or BI) is used, local/regional correlations are almost always superior. The common period for all 54 records is 1710-1988; 38 records extend back to 1500 while 23 records extend [back](#) to at least 1000. Substantially more TR records are available for Eurasia (39) compared to North America (15) with a distinct imbalance of data [by at](#) 1000 (20 vs. 3 – Figure 2a).

## Reconstruction Methodology

A similar iterative nesting method (Meko 1997; Cook et al. 2002), as utilised in D'Arrigo et al (2006) and Wilson et al. (2007), was used to develop the N-TREND2015 NH temperature reconstruction. This approach involves first normalising the TR data over a common period (1750-1950), averaging the series to derive a mean series and iteratively removing the shorter series to allow the extension of the reconstruction back (as well as forward) in time. Each nest is then scaled to have the same mean and variance as the most replicated nest (hereafter referred to as NEST 1) and the relevant time-series sections of each nest spliced together to derive the full-length reconstruction. For each nest, separate average time series were first generated for 4 longitude quadrats (Figure 1). These continental scale time series were then averaged (after again normalising to 1750-1950) to produce the final large-scale hemispheric mean [that would](#) not [be](#)-biased to data rich regions in any one continent. 37 backward nests and 17 forward nests were calculated to produce the full reconstruction length from 750 – 2011. Other methodological options (i.e. different spatial weighting procedures, and total least squares regression) were also tested to assess the sensitivity of the final NH series to varying compositing and reconstruction approaches (see Appendix C).

Full period calibration was performed over the 1880-1988 period targeting the latitudinal band 40-75 °N. We utilised herein the [land-only](#) CRUTEM4 data-set (Jones et al. 2012) although similar coherence results are obtained using different gridded temperature products (Appendix Figure B1a). Calibration experiments using NEST 1 identified mean May-August ([MJJ](#)) temperatures as the optimal large-scale instrumental target. Separate calibration and validation analyses were performed on each nested mean series allowing assessment of temporal changes in reconstruction quality. For independent validation of the TR based temperature estimates, calibration/validation was performed using a traditional split period (1880-1934 and 1935-1988) approach as well as calibrating on the combined 1880-1915/1953-1988 periods and verifying over 1916-1952. As large-scale NH instrumental temperatures show a steep increasing trend since the 19<sup>th</sup> century, the reduced variance of the simple regression based estimates make it difficult to predict the steep trend outside a short calibration period, particularly when calibration  $r^2$  values are modest. The latter calibration/validation approach allows a theoretical interpolation based assessment of the temperature estimates rather than the extrapolations of the simple early/late split period approach.

[Validation-Evaluation](#) of the reconstruction was performed using the square of the Pearson's correlation coefficient ( $r^2$ ), the Reduction of Error (RE) statistic, and the Coefficient of Efficiency (CE) statistic (Cook and Kairiukstis 1990; Cook et al. 1994). To test the robustness of the decadal to long-term signal in the reconstructed nested series, the regression model residuals (from the full period calibration) were assessed for linear trends, and autocorrelation using the Durbin-Watson statistic ([see](#) Wilson et al. 2006). A poor composite proxy fit at decadal time-scales will result in significant residual autocorrelation, while biases in centennial scale trend estimates would result in residual linear trends. The root mean square error (RMSE) was used to estimate reconstruction uncertainty. To derive the final reconstruction, NEST 1 was scaled to the instrumental data over the 1880-1988 period and the RMSE values calculated using the residual differences between the actual values and scaled estimates (referred to as  $RMSE_{scd}$ ). All

other nested series were scaled to NEST 1 over the 1710-1988 period, and the  $RMSE_{scl}$  value, calculated for each nest, provides an estimate of the change in calibration based uncertainty through time.

Although many of the individual reconstructions (Table 1) have published uncertainty estimates, some do not. Even for those that have presented reconstruction uncertainty, they are mostly related to calibration uncertainty, and conventionally, few tree-ring temperature reconstructions have attempted to quantify uncertainty related to both detrending choices and decreasing sample size back in time as the numbers of trees decrease (Esper et al. 2007; Matskovsky and Helama 2014; Wilson et al. 2014). This is problematic as the actual uncertainty of such large-scale reconstructions will be greater than that simply estimated from the nested mean composite calibration based  $RMSE_{scl}$  values. This extra uncertainty will increase back in time but there is no single way to quantify it, particularly without fundamental knowledge of the raw data used for individual tree-ring chronologies. To produce a conservative estimate of this additional error that we acknowledge may exist in our data, a theoretical worst case scenario is proposed wherein an individual input reconstruction is degraded until it has zero correlation ~~calibrates~~ with its local gridded data ~~with an  $r^2$  of zero~~. The maximal amount of error arising in this scenario is modelled by fitting a horizontal straight-line to the instrumental data and calculating the grid specific RMSE. The grand mean RMSE for all 54 grid locations comes to 1.00 °C ( $\pm 0.23$  standard deviations). This value was used to derive an additional uncertainty estimate (referred to as  $RMSE_{add}$ ) that is added to the  $RMSE_{scl}$  values but is scaled by dividing the  $RMSE_{add}$  by the square root of  $N$  (number of records) as input series replication changes through time (see Appendix Figure B2).

To compute the error for the smoothed reconstruction, we follow a method first outlined by Briffa et al. (2002a - modified by Gouirand et al. 2008) and applied by Divine et al. (2011), but modified for use with RMSE rather than the standard error of the regression estimate:

$$(RMSE_{adj})^2 = \frac{1}{r^2} \left( \frac{(RMSE_{tot})^2}{n} \right)$$

where  $RMSE_{tot} = RMSE_{scl}$  and  $RMSE_{add}$  (see Appendix Figure B2),  $n$  is the window length in years of the smoothing function (a 20-year cubic smoothing spline (Cook and Peters 1981)) and  $r^2$  is the explained variance for each smoothed nest calibration.

## Results and Discussion

### **Calibration and Validation of the N-TREND2015 reconstruction**

The N-TREND2015 reconstruction and associated quality metrics are shown in Figure 2 (B-E – calibration/validation results detailed in Appendix Table B1). NEST 1, representing the period 1710-1988, with an unfiltered calibration  $r^2$  of 0.41, does not explain the greatest amount of instrumental variance in the full record. NEST 11 (relevant for the period 1510-1520) explains the most temperature variance at 47%. This appears to occur because more recent nests contain shorter and noisier TR records that, at the hemisphere scale considered here, are somewhat detrimental to the overall skill of the reconstruction. However, as the N-TREND2015 network is also used for spatial field reconstruction (see Part II, Anchukaitis et al. ~~this volume in prep~~), removing the weaker shorter series would be deleterious for those locations where only short records exist. Overall, unfiltered calibration  $r^2$  values are  $> 0.40$  from 1200 – 1988 with values dropping below 0.30 before 918 and after 2004 (Figure 2F). ~~For comparison, These are better than those obtained either by~~ DWJ06 (RCS version) which produced unfiltered  $r^2$  values  $> 0.30$  from 1556 – 1978 and were never  $> 0.40$  ~~— or by~~ SCH2015, which also never ~~attained values  $\rightarrow$  exceeded~~ 0.40 ~~but and~~

demonstrated calibration  $r^2$  values  $> 0.30$  only from 1363 – 1976. When the data are smoothed with a 20-year cubic smoothing spline, N-TREND2015 calibration  $r^2$  values substantially increase, as expected, and are  $> 0.80$  from 910 – 1989, only dropping below 0.70 before 760 and after 2010 (Figure 2F). DWJ06 and SCH2015 explained  $> 0.70$  of the smoothed temperature variance between 1340 and 1995 and 881 and 1976, respectively. ~~Both-Neither~~ records ~~never~~ attains  $r^2$  values ~~above~~  $> 0.80$  when the data are smoothed. Spectral coherency analysis (Appendix Figure B1b) clearly shows the stronger coherence at decadal and longer time-scales, but also highlights the improved high-mid frequency coherence in the N-TREND2015 data compared to DWJ06. SCH2015 explains a similar amount of mid-frequency variance but shows dampened coherence at longer time-scales.

No significant autocorrelation is noted in the N-TREND2015 regression model residuals for any nest except after 2006 where the proxy network is quite sparse and the calibration is already weak (unfiltered  $r^2 \leq 0.20$ ). No residual linear trends are found for any nest spanning the period 760 – 2010 (Figure 2G). With the relatively strong smoothed  $r^2$  values (Figure 2F) and generally robust residuals for most nests, the decadal and longer scale fidelity in the N-TREND2015 reconstruction is a significant improvement over previous studies back to the 10<sup>th</sup> century. However, prior to 1880, there is a distinct divergence (Figure 2D+E) between the instrumental data (warmer) and N-TREND2015 reconstruction (cooler) potentially ~~related to~~ reflecting the well-previously discussed “warm bias” in early instrumental summer temperature series (Parker 1994; Frank et al. 2007b; Böhm et al. 2010). Also, there is a hint of under-prediction of temperature estimates in the recent two decades which is likely mainly related to the decreasing number of TR records through this period (Figure 2A).

Validation metrics are presented in Figures 2H-J (individual period validation results detailed in Appendix Table B1). Results are rather variable depending on the validation period used. Validation  $r^2$  values are greater than zero for all nests, but are strongest for the late 1935-1988 period and weakest for the mid 1916-1952 period. Maximum  $r^2$  values are attained from about the 12<sup>th</sup> century to the recent period. While RE values are positive for all nests, CE values are positive only since the late 14<sup>th</sup> century. CE values first go below zero before 1373 using the 1935-1988 validation period when the Quebec MXD (QUEx) series drops out from the nesting process. CE is a particularly stringent validation metric and care is needed when using this metric alone to assess reconstruction quality (Cook et al. 1994). One further driver of the poor CE values is related to the reconstruction’s warmer values centred around 1950 versus the CRUTEM instrumental data showing peak mid-20<sup>th</sup> century warming in the 1940s (Figure 2E). This slight misfit appears to be related to both the spatial ~~representivity~~ representativeness of the TR network and the fact that mean MJJA temperatures is not the optimal season for some TR records (Table 1). Appendix Figure B3 compares an instrumental based reconstruction, which utilises the 0.5 latitude x 0.5 longitude CRUTS3.20 (Harris et al. 2014) gridded optimal calibrated seasonal instrumental series at the locations of the 54 TR series against the MJJA CRUTEM4 (40-75°N – Jones et al. 2012) hemisphere series. The instrumental based composite explains 78% of the CRUTEM4 variance (95% when smoothed), but places the mid-20<sup>th</sup> century warming peak in the 1950s and under-predicts the recent period similar to that observed for the N-TREND2015 reconstruction (Figure 2E). These observations suggest that despite the increased number of TR proxy records, substantially more than 54 instrumental records are needed to attain a ~~true-robust~~ complete representation of the large-scale decadal to high frequency climate signals (see also discussion in Diaz 1996, Jones and Briffa 1996 and Anchukaitis et al. this volume in prep).

Compared to previous millennial long-length TR NH reconstructions (Esper et al. 2002; D’Arrigo et al. 2006; Frank et al. 2007a; Schneider et al. 2015) the N-TREND2015 reconstruction shows almost no late 20<sup>th</sup> century divergence (D’Arrigo et al. 2008 – see also discussion in Wilson et al. 2007) from the instrumental data (Figure 2D+E) and is statistically more robust for a longer period of time. In this respect,

N-TREND2015 therefore represents a substantial update of these earlier studies. Fidelity weakens, as expected, when replication is low, especially over North America. Balancing calibration/validation metrics and site replication (Figure 1), we therefore define the period 918-2004, represented by at least 3 TR records for each major continental mass, as a compromise period for “robust” comparison with other records although [the](#) comparative analyses below use mostly the full period.

### **Comparison to other NH reconstructions**

N-TREND2015 shows a modest increase in temperatures from 750 to ~850 followed by a prolonged warm period until the end of the 11<sup>th</sup> century, a brief cooler 50-year period and then a short period of peak medieval warmth in the 1160s (Figure 3). 1161-1170 is the 3<sup>rd</sup> warmest decade in the reconstruction followed by 1946-1955 (2<sup>nd</sup>) and 1994-2003 (1<sup>st</sup> – see Table 2). It should be noted that these three decades cannot be statistically distinguished when uncertainty estimates are taken into account. Following 2003, 1168 is the 2<sup>nd</sup> warmest year, although caution is advised regarding the inter-annual fidelity of the reconstruction (Appendix Figure B1B). The period 1200-1800 is characterised by substantial multi-decadal variability with a range of about 1<sup>st</sup> ~~degree~~ Celsius between warm peaks and cold troughs. The two coldest ~~decadal periods~~ decades, 1812-1821 and 1832-1841 (Table 2), are followed by a near continuous warming to present.

Formatted: Superscript

Figure 3 compares N-TREND2015 with DWJ06 (D’Arrigo et al. 2006) and SCH2015 (Schneider et al. 2015). Prior to 2015, DWJ06, using 19 predominantly RW records, was the most comprehensive TR specific millennial length NH reconstruction. SCH2015 is a new NH summer temperature reconstruction based on 16 MXD records. N-TREND2015 is not independent from these two studies. 3 records (ICE, TAY and YAK) were used from DWJ06 and 5 from SCH2015 (QUEx, TYR, POLx, MAN, and ALT). N-TREND2015 also shares common data with SCH2015 for some sites (e.g. ICE, EFmean, ALPS), but has used the original published versions rather than the SCH2015 processed versions. DWJ06 and SCH2015 can be seen as two end members representing quite different TR parameters - both reflecting predominantly summer temperatures but each expressing different spectral properties and potential biases. RW generally demonstrates strong biological persistence and [higher](#) autocorrelation ([Helama et al. 2009b](#); Esper et al. 2015) while MXD has lower temporal autocorrelation. ~~However~~ ~~Although~~, MXD ~~may struggle~~ ~~has been shown~~ to capture secular scale [summer temperature](#) trends ([Helama et al. 2009b](#), but see [Luckman and Wilson 2005](#); Esper et al. 2012) ~~and it~~ may also be influenced by light availability ~~as well a summer temperatures~~ (Stine and [Huybers 2014](#)). Figure 1 clearly shows the superiority of MXD (and BI) based TR measurements for modelling inter-annual local scale summer temperatures. However, there is still uncertainty as to the potential spectral biases of using these different TR parameters to capture inter-annual, decadal and longer-scale climate.

DWJ06 shows similar warm and cool periods to N-TREND2015, particularly after 1400. D’Arrigo et al. (2006) advised caution when using the pre-1400 period of the reconstruction due to the low number of TR records (10 in 1300; 6 in 1000). N-TREND2015 should therefore represent a considerable improvement in this period due to the greater number of records used (32 in 1300; 23 in 1000). In fact, the N-TREND2015 reconstruction is substantially different before 1400 with it showing a more prolonged warming from the 9<sup>th</sup> to 12<sup>th</sup> centuries compared to the short-term warmth from 950-1100 in DWJ06. SCH2015 displays more modest warmth during the 850-1050 period with a peak at the end of 9<sup>th</sup> century and overall muted multi-decadal variability until 1300. From 1300-1570, SCH2015 shows warmer conditions than N-TREND2015 although overall, multi-decadal trends are similar. The decadal cooling of the 1450s is however more pronounced in N-TREND2015 than DWJ06 and SCH2015. The three records are quite similar after 1600 with

similar short term cool periods coinciding with known volcanic events. Only N-TREND2015 expresses the post mid-20<sup>th</sup> century warming trend. N-TREND2015 correlates (918-1999 – unfiltered (filtered)) marginally stronger with DWJ06 ( $r = 0.65$  (0.73)) than with SCH2015 ( $r = 0.61$ (0.61)), but for running shorter 50-year periods, N-TREND2015 and SCH2015 are clearly more coherent (Figure 3 – lower panel). These correlations highlight the greater coincidence in multi-decadal and longer term signals between N-TREND2015 and DWJ06, and the greater high-frequency coherence with SCH2015 – the latter’s inter-annual variability driven by the stronger high-frequency signal within the MXD data used.

Figure 4 compares the aforementioned TR composites (plus Frank et al. 2007a – FRK07) with recently published multi-proxy (MP) Northern Hemisphere temperature reconstructions (Amman and Wahl 2007; Hegerl et al. 2007; Ljungqvist 2010; Mann et al. 2008) and their mean. Overall, N-TREND2015 correlates strongest ( $r = 0.76$ ) with the MP mean series although only marginally better than DWJ06 ( $r = 0.72$ ) and substantially better than SCH2015 ( $r = 0.49$ ) and FRK07 ( $r = 0.48$ ). N-TREND2015 and the MP composite show common warm conditions from the mid-9<sup>th</sup> to the end of the 11<sup>th</sup> centuries although the MP series suggests a warmer medieval period until the mid-12<sup>th</sup> century. When divided into four longitudinal quadrants (Figure 1), this protracted period of medieval warming appears to be spatially focused over eastern North America (although in that quadrant the mMedieval period is presented expressed by only a single record from Quebec – Table 1) and Western Eurasia and extends to the end of the 12<sup>th</sup> century (Appendix C2) suggesting a common dynamical link with North Atlantic variability, the thermohaline circulation/meridional overturning circulation and purported prolonged positive North Atlantic Oscillation anomalies through this period (Trouet et al. 2009; Helama et al. 2009a; Cunningham et al. 2013). Although a persistently ly positive NAO during the MCA-medieval period has recently being-been challenged by Ortega et al. (2015), the sharp rise in temperatures around 1150, which remains largely positive for around three centuries in eastern North America and western Eurasia coincides with the timing of an inferred positive phase of the NAO during winter (Trouet et al. 2009). Focusing on summer, Loader et al. (2013) suggested that the period 900-1200 was dominated by a northerly shift of the NH storm tracks, yielding warm summers in the high latitudes of NW Europe. Possibly this could be related to This warming is also reflected in an extended period of warm sea surface temperatures in the North Atlantic (Cunningham et al. 2013), associated with a positive phase of the Atlantic Multidecadal Oscillation (AMO) during the MCA-medieval period, as suggested by Oglesby et al. (2012), leading to a decrease in the mid-latitude – Arctic temperature gradient.

From ~1100 to ~1700, N-TREND shows a prolonged period of multi-decadal variability embedded within a general cooling trend. Common cool conditions are observed in all the proxy records until the mid-19<sup>th</sup> century after which temperatures rise until present. As noted above, SCH2015 expresses muted medieval warmth, shows warmer conditions during the late 14<sup>th</sup> and early 15<sup>th</sup> centuries, and does not capture recent warming well. FRK07 has a shorter period of warmth during the medieval period (similar to DWJ06), but is substantially cooler in the 13<sup>th</sup> century.

Overall, comparing the TR and MP reconstructions provides a broadly coherent picture of warm conditions around the end of the 10<sup>th</sup> century, a cooling trend for about 500 years, a prolonged cool period from ca. 1450-1850 and steep warming until present. However, there are subtle differences between the reconstructions that likely reflect (1) reconstruction composition (i.e. older studies used less data; including non-annual proxy archives with associated dating uncertainties), (2) varying spectral properties of different proxy types (including TR-tree-ring RW and MXD in comparison with MP networks) and (3) the different reconstruction’s seasonal target (summer vs. annual). The issue of seasonality is particularly important for understanding the stochastic nature of climate. The current debate over the so-called “temperature hiatus” (Meehl et al. 2011; Cohen et al. 2012; Kosaka and Xie 2013; Trenberth and Fasullo 2013; England et al.

2014; Karl et al. 2015) exemplifies the consequence of mixing proxy records with varying seasonal sensitivities. The “hiatus” is observed in large scale winter temperatures and no slowdown is noted in NH summer temperatures (Cohen et al. 2012). Differences in timing and duration of warm and cold periods between TR and MP reconstructions could well be related to the same issue – highlighting the need for the palaeoclimate community to carefully evaluate the seasonal signal embedded in their proxies.

### **Proxy/model Comparison – long term context**

A palaeoclimate aspiration is that independently derived reconstructions and models of past climate should express similar trends (where externally forced) and variance structure through time. Agreement in hindcast modelling ability will provide some confidence in scenario forecasts of the future. In comparison to the challenging regional model-data assessments (Zorita et al. 2010, Deser et al. 2012), hemispheric scale comparison is generally more straightforward: heterogeneous local/regional internally-~~generated-forced~~ variability is often dampened through averaging, allowing the influence of external forcing factors such as volcanic events and anthropogenic emissions to be addressed (Fernández-Donado et al. 2013, but see Anchukaitis et al. 2012). The strong coherence between the North American and Eurasian TR composites (Appendix Figure C3) indicates the dominance of large scale forcing on such proxy records.

Figure 5 compares the TR NH reconstructions with 10 CMIP5 ~~fully~~-coupled ocean-atmosphere global climate models (See Appendix Table A2 for information) for land only MJJA temperatures over the same 40-75°N spatial domain. The 20-year low-pass smoothed time-series are shown as anomalies with respect to 1400-1850. Overall, N-TREND2015 sits well within the range of the model ensemble back to ~1200 (Figure 5A). Very good agreement is observed from ~1500 to the early 19<sup>th</sup> century, although this partially reflects the common reference period for anomaly calculations. Prior to ~1500, decadal variance of the reconstruction is generally greater than the models with the 1430s and 1160s temperature estimates well above the model range. Warming from the 19<sup>th</sup> to 20<sup>th</sup> century is in general good agreement with most models although N-TREND2015 appears to show more prolonged cooler temperatures in the 19<sup>th</sup> century before the start of warming.

In two important periods of climate transition, the cooling from the ~~m~~Medieval ~~p~~Period (900-1250) to Little Ice Age (LIA - 1450-1850), and the post LIA warming to present, the TR reconstructions and model estimations differ. During the ~~m~~Medieval to LIA transition (Figure 5C), the amplitude of temperature change from warm to cool, as expressed by N-TREND2015, is greater than most models and other TR reconstructions. The only model that agrees with this reconstruction is CCSM4. Most model results produce a cooler medieval period relative to the 1450-1850 mean, more in-line with DWJ06 and SCH2015. However, for the LIA to 20<sup>th</sup> century transition (Figure 5D) N-TREND2015 agrees well with the models - ~~the exceptions being~~ except with CCSM4 and HADCM3 - which is likely due to differences in these models' response to anthropogenic aerosols (see Meehl et al. 2012 and Schurer et al. 2014). Overall, N-TREND2015, DWJ06 and SCH2015 share reasonable multi-decadal coherence ( $r = 0.58, 0.44$  and  $0.64$ , respectively) with the multi-model mean from 1251-1850 CE (inset table in Figure 5B) but correlations are substantially weaker for 850-1250 CE.

Comparing the spectral properties (Figure 5E) of the three TR reconstructions to the 10 CMIP5 models allows a general assessment of potential frequency biases. SCH2015 (using only MXD data) contains more high frequency variance compared to DWJ06 (using mostly RW data) with N-TREND2015 sitting in the middle. Both N-TREND2015 and SCH2015 sit within the spectral power range of the models at these high frequencies despite the fact that they show only modest coherence at high frequencies with NH

instrumental temperatures (Appendix Figure B2). Only by the addition of more long MXD data-sets from key areas currently lacking such high resolution archives (e.g. much of North America, eastern Eurasia) could the large-scale high frequency coherence be improved. At mid and lower frequencies, the spectral density of all three TR reconstructions sits within the model range.

Most attribution studies, and many proxy/model comparative studies (Hegerl et al. 2007; Mann et al. 2012), have focused on the post 1400 period, when relatively ample and well dated proxy data are available. Prior to that period there is a distinct mismatch between the proxies and models during the ~~medieval period~~ (Fernández-Donado et al. 2013; Schurer et al. 2014). ~~A~~ although the response to forcing is detectable in some reconstructions, ~~there is a tendency for less prominent forcing at this time which makes detection more difficult~~ (Schurer et al., 2013). Since N-TREND2015 ~~contains more data during the medieval period than~~ ~~—compared to~~ all previous TR reconstructions, ~~contains more data during the medieval period~~, it should provide ~~a~~ more robust ~~assessment estimates~~ of medieval summer season temperature. The relatively simple comparison shown in Figure 5 suggests medieval NH temperatures were likely warmer than ~~that expressed those simulated~~ by the CMIP5 models, although a detailed evaluation of this observation is beyond the scope of the current work. In a recent appraisal of proxy and CMIP5 model output over the Last Glacial Maximum, mid-Holocene, and the recent period, a substantial disagreement ~~of past climates between palaeoclimate and modelled estimates~~ was reported, although the models were marginally better at representing modern late Holocene climate (Harrison et al. 2015). However, underlying this comparison is the quality of the proxy data used and although N-TREND2015 represents an improvement on previous studies, more data from key regions such as North America and eastern Eurasia are needed.

### ***Proxy/model Comparison – volcanic forcing***

Coherency analysis (Appendix Figure B1) shows that large-scale TR composite reconstructions are generally poor at representing mean hemisphere-scale inter-annual variability, mainly because of the relatively small numbers of TR proxy series used for the hemispheric composites (see discussion in Diaz 1996 and Jones and Briffa 1996) and the biological memory effects in RW data (Esper et al. 2015). However, after major volcanic events, short term cooling can be much more spatially homogenous, allowing even relatively small networks of TR records to potentially capture the amplitude of cooling after such events (Hegerl et al. 2003; Schneider et al. 2015; Stoffel et al. 2015, but see Anchukaitis et al. 2012). Though it has long been known that TR records can reflect volcanically forced cool summer conditions (Briffa et al. 1998), it is less well known that MXD is the superior TR parameter for assessing such short term climate perturbations (Jones et al. 1995; Frank et al. 2007b; Anchukaitis et al. 2012; D'Arrigo et al. 2013; Esper et al. 2015). We briefly return to this issue by comparing the response of N-TREND2015, DWJ06, SCH2015 and the CMIP5 models to selected volcanic events. Esper et al. (2013) showed that the choice of candidate volcanic events can greatly influence ~~greatly~~ the observed estimated amount of cooling experienced on average over large regions. Herein, we use three different lists of ~~candidate significant~~ volcanic events (Table 3 – combined information from Gao et al. (2008) ~~and~~; Crowley and Unterman (20123); ~~Plummer et al. 2012~~; Esper et al. 2013; Sigl et al. 2013) to assess the mean response of each TR NH series and CMIP5 model to large volcanic events using superposed epoch analysis (SEA). It should be noted that the CMIP5 models ~~use were forced with~~ either the Gao et al. (2008) or Crowley and Unterman (20132) volcanic forcing series (Appendix Table A2).

Overall the mean TR proxy response is substantially muted using the Esper events (minimal cooling in year T+1) compared to the Gao/Crowley and ~~Sigl/Plummer~~Sigl lists (Figure 6). A muted response may

~~occur if some event dates are in error, conversely, an amplified response can occur if event dates are related, in a circular manner, to tree-ring evidence (Esper et al. 2013). The latter condition may in fact be applicable to some ice core records (e.g. the Gao/Crowley), which would partly explain the stronger response observed using such lists. We concede, the amplification effect may be limited to just a few events, such as the 1450s eruption(s), which are not firmly dated by any evidence (Plummer et al. 2012) other than tree-rings (Gao et al. 2006). This, however, is not the case for the Sigl et al. (2013) records. This observation is partly expected as some of the events in the latter lists were identified and/or dated using tree ring evidence and there is some degree of circularity.~~ Using the Gao/Crowley and ~~Sigl/Sigl/Plummer~~ lists, N-TREND2015 and SCH2015 express a maximum cooling in year T+1 of  $\sim 0.4-0.54$  °C while DWJ06 does not reach peak cooling until years T+5 or T+6 and with a more modest cooling of  $\sim 0.2-0.3$  °C. This lagged and muted response of the RW dominated DWJ06 reconstruction has been detailed elsewhere and is mainly related to autocorrelated biological memory in RW data (Krakauer and Randerson, 2003; Frank et al., 2007b; Anchukaitis et al. 2012; Esper et al. 2015). As N-TREND2015 includes RW as well as MXD data, there is potential for this biological persistence bias to influence the record's ability to express true large-scale temperature response (and recovery) to volcanic events. The lower panel of Figure 6 compares the mean response of the three reconstructions to the 10 CMIP5 simulations. For the Gao/Crowley ~~and Sigl~~ events, the models on average show marginally greater cooling ( $\sim 0.6-0.7$  °C) in year T+1 than N-TREND2015 and SCH2015, but similar amounts of cooling for the Esper ( $\sim 0.2$  °C) ~~events and Sigl/Plummer~~ ( $\sim 0.4$  °C) events. Most importantly, the rate of recovery after volcanic cooling is slower for N-TREND2015 than SCH2015, with the latter record agreeing well with the models—~~especially using the Gao/Crowley events.~~

Dividing the N-TREND2015 data-set into its three TR parameter types RW, MXD and MIX (Figure 7) and producing a parameter specific hemispheric mean composite from each, highlights the behaviours mentioned above, although the MIX data displays a slower recovery than RW when using the Gao/Crowley and ~~Sigl/Sigl/Plummer~~ event lists. Overall, the response of each parameter composite appears specific to different eruptions (Table 3, Figure 7). In 1258 (1257, Samalas eruption – Table 3), the MXD response is immediate, RW data is strongest in year T+1 and in the MIX composite at year T+2. The Kuwae eruptions of 1452 (~~not identified as a significant eruption event – Table 3~~) and 1458 are clearly captured by MXD data in 1453, and although cooling starts in 1453, peak cold conditions in the MIX series does not occur until 1459. All three parameter types start decreasing after the unknown 1809 event, but only MXD and MIX contain a defined cool peak in 1816. In 1883 (Krakatau), MXD and MIX respond similarly, with cold conditions in 1884, while in 1912 (~~NovaputraKatmai~~), the response is immediate in MXD and RW, but lagged by 1 year in MIX. By just looking at the response to individual volcanic eruptions, it is difficult to ascertain whether there are true lagged biases in the MIX composite compared to MXD. A more detailed analysis is needed to examine the spatiotemporal response of the individual TR parameters, which is beyond the scope of this paper but is partly addressed in Anchukaitis et al. (~~this volume in prep~~).

In a recently published new TR reconstruction of NH JJA mean temperatures, Stoffel et al. (2015) utilised 30 regional records (using both RW and MXD) to assess large-scale volcanic response expressed in their composite record by comparison to ~~a modified version of the IPSL-CM5A-LR model (Schmidt et al. 2012). Their model version simulated less cooling and they stated~~ that with respect to volcanic forcing their work had reconciled previously reported dissimilarities between climate simulations and TR based reconstructions. Although the “missing ring” hypothesis (Mann et al. 2012) is now thoroughly rejected (Anchukaitis et al. 2012; D'Arrigo et al. 2013; Esper et al. 2015), ~~Stoffel et al. (2015) and the results detailed herein, indicate that~~ there are still substantial uncertainties that ~~still~~ need to be addressed before we fully understand the temporal and spatial dynamical response of the climate system to volcanic forcing ~~as well as the response of specific TR parameter to volcanically forced cool summers.~~

## Conclusion

Over the last 20 years, palaeoclimatology has augmented and improved the available proxy data that can be used to produce large-scale, high resolution, reconstructions of the Earth's climate for the late Holocene. This improvement is clearly shown in synthesis figures of NH temperatures in the IPCC reports, which have moved from a single qualitative representation (FAR – Houghton et al. 1990), to a quantitative ensemble (AR4/5 – Masson-Delmotte et al., 2013 - see discussion in Frank et al. 2010). It is quite understandable that when few data existed, there was no real option but to amalgamate all possible records together, ignoring the local/regional seasonal signal they represented, or their quality and individual biases. However, substantial investment and progress has been made over the past decade and many high quality records now exist from different proxy archives for multiple locations around the NH (Jones et al. 2009; PAGES2 Consortium 2013). This data richness now allows the palaeoclimate-community to focus more specifically on the unique properties of different proxy archives allowing selection for optimising seasonal and temporal fidelity.

This paper (and its part II companion, Anchukaitis et al. *this volume in prep*), represents a tree-ring community initiative to improve the Northern Hemisphere-scale proxy network (temporally and spatially), and reconstruct NH summer temperatures back to the 8<sup>th</sup> century AD. Following recommendations detailed in Wilson et al. (2007), we have focused on using published tree-ring reconstructions which reflect a statistically and biologically reasonable amount of local/regional summer temperature variance, and have utilised a substantially increased data-base of records compared to previous studies (D'Arrigo et al. 2006; Frank et al. 2007a; Wilson et al. 2007; Schneider et al. 2015; Stoffel et al. 2015). Consequently, the increased replication in the N-TREND2015 reconstruction alone represents a substantial improvement over previous TR, NH millennial reconstructions and the final *outcome record* appears reasonably insensitive to the *statistical methods used to derive the record* (Appendix Figure C1). The N-TREND2015 reconstruction expresses reasonable fidelity, at all time-scales, from 918-2004 (Figure 2) and reasonably captures late 20<sup>th</sup> century warming. *However, Nevertheless*, N-TREND2015 is far from perfect!

How *good-well* should a TR reconstruction calibrate locally before it can be considered for inclusion in such a large-scale data-set? Wilson et al. (2007) specifically used a minimum *correlation* of 0.4 (only *2016%* explained variance) against local gridded *temperature* data, but this value is still rather low, and for realistic reconstructions of local, regional, and hemisphere-scale temperatures, greater fidelity should be required, especially when the number of input series is modest. Of the records used herein, local scale calibration  $r^2$  values range from ~10-70% explained variance (Figure 1) with all RW based records explaining less than 40% variance. Considering *as well*, the well-known *lagged* high frequency biases of RW data due to biological persistence (Krakauer and Randerson, 2003; Frank et al., 2007b; Anchukaitis et al. 2012; Esper et al. 2015), it could be argued that RW derived reconstructions should not be deemed robust estimates of local temperatures without the inclusion of (or replacement with) MXD or BI data. *We admit that -* This is a rather contentious statement but it is *particularly-relevant more defensible* if the study focus is on addressing past climate response to volcanic forcing. At the very least, whether RW is used alone or in combination with MXD/BI data, local calibration needs to express "reasonable" fidelity and multiple statistical metrics exist to assess reconstruction quality (e.g.  $r^2$ , RE, CE and residual analysis – Cook et al. 1994; Wilson et al. 2006; Macias-Fauria et al. 2012). As a minimum, local climate calibration for any single TR parameter or combination of multiple TR parameters must be statistically significant with no significant long-term trends within the *model-calibration* residuals. Any TR records expressing local based divergence (D'Arrigo et al. 2008) should not be considered in a large composite database unless the cause of the

divergence can be truly identified as unique to the recent period (Wilson et al. 2007). ~~This situation can potentially be further confused if~~ ~~To potentially confuse this latter situation, however, is~~ the quality and/or homogeneity of the instrumental data ~~where homogeneity~~ cannot be assumed to be consistent back in time (Parker 1994; Frank et al. 2007b). Correlations between some TR records and pre-1950 gridded data weaken considerably, especially for locations in central Asia (Table 1 and Cook et al. 2012), which make defining an acceptable calibration threshold difficult.

Although the number of TR records used in N-TREND2015 represents a substantial increase over previous TR NH composites (D'Arrigo et al. 2006; Frank et al. 2007a; Wilson et al. 2007; Schneider et al. 2015; Stoffel et al. 2015), there are still vast regions of the NH that are poorly represented (see Figure 1 and discussion in Anchukaitis et al. (~~this volume in prep~~)). Only 3 millennial length temperature-sensitive TR series exist for North America and only one of those incorporates MXD. Very few long TR records have been developed for eastern Eurasia and no millennial long-length MXD series exist between 90°E and 170°E. ~~We appreciate that wW~~here forests grow in non-temperature limiting environments (e.g. mid latitude eastern North America, Siberian Steppes, low elevation central Europe), it may never be possible to develop temperature-sensitive chronologies using traditional TR parameters, ~~but. However,~~ stable isotopes may help in some regions (Young et al. 2012). Nevertheless, there are substantial swaths of high latitude boreal forest and high altitude environments where tree-growth is predominantly limited by summer temperatures, and a strategically focussed sample collection in these regions (including sub-fossil material to extend living data-sets) is needed.

N-TREND2015 represents the latest iteration of ~~the~~ tree-ring based NH reconstructions. It will not be the last, but we hope that the creation of the N-TREND consortium will facilitate further expansion of the NH tree-ring network. A strategic plan is needed to update existing records to the present and develop new long millennial records from regions where little or no data exists. It is also paramount that MXD or BI records are concurrently developed with RW data to produce locally, as well as regionally, robust temperature reconstructions. As the network expands, further assessment of the unique properties of RW, MXD and BI can be performed that should lead towards even further, substantial refinement of large-scale temporal and spatial scale summer temperature reconstructions.

## Acknowledgement

The N-TREND consortium is not specifically funded, but many individuals acknowledge relevant funds; RW: UK Natural Environment Research Council (NERC - NE/K003097/1) and Leverhulme Trust project (F/00 268/BG); KJA: National Science Foundation - NSF AGS-1501856 and NSF AGS-1501834; KB and TO: [NERC \(NE/G018863/1\)](#); RD: National Science Foundation Paleoclimatic Perspectives of Climatic Change; SH: Academy of Finland; GH and AS: ERC advanced grant TITAN (EC-320691), NCAS and GH specifically with a Royal Society Wolfson Research Merit Award; HL: The Swedish Science Council (VR); 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 reconstructions can be found at: <https://ntrenddendro.wordpress.com/>. Lamont-Doherty Earth Observatory contribution No. XXXX.

## References

Formatted: Superscript

Ammann, C.~~M.~~, Joos, F., Schimel, D. ~~S.~~, Otto-Bliesner, B. ~~L.~~, & Tomas, R. ~~A.~~, 2007. Solar influence on climate during the past millennium: results from transient simulations with the NCAR Climate System Model. *Proc. Natl. Acad. Sci. U.S.A.*, 104(10), 3713–8.

Anchukaitis, K.~~J.~~, D'Arrigo, R.~~D.~~, Andreu-Hayles, L., Frank, D., Verstege, A., Buckley, B.~~M.~~, Curtis, A., Jacoby, G.~~C.~~, Cook, E.~~R.~~, 2013. Tree-ring reconstructed summer temperatures from northwestern North America during the last nine centuries. *Journal of Climate* 26, 10, 3001–3012, doi: 10.1175/JCLI-D-11-00139.1

Anchukaitis, K.~~J.~~, Breitenmoser, P., Briffa, K.~~R.~~, Buchwal, A., Büntgen, U., Cook, E.~~R.~~, D'Arrigo, R.~~D.~~, Esper, J., Evans, M.~~N.~~, Frank, D., Grudd, H., Gunnarson, B., Hughes, M.~~K.~~, Kirilyanov, A.~~V.~~, Körner, C., Krusic, P., Luckman, B., Melvin, T.~~M.~~, Salzer, M.~~W.~~, Shashkin, A.~~V.~~, Timmreck, C., Vaganov, E.~~A.~~, Wilson, R., 2012. Tree rings and volcanic cooling. *Nature Geoscience* 5, 836– 837, doi:10.1038/ngeo1645

[Anchukaitis, K., Wilson, R., Briffa, K., Büntgen, U., Cook, E., D' Arrigo, R., Davi, N., Esper, J., Frank, D., Gunnarson, B., Hegerl, G., Helama, S., Klesse, S., Krusic, P., Linderholm, H., Myglan, V., Osborn, T., Rydval, M., Schneider, L., Schurer, A., Wiles, G., Zhang, P. and Zorita, E. Last millennium Northern Hemisphere summer temperatures from tree rings: Part I: the spatial context. In Preparation. To be submitted to \*Quaternary Science Reviews\*.](#)

Biondi, F., Perkins, D.~~L.~~, Cayan, D.~~R.~~, Hughes, M.~~K.~~, 1999. July temperature during the second millennium reconstructed from Idaho tree rings. *Geophysical Research Letter* 26, 1445-1448.

Björklund, J.~~A.~~, Gunnarson, B.~~E.~~, Seftigen, K., Esper, J., Linderholm, H.~~W.~~, 2014. Blue intensity and density from northern Fennoscandian tree rings, exploring the potential to improve summer temperature reconstructions with earlywood information. *Climate of the Past* 10, 877-885, doi: 10.5194/cp-10-877-2014

Böhm, R., Jones, P.~~D.~~, Hiebl, J., Frank, D., Brunetti, M., Maugeri, M., 2010. The early instrumental warm-bias: a solution for long central European temperature series 1760 – 2007. *Climatic Change* 101(1-2), 41-67.

Briffa, K.~~R.~~, Jones, P.~~D.~~, Schweingruber, F.~~H.~~, Karlen, W., Shiyatov, G., 1996. Tree ring variables as proxy climate indicators: problems with low-frequency signals, in: Jones, P.D., Bradley, R.S., Jouzel, J. (Eds.), *Climatic Variations and Forcing Mechanisms of the Last 2000 Years*. Springer-Verlag, Berlin, pp. 9-41.

Briffa, K., Jones, P., Schweingruber, F., Osborn, T., 1998. Influence of volcanic eruptions on Northern Hemisphere summer temperature over the past 600 years. *Nature* 393, 450–455, doi:10.1038/30943

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.

Briffa, K.~~R.~~, Osborn, T.~~J.~~, Schweingruber, F.~~H.~~, Jones, P.~~D.~~, Shiyatov, S.~~G.~~, Vaganov, E.~~A.~~, 2002a. Tree-ring width and density data around the Northern Hemisphere: Part 1, Local and regional climate signals. *The Holocene* 12(6), 737–757.

Briffa, K.R., Osborn, T.J., Schweingruber, F.H., Jones, P.D., Shiyatov, S.G., Vaganov, E.A., 2002b. Tree-ring width and density data around the Northern Hemisphere: Part 2, spatio-temporal variability and associated climate patterns. *The Holocene* 12(6), 759-789.

Briffa, K.R., Melvin, T.M., Osborn, T.J., Hantemirov, R.M., Kirilyanov, A.V., Mazepa, V., Shiyatov, S.G., Esper, J. 2013. Reassessing the evidence for tree-growth and inferred temperature change during the Common Era in Yamalia, northwest Siberia. *Quaternary Science Reviews* 72, 83-107. doi: 10.1016/j.quascirev.2013.04.008

Büntgen, U., Frank, D.C., Nievergelt, D., Esper, J., 2006. Summer temperature variations in the European Alps, AD 755-2004. *Journal of Climate* 19(21), 5606-5623.

Büntgen, U., Kyncl, T., Ginzler, C., Jacks, D.S., Esper, J., Tegel, W., Heussner, K.U., Kyncl, J., 2013. Filling the Eastern European gap in millennium-long temperature reconstructions. *Proceedings of the National Academy of Science, USA* 110, 1773-1778.

Campbell, R., McCarroll, D., Robertson, I., Loader, N.J., Grudd, H., Gunnarson, B.E., 2011. Blue intensity in *Pinus sylvestris* tree rings: A manual for a new palaeoclimate proxy. *Tree-Ring Research* 67(2), 127-134.

Cohen, J.L., Furtado, J.C., Barlow, M., Alexeev, V.A., Cherry, J.E., 2012. Asymmetric seasonal temperature trends. *Geophysical Research Letters* 39(4), doi: 10.1029/2011GL050582

Cook, E. R., Peters, K., 1981. The smoothing spline: A new approach to standardizing forest interior tree-ring width series for dendroclimatic studies. *Tree-Ring Bulletin* 41, 45-53.

Cook, E., Kairiukstis, L., 1990. *Methods of Dendrochronology*, Springer, New York.

Cook, E.R., Briffa, K.R., Jones, P.D., 1994. Spatial regression methods in dendroclimatology: a review and comparison of two techniques. *International Journal of Climatology* 14(4), 379-402.

Cook, E.R., Briffa, K.R., Meko, D.M., Graybill, D.A., Funkhouser, G., 1995. The 'segment length curse' in long tree-ring chronology development for palaeoclimatic studies. *The Holocene* 5, 229-237.

Cook, E.R., D'Arrigo, R., Mann, M.E., 2002. A well-verified, multiproxy reconstruction of the winter North Atlantic Oscillation index since AD 1400. *Journal of Climate* 15, 1754-1764.

Cook, E.R., Krusic, P.J., Anchukaitis, K.J., Buckley, B.M., Nakatsuka, T., Sano, M., PAGES Asia2k Members, 2012. Tree-ring reconstructed boreal summer temperature anomalies for temperate East Asia since 800 C.E.. *Climate Dynamics* 41(11-12), 2957-2972, doi: 10.1007/s00382-012-1611-x

Crowley, T.J., 2000. Causes of Climate Change Over the Past 1000 Years. *Science*, 289(5477), 270-277. <http://doi.org/10.1126/science.289.5477.270>

[Crowley, T.J., Unterman, M.B., 2013. Technical details concerning development of a 1200 yr proxy index for global volcanism. \*Earth System Science Data\* 5, 187-197, doi:10.5194/essd-5-187-2013](#)

Cunningham, L.~~K~~, Austin, W.~~E.N~~, Knudsen, K.~~L~~, Eiriksson, J., Scourse, J.~~D~~, Wanamaker Jr., A.~~D~~, Butler, P.~~G~~, Cage, A.~~G~~, Richter, T., Husum, K., Hald, M., Andersson, C., Zorita, E., Linderholm, H.W., Gunnarson, B.E., Sicre, M.-A., Sejruo, H.P., Jiang, H., Wilson, R., 2013. Reconstructions of surface ocean conditions from the northeast Atlantic and Nordic Seas during the last millennium. *The Holocene* 23, 921–935, doi:10.1177/0959683613479677

D'Arrigo, R., Buckley, B., Kaplan, S., Woollett, J., 2003. Interannual to multidecadal modes of Labrador climate variability inferred from tree rings. *Climate Dynamics* 20(2-3), 219–228.

D'Arrigo, R., Mashig, E., Frank, D., Jacoby, G., Wilson, R., 2004. Reconstructed Warm Season Temperatures For Nome, Seward Peninsula, Alaska. *Geophysical Research Letters* 31(9), doi: 10.1029/2004GL019756

D'Arrigo, R., Wilson, R., Jacoby, G., 2006. On the long-term context for late 20th century warming. *Journal of Geophysical Research* (1984-2012) 111(D3), doi: 10.1029/2005JD006352

D'Arrigo, R., Wilson, R., Liepert, B., Cherubini, P., 2008. On the 'Divergence Problem' in Northern Forests: A Review of the Tree-Ring Evidence and Possible Causes. *Global and Planetary Change* 60, 289–305, doi: 10.1016/j.gloplacha.2007.03.004

D'Arrigo, R., Wilson, R., Anchukaitis, K., 2013. Volcanic Cooling Signal in Tree-Ring Temperature Records for the Past Millennium. *Journal of Geophysical Research: Atmospheres* 118(16), 9000-9010, doi: 10.1002/jgrd.50692

D'Arrigo, R., Wilson, R., Wiles, G., Anchukaitis, K., Solomina, O., Davi, N., Deser, C., Dolgova, E., 2014. Tree-ring reconstructed temperature index for coastal northern Japan: implications for western North Pacific variability. *International Journal of Climatology*, doi: 10.1002/joc.4230

Davi, N., Jacoby, G., Wiles, G., 2003. Boreal temperature variability inferred from maximum latewood density and tree-ring width data, Wrangell Mountain region, Alaska. *Quaternary Research* 60, 252–262.

Davi, N., D'Arrigo, R., Jacoby, G., Cook, E.R., Anchukaitis, K., Nachin, B., Rao, M.P., Leland, C., 2015. A long-term context (931–2005 C.E.) for rapid warming over Central Asia. *Quaternary Science Reviews* 121, 89-97, doi: 10.1016/j.quascirev.2015.05.020

Deser, C., Knutti, R., Solomon, S., Phillips, A.S., 2012. Communication of the role of natural variability in future North American climate. *Nature Climate Change* 2, 775–779.

Deslauriers, A., Anfodillo, T., Rossi, S., Carraro, V., 2007. Using simple causal modeling to understand how water and temperature affect daily stem radial variation in trees. *Tree Physiology* 27(8), 1125–1136.

Diaz, H.~~F~~, 1996. Temperature changes on long time and large spatial scales: inferences from instrumental and proxy records, in: Jones, P.D., Bradley, R.S., Jouzel, J. (Eds.), *Climate variations and forcing mechanisms of the last 2,000 years*. Springer Berlin Heidelberg, pp 585–601.

Divine, D., Isaksson, E., Martma, T., Meijer, H.A.J., Moore, J., Pohjola, V., van de Wal, R.S.W., Godtlielsen, F., 2011. Thousand years of winter surface air temperature variations in Svalbard and northern Norway reconstructed from ice core data. *Polar Research* 30, doi: 10.3402/polar.v30i0.7379

Dorado-Linan, I., Büntgen, U., Gonzalez-Rouco, F., Zorita, E., Montavez, J.P., Gomez-Navarro, J.J., Brunet, M., Heinrich, I., Helle, G., Gutierrez, E., 2012. Estimating 750 years of temperature variations and uncertainties in the Pyrenees by tree-ring reconstructions and climate simulations. *Climate of the Past* 8, 919-933.

Edwards, T.L., Crucifix, M., Harrison, S.P., 2007. Using the past to constrain the future: how the palaeorecord can improve estimates of global warming. *Progress in Physical Geography* 31 (5), 481-500.

England, M.H., McGregor, S., Spence, P., Meehl, G.A., Timmermann, A., Cai, W., Gupta, A.S., McPhaden, M.J., Purich, A., Santoso, A., 2014. Recent intensification of wind-driven circulation in the Pacific and the ongoing warming hiatus. *Nature Climate Change* 4, 222–227, doi: 10.1038/nclimate2106

Esper, J., Cook, E.R., Schweingruber, F., 2002. Low-frequency signals in long tree-ring chronologies for reconstructing past temperature variability. *Science* 295, 2250-2253.

Esper, J., Cook, E., Peters, K., Krusic, P., Schweingruber, F., 2003. Detecting low-frequency tree-ring trends by the RCS method. *Tree-Ring Research* 59, 81–98.

Esper, J., Frank, D., Wilson, R., 2004. Climate Reconstructions: Low-Frequency Ambition and High-Frequency Ratification. *Eos, Transactions American Geophysical Union* 85(12), 113-120.

Esper, J., Frank, D.C., Büntgen, U., Verstege, A., Luterbacher, J., Xoplaki, E., 2007. Long-term drought severity variations in Morocco. *Geophysical Research Letters* 34, doi: 10.1029/2007GL030844

Esper, J., Schneider, L., Krusic, P.J., Luterbacher, J., Büntgen, U., Timonen, M., Sirocko, F., Zorita, E., 2013. European summer temperature response to annually dated volcanic eruptions over the past nine centuries. *Bulletin of Volcanology* 75(7), 1-14, doi: 10.1007/s00445-013-0736-z

Esper, J., DÜthorn, E., Krusic, P., Timonen, M., Büntgen, U., 2014. Northern European summer temperature variations over the Common Era from integrated tree-ring density records. *Journal of Quaternary Science* 29, 487-494.

Esper, J., Schneider, L., Smerdon, J., Schöne, B., Büntgen, U., [in press 2015](#). Signals and memory in tree-ring width and density data. *Dendrochronologia*, [35, 62-70](#).

Fernández-Donado, L., González-Rouco, J.F., Raible, C.C., Ammann, C.M., Barriopedro, D., García-Bustamante, E., Jungclauss, J.H., Lorenz, S.J., Luterbacher, J., Phipps, S.J., Servonnat, J., Swingedouw, D., Tett, S.F.B., Wagner, S., Yiou, P., Zorita, E., 2013. Large-scale temperature response to external forcing in simulations and reconstructions of the last millennium. *Climate of the Past* 9, 393–421, doi: 10.5194/cp-9-393-2013

Frank, D., Esper, J., Cook, E.R., 2007a. Adjustment for proxy number and coherence in a large-scale temperature reconstruction. *Geophysical Research Letters* 34, doi:10.1029/2007GL030571

Frank, D., Büntgen, U., Böhm, R., Maugeri, M., Esper, J., 2007b. Warmer early instrumental measurements versus colder reconstructed temperatures: shooting at a moving target. *Quaternary Science Reviews* 26(25), 3298-3310

Frank, D., Esper, J., Zorita, E., Wilson, R., 2010. A noodle, hockey stick, and spaghetti plate: a perspective on high resolution paleoclimatology. *Wiley Interdisciplinary Reviews Climate Change* 1(4), 507-516, doi: 10.1002/wcc.53

Frank, J., Frank, D., Raible, C., Esper, J., 2013. Spectral biases in tree-ring climate proxies. *Nature Climate Change* 3, 360-364.

Fritts, H.C., 1976. *Tree Rings and Climate*. Academic Press, San Diego, Calif..

[Gao, C., A. Robock, S. Self, J. B. Witter, J. P. Steffenson, H. B. Clausen, M.-L. Siggaard-Andersen, S. Johnsen, P. A. Mayewski, and C. Ammann \(2006\), The 1452 or 1453 A.D. Kuwae eruption signal derived from multiple ice core records: Greatest volcanic sulfate event of the past 700 years, \*J. Geophys. Res.\*, 111, D12107, doi:10.1029/2005JD006710](#)

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, doi: 10.1029/2008JD010239

Guo, Z., and Zhou, T., 2013. Why does FGOALS-gl reproduce a weak Medieval Warm Period but a reasonable Little Ice Age and 20th century warming? *Adv. Atmos. Sci.*, 30(6), 1758–1770. <http://doi.org/10.1007/s00376-013-2227-8>

Gennaretti, F., Arseneault, D., Nicault, A., Perreault, L., Bégin, Y., 2014. Volcano-induced regime shifts in millennial tree-ring chronologies from Northeastern North America. *Proceedings of the National Academy of Sciences, USA* 111, 10077–10082.

Gouirand, I., Linderholm, H., Moberg, A., Wohlfarth, B., 2008. On the spatio-temporal characteristics of Fennoscandian tree-ring based summer temperature reconstructions. *Theoretical and Applied Climatology* 91, doi: 10.1007/s00704-007-0311-7

Hansen, J., Ruedy, R., Sato, M., Lo, K., 2010. Global surface temperature change. *Reviews of Geophysics* 48(4), RG4004, doi:10.1029/2010RG000345

Harris, I., Jones, P., Osborne, T., Lister, D., 2014. Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset. *International Journal of Climatology* 34(3), 623–642, doi: 10.1002/joc.3711

Harrison, S., Bartlein, P., Izumi, K., Li, G., Annan, J., Hargreaves, J., Braconnot, P. & Kageyama, M. (2015). Evaluation of CMIP5 palaeo-simulations to improve climate projections. *Nature Climate Change*, 5(8), 735-743.

Hegerl G., Crowley T., Baum S., Yul Kim K, Hyde W., 2003. Detection of volcanic, solar and greenhouse gas signals in paleo-reconstructions of Northern Hemispheric temperature. *Geophysical Research Letters* 30:461–464

Hegerl, G.~~C.~~, Crowley, T.~~J.~~, Hyde, W.~~F.~~, Frame, D.~~J.~~, 2006. Climate sensitivity constrained by temperature reconstructions over the past seven centuries. *Nature* 440, 1029-1032, doi: 10.1038/Nature 04679

Hegerl, G.~~C.~~, Crowley, T.~~J.~~, Allen, M., Hyde, W.~~F.~~, Pollack, H.~~N.~~, Smerdon, J., Zorita, E., 2007. Detection of human influence on a new, validated 1500-year temperature reconstruction. *Journal of Climate* 20(4), 650–666.

Helama, S., Timonen, M., Holopainen, J., Ogurtsov, M.~~G.~~, Mielikäinen, K., Eronen, M., Lindholm, M., Meriläinen, J., 2009a. Summer temperature variations in Lapland during the Medieval Warm Period and the Little Ice Age relative to natural instability of thermohaline circulation on multi-decadal and multi-centennial scales. *Journal of Quaternary Science* 24, 450-456.

[Helama, S., Makarenko, N.G., Karimova, L.M., Kruglun, O.A., Timonen, M., Holopainen, J., Meriläinen, J., Eronen, M., 2009b. Dendroclimatic transfer functions revisited: Little Ice Age and Medieval Warm Period summer temperatures reconstructed using artificial neural networks and linear algorithms. \*Annales Geophysicae\* 27 \(3\), 1097-1111, doi: 10.5194/angeo-27-1097-2009](#)

Helama, S., Vartiainen, M., Holopainen, J., Mäkelä, H.M., Kolström, T., Meriläinen, J., 2014. A palaeotemperature record for the Finnish Lakeland based on microdensitometric variations in tree rings. *Geochronometria* 41(3), 265-277.

Hughes, M.~~K.~~, Vaganov, E.~~A.~~, Shiyatov, S., Touchan, R., Funkhouser, G., 1999. Twentieth-century summer warmth in northern Yakutia in a 600 year context. *The Holocene* 9, 603– 608.

Hughes, M.~~K.~~, 2011. Dendroclimatology in High-Resolution Paleoclimatology, in: Hughes, M.K., Swetnam, T.W., Diaz, H.F. (Eds.), *Dendroclimatology*. Springer, Netherlands, pp. 17-34.

Jacoby, G.~~C.~~, Lovelius, N.~~V.~~, Shumilov, O.~~I.~~, Raspopov, O.~~M.~~, Karbainov, J.~~M.~~, Frank, D.~~C.~~, 2000. Long-term temperature trends and tree growth in the Taymir region of northern Siberia. *Quaternary Research* 53(3), 312– 318.

Jones, P.~~D.~~, Briffa, K.~~R.~~, Schweingruber, F.~~H.~~, 1995. Tree-ring evidence of the widespread effects of explosive volcanic eruptions. *Geophysical Research Letters* 22, 1333–1336.

Jones, P.~~D.~~, Briffa, K.~~R.~~, 1996. What can the instrumental record tell us about longer timescale paleoclimatic reconstructions? in: Jones, P.D., Bradley, R.S., Jouzel, J. (Eds.). *Climatic Variations and Forcing Mechanisms of the Last 2000 Years*. Springer, Berlin, pp. 625–644.

Jones, P.~~D.~~, Briffa, K.~~R.~~, Barnett, T.~~P.~~, Tett, S.~~F.B.~~, 1998. High-resolution paleoclimatic records for the last millennium: Interpretation, integration and comparison with general circulation model control-run temperatures. *The Holocene*, 8(4), 455– 471.

Jones, P.~~D.~~, Briffa, K.~~R.~~, Osborn, T.~~J.~~, Lough, J.~~M.~~, van Ommen, T.~~D.~~, Vinther, B.~~M.~~, Luterbacher, J., Wahl, E.~~R.~~, Zwiars, F.~~W.~~, Mann, M.~~E.~~, Schmidt, G.~~A.~~, Ammann, C.~~M.~~, Buckley, B.~~M.~~, Cobb, K.~~M.~~, Esper, J., Goosse, H., Graham, N., Jansen, E., Kiefer, T., Kull, C., Küttel, M., MosleyThompson, E., Overpeck, J.~~T.~~, Riedwyl, N., Schulz, M., Tudhope, A.~~W.~~, Villalba, R., Wanner, H., Wolff, E., Xoplaki, E., 2009. High-resolution

palaeoclimatology of the last millennium: A review of current status and future prospects. *The Holocene*, 19, 3–49.

Jones, P.~~D.~~, Lister, D.~~H.~~, Osborn, T.~~J.~~, Harpham, C., Salmon, M., Morice, C.P., 2012. Hemispheric and large-scale land surface air temperature variations: an extensive revision and an update to 2010. *Journal of Geophysical Research* 117, D05127, doi: 10.1029/2011JD017139

Jungclaus, J.~~H.~~, Lohmann, K., and Zanchettin, D., 2014. Enhanced 20th-century heat transfer to the Arctic simulated in the context of climate variations over the last millennium. *Climate of the Past*, 10(6), 2201–2213. <http://doi.org/10.5194/cp-10-2201-2014>

Kaplan, J.~~O.~~, Krumhardt, K.~~M.~~, Ellis, E.~~C.~~, Ruddiman, W.~~F.~~, Lemmen, C., and Goldewijk, K.~~K.~~, 2010. Holocene carbon emissions as a result of anthropogenic land cover change. *Holocene*, 21(5), 775–791. <http://doi.org/10.1177/0959683610386983>

Karl, T.~~R.~~, Arguez, A., Huang, B., Lawrimore, J.~~H.~~, McMahon, J.~~R.~~, Menne, M.~~J.~~, Peterson, T.~~C.~~, Vose, R.~~S.~~, Zhang, H.~~M.~~, 2015. Possible artifacts of data biases in the recent global surface warming hiatus. *Science* 348(6242), 1469–1472.

Klesse, S., Ziehmer, M., Trouet, V., Frank, D.C., 2014. Synoptic drivers of 400 years of summer temperature and precipitation variability on Mt. Olympus, Greece. *Climate Dynamics* 45(3-4), 807–824, doi: 10.1007/s00382-014-2313-3

Korner, C., 2003. Carbon limitation in trees. *Journal of Ecology*, 91(1), 4–17.

Kosaka, Y., Xie, S.P., 2013. Recent global warming hiatus tied to equatorial Pacific surface cooling. *Nature* 501, 403–407.

Krakauer, N.~~Y.~~, Randerson, J.~~F.~~, 2003. Do volcanic eruptions enhance or diminish net primary production? Evidence from tree rings. *Global Biogeochemical Cycles* 17(4), doi: 10.1029/2003GB002076

Landrum, L., Otto-Bliesner, B. L., Wahl, E. R., Conley, A., Lawrence, P. ~~J.~~, Rosenbloom, N., & Teng, H., 2013. Last Millennium Climate and Its Variability in CCSM4. *Journal of Climate*, 26(4), 1085–1111. <http://doi.org/10.1175/JCLI-D-11-00326.1>

Lawrimore, J.~~H.~~, Menne, M.~~J.~~, Gleason, B.~~E.~~, Williams, C.~~H.~~, Wuertz, D.~~B.~~, Vose, R.~~S.~~, Rennie, J., 2011. An overview of the Global Historical Climatology Network monthly mean temperature data set, version 3. *Journal of Geophysical Research: Atmospheres* (1984–2012) 116, D19121, doi:10.1029/2011JD016187

Linderholm, H.~~W.~~, Björklund, J., Seftigen, K., Gunnarson, B.~~E.~~, Fuentes, M., 2014. Fennoscandia revisited: a spatially improved tree-ring reconstruction of summer temperatures for the last 900 years. *Climate Dynamics* 45(3-4), 933–947, doi: 10.1007/s00382-014-2328-9

Ljungqvist, F.~~C.~~, 2010. A new reconstruction of temperature variability in the extra-tropical Northern Hemisphere during the last two millennia. *Geografiska Annaler: Series A, Physical Geography* 92(3), 339–351.

Ljungqvist, F.~~G.~~, Krusic, P.~~J.~~, Brattström, G., Sundqvist, H.~~S.~~, 2012. Northern Hemisphere temperature patterns in the last 12 centuries. *Climate of the Past* 8, 227-249.

Loader, N.~~J.~~, Young, G.~~H.F.~~, Grudd, H., McCarroll, D., 2013. Stable carbon isotopes from Torneträsk, northern Sweden provide a millennial length reconstruction of summer sunshine and its relationship to Arctic circulation. *Quaternary Science Reviews* 62, 97-113.

Luckman, B.~~H.~~, Wilson, R., 2005. Summer temperature in the Canadian Rockies during the last millennium – a revised record. *Climate Dynamics* 24, 131-144, doi:10.1007/s00382-004-0511-0

Macias-Fauria M., Grinsted, A., Helama, S., Holopainen, J., 2012. Persistence matters: Estimation of the statistical significance of paleoclimatic reconstruction statistics from autocorrelated time series. *Dendrochronologia* 30, 179-187.

Mann, M.~~E.~~, Bradley, R.~~S.~~, Hughes, M.~~K.~~, 1998. Global-scale temperature patterns and climate forcing over the past six centuries. *Nature* 392(6678), 779-787.

Mann, M.~~E.~~, Bradley, R.~~S.~~, Hughes, M.~~K.~~, 1999. Northern hemisphere temperatures during the past millennium: Inferences, uncertainties and limitations. *Geophysical Research Letters* 26, 759–762.

Mann, M.E., Zhang, Z., Hughes, M.K., Bradley, R.S., Miller, S.K., Rutherford, S., Ni, F., 2008. Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia. *Proceedings of the National Academy of Sciences, USA* 105(36), doi: 10.1073/pnas.0805721105

[Mann, M., Zhang, Z., Rutherford, S., Bradley, R., Hughes, M., Shindell, D., Ammann, C., Faluvegi, G. and Ni, F. 2009. Global signatures and dynamical origins of the Little Ice Age and Medieval Climate Anomaly. \*Science\* 326 \(5957\): 1256-1260.](#)

Mann, M.~~E.~~, Fuentes, J.~~D.~~, Rutherford, S., 2012. Underestimation of volcanic cooling in tree-ring-based reconstructions of hemispheric temperatures. *Nature Geoscience*. 5(3), 202–205.

Masson-Delmotte, V., Schulz, M., Abe-Ouchi, A., Beer, J., Ganopolski, A., González Rouco, J.F., Jansen, E., Lambeck, K., Luterbacher, J., Naish, T., Osborn, T., Otto-Bliesner, B., Quinn, T., Ramesh, R., Rojas, M., Shao, X., Timmermann, A., 2013. Information from Paleoclimate Archives, in: Stocker, T.F., Qin, D., Plattner, G.K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M. (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*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 383–464, doi:10.1017/CBO9781107415324.013

Matskovsky, V.~~V.~~, Helama, S., 2014. Testing long-term summer temperature reconstruction based on maximum density chronologies obtained by reanalysis of tree-ring data sets from northernmost Sweden and Finland. *Climate of the Past* 10(4), 1473-1487.

McCarroll, D., Pettigrew, E., Luckman, A., Guibal, F., Edouard, J.L., 2002. Blue reflectance provides a surrogate for latewood density of high-latitude pine tree rings. *Arctic, Antarctic, and Alpine Research* 34, 450–453.

McCarroll, D., Loader, N.J., Jalkanen, R., Gagen, M.H., Grudd, H., Gunnarson, B.E., Kirchhefer, A.J., Friedrich, M., Linderholm, H.W., Lindholm, M., Boettger, T. Los, S.O., Remmele, S., Kononov, Y.M., Yamazaki, Y.H., Young, G.H.F., Zorita, E., 2013. A 1200-year multiproxy record of tree growth and summer temperature at the northern pine forest limit of Europe. *The Holocene* 23(4), 471-484.

Meehl, G.A., Arblaster, J.M., Fasullo, J.F., Hu, A., Trenberth, K.E., 2011. Model-based evidence of deep-ocean heat uptake during surface-temperature hiatus periods. *Nature Climate Change* 1, 360–364.

Meehl, G.A., Washington, W.M., Arblaster, J.M., Hu, A., Teng, H., Tebaldi, C., Sanderson, B.N., Lamarque, J.-F., Conley, A., Strand, W.G., White III, J.B., 2012. Climate System Response to External Forcings and Climate Change Projections in CCSM4. *Journal of Climate* 25(11), 3661–3683, doi:<http://dx.doi.org/10.1175/JCLI-D-11-00240.1>

~~D.~~Meko, D., 1997. Dendroclimatic reconstruction with time varying predictor subsets of tree indices. *Journal of Climate*, 10:687–696, ~~1997~~.

Moberg, A., Sonechkin, D., Homgren, K., Datsenko, N., Karlen, W., 2005. Highly variable Northern Hemisphere temperatures reconstructed from low-and high-resolution proxy data. *Nature* 433, 613–617, doi: 10.1038/nature03265

Oglesby, R., Feng, S., Hu, Q., Rowe, C., 2012. The role of the Atlantic Multidecadal Oscillation on medieval drought in North America: Synthesizing results from proxy data and climate models. *Global and Planetary Change* 84-85: 56–65.

Ortega, P., Lehner, F., Swingedouw, D., Masson-Delmotte, V., Raible, C.C., Casado, M., Yiou, P., 2015. A model-tested North Atlantic Oscillation reconstruction for the past millennium. *Nature* 523, 71-74.

Osborn, T.J., Briffa, K.R., 2000. Revisiting timescale-dependent reconstruction of climate from tree-ring chronologies. *Dendrochronologia* 18, 9-25.

PAGES 2k Consortium, 2013. Continental-scale temperature variability during the past two millennia. *Nature Geoscience* 6, 339–346, doi:10.1038/ngeo1797

PAGES2k-PMIP3 group, 2015. Continental-scale temperature variability in PMIP3 simulations and PAGES 2k regional temperature reconstructions over the past millennium, *Climate of the Past Discussion* 11, 2483-2555, doi:10.5194/cpd-11-2483-2015

Palacio, S., Hoch, G., Sala, A., Körner, C., Millard, P., 2014. Does carbon storage limit tree growth?. *New Phytologist* 201(4), 1096-1100.

Parker, D.E., 1994. Effects of changing exposure of thermometers at land stations. *International Journal of Climatology* 14(1), 1–31.

Payette, S., 2007. Contrasted dynamics of northern Labrador tree lines caused by climate change and migrational lag. *Ecology* 88, 770– 780.

[Plummer C, Curran M, van Ommen T, Rasmussen S, Moy A, Vance T, Clausen H, Vinther B and Mayewski P. 2012. An independently dated 2000-yr volcanic record from Law Dome, East Antarctica, including a new perspective on the dating of the 1450s CE eruption of Kuwae, Vanuatu. \*Climate of the Past\*, 8, 1929-1940, doi:10.5194/cp-8-1929-2012](#)

Pongratz, J., Reick, C., Raddatz, T., and Claussen, M., 2008. A reconstruction of global agricultural areas and land cover for the last millennium. *Global and Biogeochemical Cycles*, 22(3), n/a–n/a. <http://doi.org/10.1029/2007GB003153>

Rohde, R., Muller, R.A., Jacobsen, R., Muller, E., Perlmutter, S., Rosenfeld, A., Wurtele, J., Groom, D., Wickham, C., 2013. A New Estimate of the Average Earth Surface Land Temperature Spanning 1753 to 2011. *Geoinformatics & Geostatistics: GeoInform-Geostat: An Overview* 1:1, doi:10.4172/gigs.1000101

Rossi, S., Morin, H., Deslauriers, A., 2012. Causes and correlations in cambium phenology: towards an integrated framework of xylogenesis. *Journal of Experimental Botany* 63(5), 2117-2126.

Rossi, S., Anfodillo, T., Čufar, K., Cuny, H.E., Deslauriers, A., Fonti, P., Frank, D., Gričar, J., Gruber, A., King, G.M., Krause, C., Morin, H., Oberhuber, W., Prislán, P., Rathgeber, G.B.K., 2013. A meta-analysis of cambium phenology and growth: linear and non-linear patterns in conifers of the northern hemisphere. *Annals of Botany* 112(9), 1911-1920.

Rossi, S., Deslauriers, A., Lupi, C., Morin, H., 2014. Control over growth in cold climates, in: Tausz, M., Grulke, N. (Eds.), *Trees in a Changing Environment*. pp. 191-219. Springer, Netherlands, pp. 191-219.

Rydval, M., Larsson, L.Å., McGlynn, L., Gunnarson, B.E., Loader, N.J., Young, G.H.F., Wilson, R., 2014. Blue intensity for dendroclimatology: Should we have the blues? Experiments from Scotland. *Dendrochronologia* 32(3), 191-204.

Rydval, M., Gunnarson, B.E., Loader, N.J., Druckenbrod, D., Moreton, S. and Wilson, R., in preparation. Reconstructing 800 years of summer temperatures in Scotland.

Schneider, L., Smerdon, J.E., Büntgen, U., Wilson, R., Myglan, V.S., Kirilyanov, A.V., Esper, J., 2015. Revising midlatitude summer temperatures back to A.D. 600 based on a wood density network. *Geophysical Research Letters* 42(11), 4556-4562, doi: 10.1002/2015GL063956

Schmidt, G. A., Jungclaus, J.-H., Ammann, C. M., Bard, E., Braconnot, P., Crowley, T. J., ... Vieira, L. E. A., 2012. [Climate forcing reconstructions for use in PMIP simulations of the Last Millennium \(v1.1\). \*Climate forcing reconstructions for use in PMIP simulations of the last millennium \(v1.0\). Geoscientific Model Development\* Geosci. Model Dev., 5\(1\), 185-191. doi:10.5194/gmd-5-185-2012](#)<http://doi.org/10.5194/gmd-4-33-2011>

Schmidt, G.-A., Kelley, M., Nazarenko, L., Ruedy, R., Russell, G. L., Aleinov, I., ... Zhang, J., 2014. Configuration and assessment of the GISS ModelE2 contributions to the CMIP5 archive. *Journal of Advances in Modeling Earth Systems*, 6(1), 141–184. <http://doi.org/10.1002/2013MS000265>

Schurer, A. ~~P.~~, Hegerl, G. ~~C.~~, Mann, M. ~~E.~~, Tett, S. ~~F.-B.~~, & Phipps, S. ~~J.~~, 2013. Separating Forced from Chaotic Climate Variability over the Past Millennium. *Journal of Climate*, 26(18), 6954–6973. <http://doi.org/10.1175/JCLI-D-12-00826.1>

Schurer, A. ~~P.~~, Tett, S. ~~F.-B.~~, Hegerl, G. ~~C.~~, 2014. Small influence of solar variability on climate over the past millennium. *Nature Geoscience*. 7, 104–108, doi: 10.1038/ngeo2040

Sigl, M., McConnell, J. ~~R.~~, Layman, L., Maselli, O., McGwire, K., Pasteris, D., Dahl-Jensen, D., Steffensen, J. ~~P.~~, Vinther, B., Edwards, R., Mulvaney, R., Kipfstuhl, S., 2012. A new bipolar ice core record of volcanism from WAIS Divide and NEEM and implications for climate forcing of the last 2000 years. *Journal of Geophysical Research: Atmospheres* 118(3), 1151-1169, doi: 10.1029/2012JD018603

Smerdon, J. ~~E.~~, Kaplan, A., Zorita, E., González-Rouco, J. ~~F.~~, Evans, M. ~~N.~~, 2011. Spatial performance of four climate field reconstruction methods targeting the Common Era. *Geophysical Research Letters* 38, L11705, doi:10.1029/2011GL047372

~~St.~~ George, S. ~~S.~~, 2014. An overview of tree-ring width records across the Northern Hemisphere. *Quaternary Science Reviews* 95, 132-150.

~~St.~~ George, S. ~~S.~~, Ault, T. ~~R.~~, 2014. The imprint of climate within Northern Hemisphere trees. *Quaternary Science Reviews* 89, 1-4.

Stine, A. ~~R.~~, Huybers, P., 2014. Arctic tree rings as recorders of variations in light availability. *Nature Communications* 5, 1–8, doi:10.1038/ncomms4836

Steinhilber, F., Beer, J., & Fröhlich, C., 2009. Total solar irradiance during the Holocene. ~~Geophys. Res. Lett~~~~Geophysical Research Letters~~, 36(19), L19704. <http://doi.org/10.1029/2009GL040142>

Stoffel, M., Khodri, M., Corona, C., Guillet, S., Poulain, V., Bekki, S., Guiot, J., Luckman, B., Oppenheimer, C., Lebas, N., Beniston, M & Masson-Delmotte, V. (2015). Estimates of volcanic-induced cooling in the Northern Hemisphere over the past 1,500 years. *Nature Geoscience*. 8, 784–788 (2015) doi:10.1038/ngeo2526.

Stokes, M. ~~A.~~, Smiley, T. ~~L.~~, 1968. An introduction to tree-ring dating. University of Chicago Press, Chicago, Illinois, USA.

Szeicz, J., MacDonald, G., 1995. Dendroclimatic reconstruction of summer temperatures in northwestern Canada since A. D. 1638 based on age-dependent modeling, *Quaternary Research* 44(2), 257–266.

~~Tingley and Huybers~~, Tingley, M. ~~P.~~, & Huybers, P., 2013. Recent temperature extremes at high northern latitudes unprecedented in the past 600 years. *Nature*, 496(7444), 201–205. doi:10.1038/nature11969

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.

Trenberth, K. ~~E.~~, Fasullo, J. ~~T.~~, 2010. Simulation of present-day and twenty-first-century energy budgets of the Southern Oceans. *Journal of Climate* 23(2), 440–454, doi: 10.1175/2009jcli3152.1

Trouet, V., Esper, J., Graham, N.E., Baker, A., Scourse, J.D., Frank, D.C., 2009. Persistent positive North Atlantic Oscillation mode dominated the medieval climate anomaly. *Science* 324, 78–80.

Vaganov, E.A., Hughes, M.K., Shashkin, A.V., 2006. Growth dynamics of conifer tree rings images of past and future environments. Springer, Berlin, New York.

Vaganov, E.A., Anchukaitis, K.J., Evans, M.H., 2011. How well understood are the processes that create dendroclimatic records? A mechanistic model of climatic control on conifer tree-ring growth dynamics, in: Hughes, M.K., Swetnam, T.W., Diaz, H.F. (Eds.), *Dendroclimatology: Progress and Prospects, Developments in Paleocological Research*. Springer-Verlag, Netherlands, 11(2), pp. 37-75, doi:10.1007/978-1-4020-5725-0\_3

Vieira, L. E.A., Solanki, S.K., Krivova, N. A., & Usoskin, I., 2011. Evolution of the solar irradiance during the Holocene. *Astronomy & Astrophysics*, 531, A6. <http://doi.org/10.1051/0004-6361/201015843>

Wahl, E.R., Ammann, C.M., 2007. Robustness of the Mann, Bradley, Hughes reconstruction of Northern Hemisphere surface temperatures: examination of criticisms based on the nature and processing of climate evidence. *Climatic Change* 85(1-2), 33–69, doi: 10.1007/s10584-006-9105-7

Wiles, G.C., D'Arrigo, R.D., Barclay, D., Wilson, R., Jarvis, S., Vargo, L., Frank, D., 2014. Surface air temperature variability reconstructed with tree rings for the Gulf of Alaska over the past 1200 years. *The Holocene* 24(2), 198-208, doi: 10.1177/0959683613516815

Wilson, R., Tudhope, A., Brohan, P., Briffa, K., Osborn, T. and Tett, S. 2006. 250-years of reconstructed and modeled tropical temperatures. *Journal of Geophysical Research*. 111, C10007, doi:10.1029/2005JC003188

Wilson, R., D'Arrigo, R., Buckley, B., Büntgen, U., Esper, J., Frank, D., Luckman, B., Payette, S., Vose, R., Youngblut, D., 2007. A matter of divergence – tracking recent warming at hemispheric scales using tree-ring data. *Journal of Geophysical Research:– Atmospheres* (1984-2012) 112(D17), doi: 10.1029/2006JD008318

Wilson, R., Rao, R., Rydval, M., Wood, C., Larsson, L.Å., Luckman, B.H., 2014. Blue Intensity for Dendroclimatology: The BC Blues: A Case Study from British Columbia, Canada. *The Holocene* 24(11), 1428-1438, doi: 10.1177/0959683614544051

Wu, T., Li, W., Ji, J., Xin, X., Li, L., Wang, Z., Zhang, Y., Li, J., Zhang, F., and Wei, M., Wu, T., Li, W., Ji, J., Xin, X., Li, L., Wang, Z., ... Zhang, J., 2013. Global carbon budgets simulated by the Beijing Climate Center Climate System Model for the last century. *J. Geophys. Res.: Atm.*, 118(10), 4326–4347. <http://doi.org/10.1002/jgrd.5032>

Young, G.H.F., Bale, R.J., Loader, N.J., McCarroll, D., Nayling, N., Vousden, N., 2012. Central England temperature since AD 1850: the potential of stable carbon isotope in British oak trees to reconstruct past summer temperatures. *Journal of Quaternary Science* 27(6), 606–614.

Youngblut, D.K., Luckman, B.H., 2008. Maximum June-July temperatures in the southwest Yukon region over the last three hundred years reconstructed from tree-rings. *Dendrochronologia* 25(3), 153-166.

| Zhang, P., Linderholm, H.~~W.~~, Gunnarson, B.~~E.~~, Björklund, J., Chen, D., 2015. 1200 years of warm-season temperature variability in central Fennoscandia inferred from tree-ring density. *Climate of the Past Discussion* 11, 489-519, doi: 10.5194/cpd-11-489-2015

# Tables and Figures

Regional Grouping	Code	Lat	Long	Period Used	Parameter	Chronologies Type	Reference	Correlation Tmean (post 1950)
<b>NORTH AMERICA</b>								
Seward	NTR	65.11 to 65.22	162.18 to 162.27	1710-2001	MXD	STD	D'Arrigo et al. (2004)	JA: 0.47 (0.40)
Coastal Alaska	GOA	60.01 to 60.45	-149.31 to -141.42	800-2010	RW	RCS	Wiles et al. (2014)	Jan-Sep: 0.42 (0.52)
Wrangells	WRAX	60-65	-145.00 to -140.00	1593-1992	MXD	STD	Davi et al. (2003); used in Wilson et al. (2007)	JA: 0.61 (0.66)
Firth	FIRT	68.39	-141.38	1073 - 2002	MXD	RCS-SF	Anchukaitis et al. (2013)	JA: 0.58 (0.64)
Southern Yukon	YUS	59-62	-140 to -133	1684-2000	RW	STD	Youngblut and Luckman [2007]; used in Wilson et al. (2007)	JA: 0.58 (0.60)
Northern Yukon	YUN	65-70	-125 to -135	1638-1988	RW	STD	Szeicz and MacDonald [1995]; used in Wilson et al. (2007)	June: 0.44 (0.55)
Int. British Columbia	IBC	49.02-50.59	-121.43 to -117.03	1600-1995	RW/MXD/BI	STD/SF	Wilson et al. (2014)	JA: 0.68 (0.74)
Icefields	ICE	52-16	-117.19	918 - 1994	RW/MXD	RCS	Luckman and Wilson (2005); Used in D'Arrigo et al. (2006)	JA: 0.74 (0.79)
Idaho	IDA	40-45	-110 to -120	1135-1992	RW	STD	Biondi et al. [1999]; used in Wilson et al. (2007)	July: 0.47 (0.51)
Coppermine	COP	67.14	-115.55	1551 - 2003	MXD	RCS	Anchukaitis et al. (2013)	MJJA: 0.59 (0.60)
Thelon	THE	64.02	-103.52	1492 - 2004	MXD	RCS	Anchukaitis et al. (2013)	MJJA: 0.67 (0.73)
Quebec	QUEX	57.30	76.00	1373-1988	MXD	RCS	Schneider et al. (2015)	MJJA: 0.70 (0.79)
Quebec	QUEW	57.30	74.00	910-2011	RW	RCS	Gennaretti et al. (2014)	JA: 0.47 (0.57)
Northern Quebec	NUQ	55-60	-70 to -65	1642-2002	RW	STD	Payette (2007); used in Wilson et al. (2007)	JJ: 0.39 (0.35)
Labrador	LAB	56.33 to 57.58	-62.25 to -61.56	1710-1998	RW/MXD	STD/RCS	D'Arrigo et al. (2003 (RW) and 2013 (MXD))	JA: 0.53 (0.52)
<b>EURASIA</b>								
Scotland	SCOT	57.08	-3.44	1200-2010	RW/BI	STD/RCS	Rydval et al. (submitted)	JA: 0.75 (0.73)
Pyrenees	PYR	42 - 43	0 - 1	1260-2005	MXD	RCS	Dorado-Linan et al. (2012)	AS: 0.53 (0.75)
W Alps - Lotschental	ALPS			755-2004	MXD	RCS	Buentgen et al. (2006).	JJAS: 0.75 (0.73)
E Alps - Tyrol	TYR	47.30	12.30	1053-2003	MXD	RCS	Schneider et al. (2015)	JAS: 0.72 (0.72)
Jaemtland	JAEM	63.30	13.25	783-2011	MXD	RCS	Zhang et al. (2015)	AMJJA: 0.75 (0.65)
Tjeggelvas, Arjeplog, Ammarnäs composite	TAA	65.54 - 66.36	16.06 - 18.12	1200-2010	MXD	RCS	Linderholm et al. (2014)	MJJA: 0.81 (0.68)
North Fenno	Efmean	66-69	19-32	750-2010	MXD	RCS	Esper et al (2014) and Matskovsky and Helama (2014)	JA: 0.76 (0.73)
Forfjorddalen	FORF	68.47	15.43	978-2005	MXD	MXD	McCarroll et al. (2013)	JA: 0.71 (0.72)
Tatra	TAT	48-49	19-20	1040-2010	RW	RCS	Büntgen et al. (2013)	MJ: 0.45 (0.47)
Mt Olympus, Greece	MOG	40.09	22.37	1521-2010	MXD	RCS	Klesse et al. (2014)	JJA: 0.54 (0.53)
South Finland	SFIN	62.19.30	28.19.30	760-2000	MXD	RCS	Helama et al. (2014)	MJJA: 0.71 (0.63)
Khibiny (Kola)	KOL	67.38 - 67.50	33.13 - 34.15	821-2005	RW/BI	RCS/STD	McCarroll et al. (2013)	JA: 0.73 (0.75)
Polar Urals	POLx	66.51	65.40	891-2006	MXD	RCS	Schneider et al. (2015)	JA: 0.80 (0.83)
Yamal	YAM	67.32	69.54	750-2005	RW	RCS	Briffa et al. (2013)	July: 0.61 (0.51)
Asia Grid 1	ASGrid1	40.15 - 46.15	60.15 - 68.15	817-1989	MX	RCS/STD	Cook et al. (2012)	JJ: 0.51 (0.60)
Asia Grid 2	ASGrid2	40.15 - 46.15	70.15 - 78.15	827-1989	MX	RCS/STD	Cook et al. (2012)	JA: 0.26 (0.59)
Asia Grid 10	ASGrid10	48.15 - 54.15	60.15 - 68.15	937-1989	MX	RCS/STD	Cook et al. (2012)	JA: 0.33 (0.52)
Asia Grid 11	ASGrid11	48.15 - 54.15	70.15 - 78.15	937-1989	MX	RCS/STD	Cook et al. (2012)	MJJA: 0.51 (0.71)
Kyrgyzstan	KYR	41.36 to 42.11	75.09 to 78.11	1689-1995	RW/MXD	STD	Wilson et al. (2007)	JA: 0.70 (0.77)
Mangazeja	MAN	66.42	82.18	1328-1990	MXD	RCS	Schneider et al. (2015)	JA: 0.79 (0.69)
Asia Grid 3	ASGrid3	40.15 - 46.15	80.15 - 88.15	800-1989	MX	RCS/STD	Cook et al. (2012)	JA: 0.44 (0.69)
Asia Grid 12	ASGrid12	48.15 - 54.15	80.15 - 88.15	800-1989	MX	RCS/STD	Cook et al. (2012)	JA: 0.62 (0.84)
Altai MXD	ALT	50.00	88.00	750-2007	MXD	RCS	Schneider et al. (2015)	MJJA: 0.65 (0.72)
Asia Grid 4	ASGrid4	40.15 - 46.15	90.15 - 98.15	800-1989	MX	RCS/STD	Cook et al. (2012)	JJ: 0.60 (0.81)
Asia Grid 13	ASGrid13	48.15 - 54.15	90.15 - 98.15	1024-1989	MX	RCS/STD	Cook et al. (2012)	JA: 0.52 (0.80)
Mongolia	OZN	51.15	99.04	931-2005	RW	RCS	Davi et al. (2015)	JJ: 0.52 (0.62)
Taymir	TAY	72.01	102.00	755 - 1997	RW	RCS	Jacoby et al. (2000); Used in D'Arrigo et al. (2006)	JJ: 0.45 (0.45)
Asia Grid 5	ASGrid5	40.15 - 46.15	100.15 - 108.15	800-1989	MX	RCS/STD	Cook et al. (2012)	JA: 0.39 (0.59)
Asia Grid 14	ASGrid14	48.15 - 54.15	100.15 - 108.15	1396-1989	MX	RCS/STD	Cook et al. (2012)	JA: 0.58 (0.70)
Asia Grid 6	ASGrid6	40.15 - 46.15	110.15 - 118.15	800-1989	MX	RCS/STD	Cook et al. (2012)	JA: 0.25 (0.50)
Asia Grid 15	ASGrid15	48.15 - 54.15	110.15 - 118.15	1396-1989	MX	RCS/STD	Cook et al. (2012)	JA: 0.15 (0.46)
Asia Grid 7	ASGrid7	40.15 - 46.15	120.15 - 128.15	1024-1989	MX	RCS/STD	Cook et al. (2012)	MJJA: 0.30 (0.66)
Asia Grid 16	ASGrid16	48.15 - 54.15	120.15 - 128.15	1510-1989	MX	RCS/STD	Cook et al. (2012)	MJJA: 0.04 (0.41)
Asia Grid 8	ASGrid8	40.15 - 46.15	130.15 - 138.15	1510-1989	MX	RCS/STD	Cook et al. (2012)	JJ: 0.58 (0.68)
Asia Grid 17	ASGrid17	48.15 - 54.15	130.15 - 138.15	1510-1989	MX	RCS/STD	Cook et al. (2012)	MJJA: 0.41 (0.57)
Asia Grid 9	ASGrid9	40.15 - 46.15	140.15 - 148.15	1510-1989	MX	RCS/STD	Cook et al. (2012)	JJ: 0.67 (0.71)
Asia Grid 18	ASGrid18	48.15 - 54.15	140.15 - 148.15	1510-1989	MX	RCS/STD	Cook et al. (2012)	JJ: 0.52 (0.53)
North Japan	NJAP	43-51	142-145	1640-1993	RW/MXD	STD	D'Arrigo et al. (2014)	JJ: 0.65 (0.62)
Yakutia	YAK	67.27 to 70.33	142.37 to 150.17	1342 - 1994	RW	RCS	Hughes et al. (1999); Used in D'Arrigo et al. (2006)	June: 0.30 (0.62)

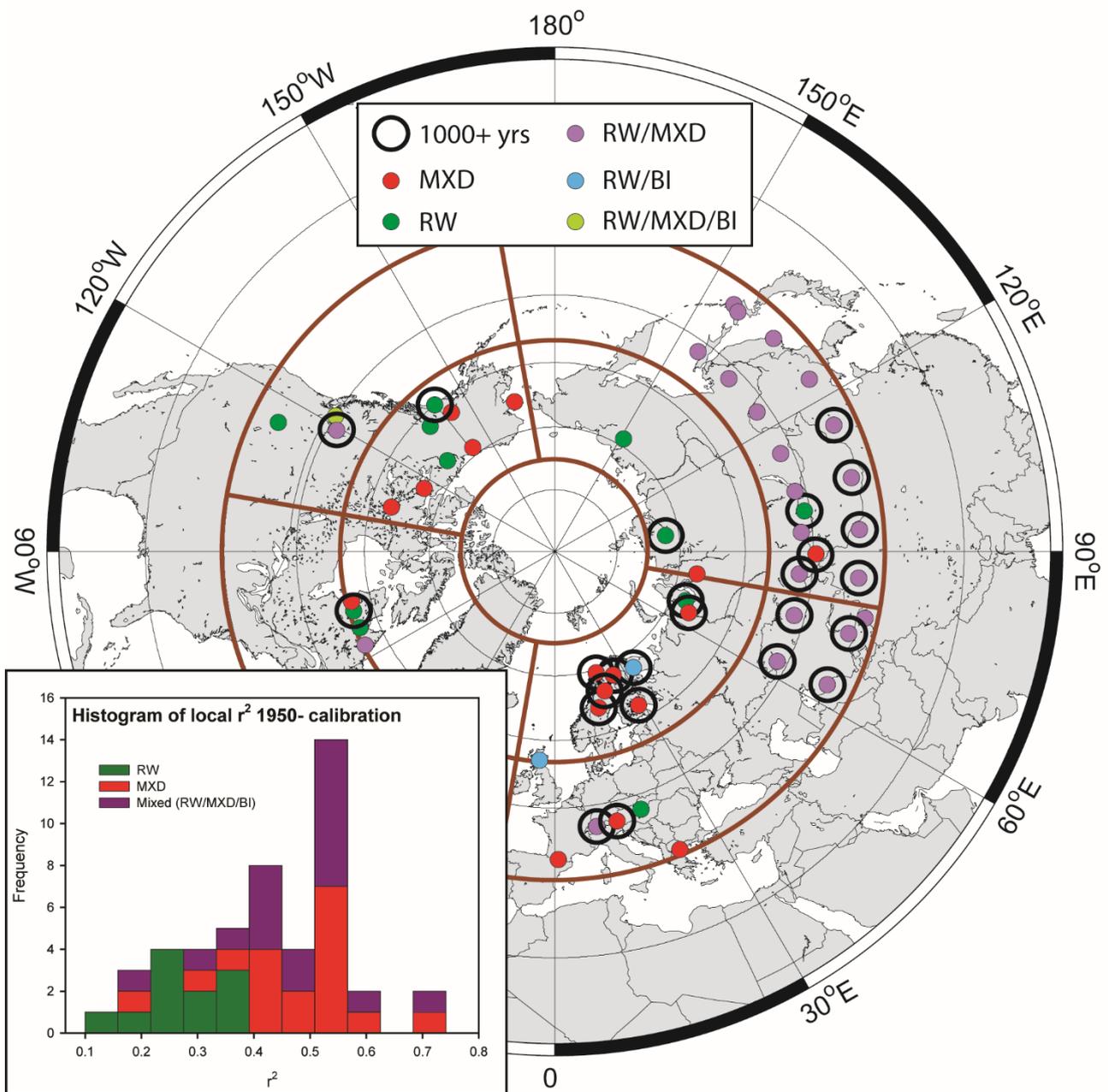
**Table 1:** Meta Information for tree-ring reconstructions used in the current N-TREND2015 network. See Appendix A Table A1 for brief information on how each series was derived. Final column presents correlations for 1901-present (1950-present) for all series against the local 1 degree (or larger where relevant) CRUTS3.2 (Harris et al. 2014) mean temperature grid for the TR site's optimal summer season. Histogram of post 1950 correlation squared values is embedded in Figure 1.

Rank	Warm	Anomaly	$\pm 2\sigma$	Cool	Anomaly	$\pm 2\sigma$
1	1994-2003	0.34	0.25	1812-1821	-1.03	0.19
2	1946-1955	0.30	0.19	1832-1841	-1.00	0.19
3	1161-1170	0.27	0.23	1462-1471	-0.96	0.19
4	1980-1989	0.19	0.19	1695-1704	-0.94	0.19
5	1425-1434	0.17	0.19	1452-1461	-0.90	0.19

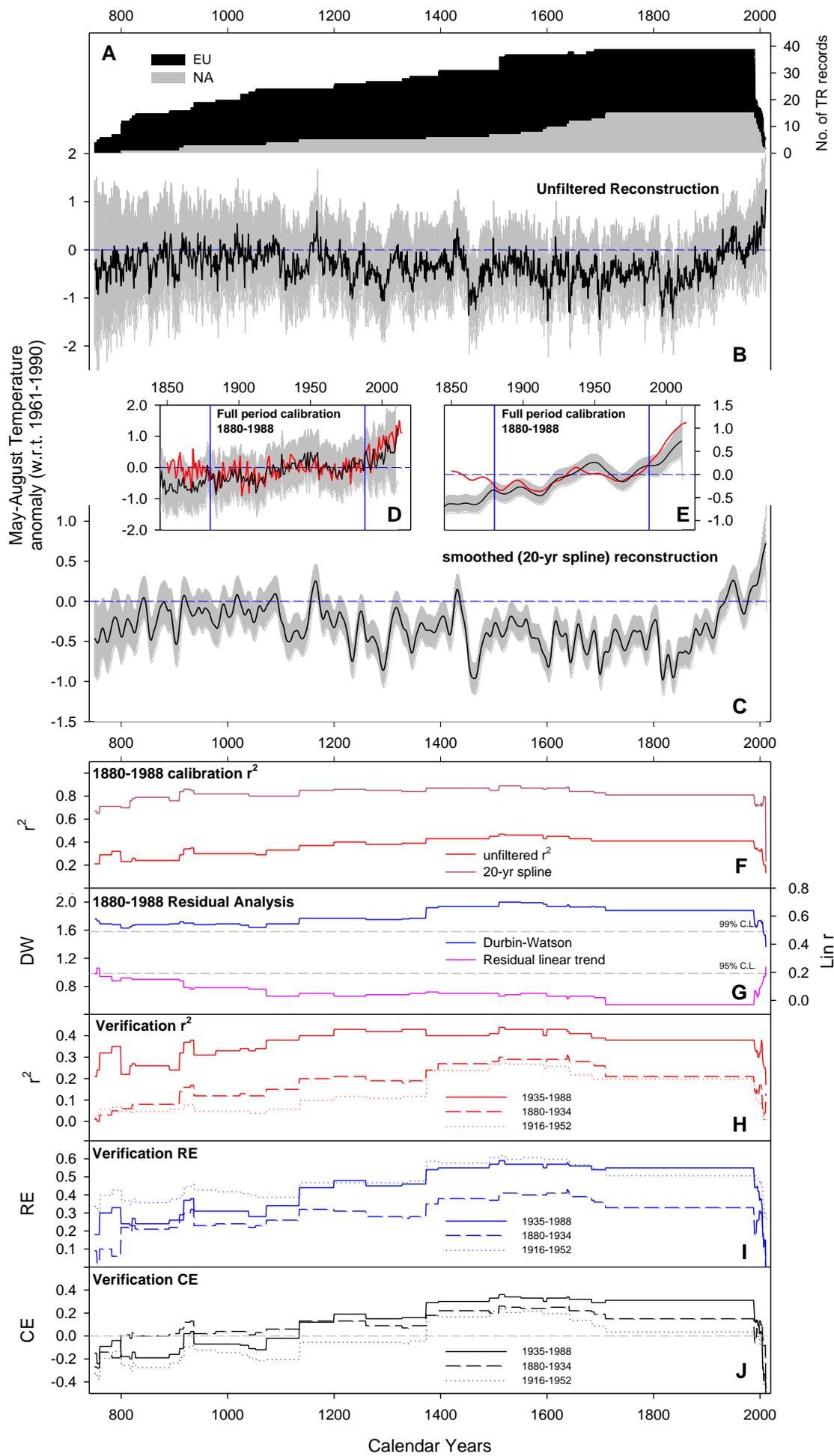
**Table 2:** Five warmest and coldest non-overlapping decadal periods in N-TREND2015. Calculated over 918-2004. Temperature anomalies (plus 2-sigma error) shown with respect to 1961-1990.

Event	Gao/Crowley	Esper	Sigl
Unknown	1227		1229
Samalas (Indonesia)	1258		1257
Katla (Iceland)		1262	
Unknown	1284		1286
Unknown			1344
Oraefajokull (Iceland)		1362	
Kuwae? (Vanuatu)	1459		1458
Barbarbunga (Iceland)		1477	
Hekla (Iceland)			
Colima (Mexico)	1584		
Ruiz (Columbia)			1594
Huaynaputina (Peru)	1600	1600	1600
Katla (Iceland)		1625	
Furnas (Azores)		1630	
Komaga-Take (Japan)		1640	
Parker Peak (Philippines)	1641		
Usu (Japan)		1663	
Shikotsu (Japan)		1667	
Unknown			1694
Fuji (Japan)		1707	
Shikotsu (Japan)		1739	
Katla (Iceland)		1755	
Laki (Iceland)	1783		
Unknown	1809		1809
Tambora (Indonesia)	1815	1815	1815
Babuyan (Philippines)	1831		
Cosiguina (Nicaragua)	1835	1835	
Shiveluch (Russia)		1854	
Makian (Indonesia)			
Askja (Iceland)		1875	
Krakatau (Krakatau)		1883	
Santa Maria (Guatemala)		1902	
Ksudach (Russia)		1907	
Katmai (USA)		1912	
Colima (Mexico)		1913	
Bezymianny (Russia)		1956	
Pinatubo (Indonesia)			1991

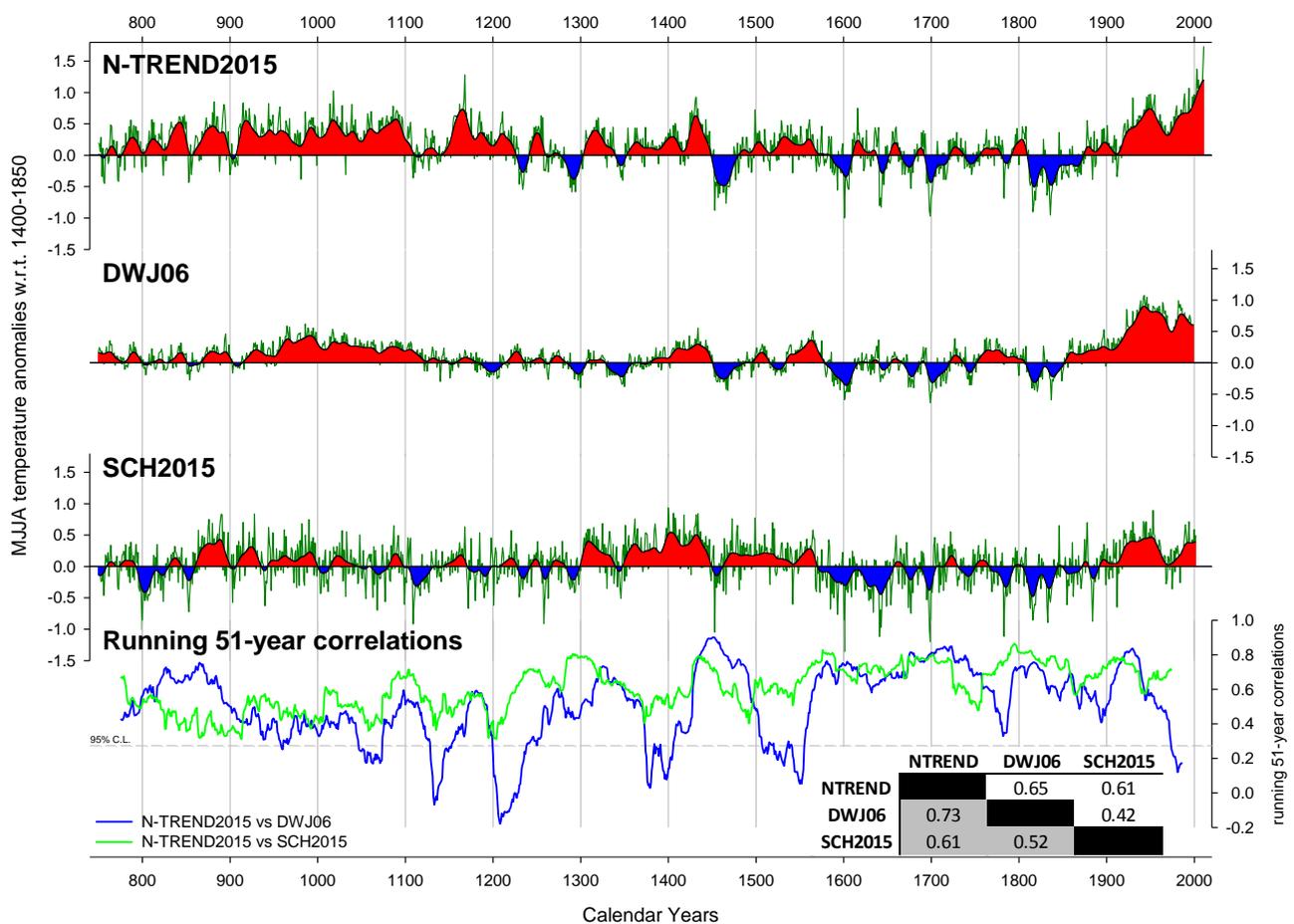
**Table 3:** Three different lists of significant volcanic events derived from: Gao/Crowley = the earliest year with common sulphate deposition values in both hemispheres  $> 15 \text{ kg km}^{-2}$  detailed in both Gao et al. (2008), Crowley and Unterman (2013); Esper = The large volcanic events list in Esper et al. (2013) which were identified via historical documents and had a Volcanic Explosivity Index  $\geq 5$ ; Sigl = volcanic events, detailed Sigl et al. (2013), with sulphate deposition values in both hemispheres  $> 15 \text{ kg km}^{-2}$ .



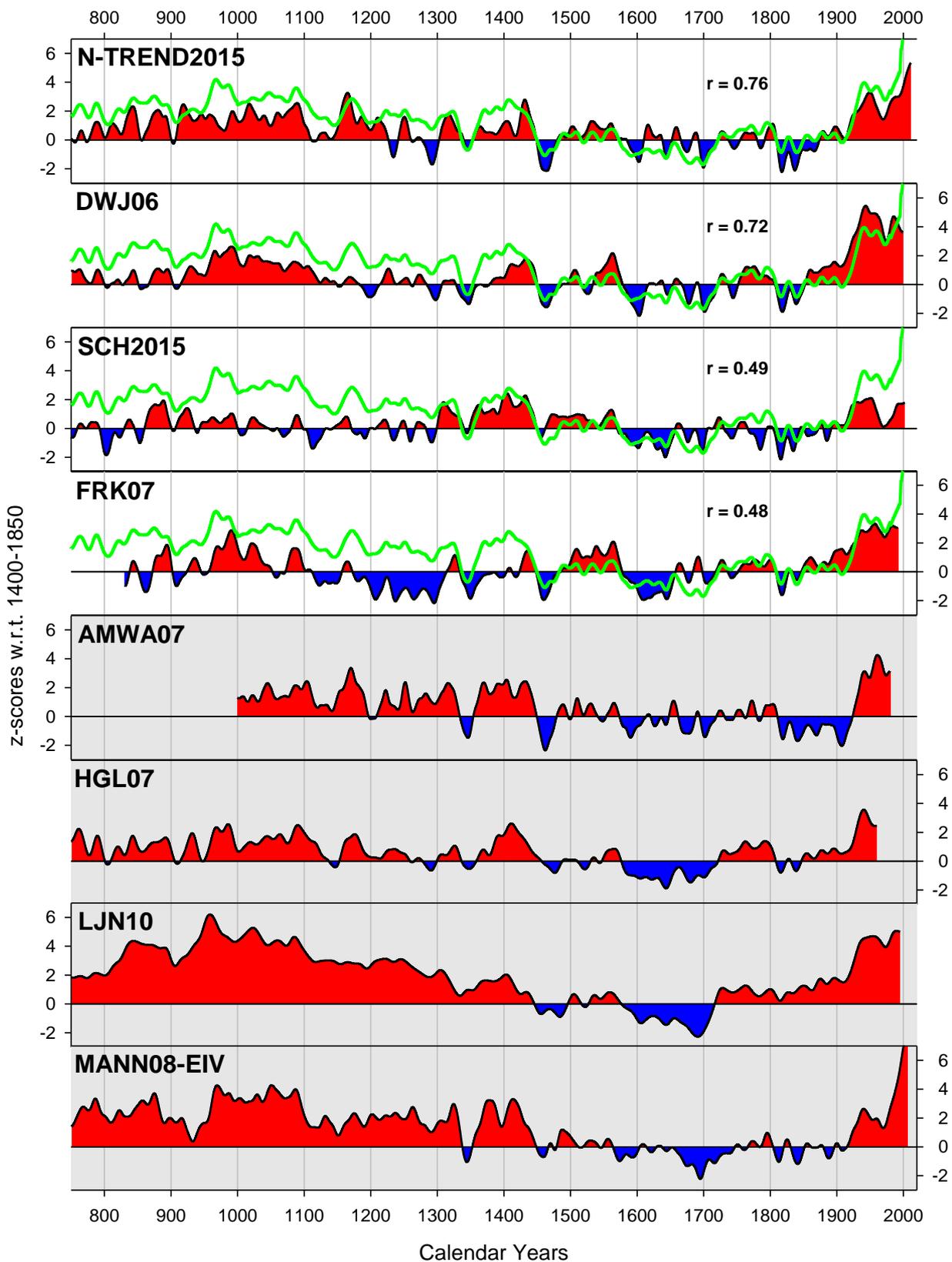
**Figure 1:** Location map of individual tree-ring reconstructions used in the current N-TREND2015 network. See Table 1 for site information and Appendix Table 1A for brief information on how each series was derived. Brown lines denote the high/low latitude and four longitude quadrants. Inset histogram shows distribution of post 1950  $r^2$  values for local calibration, for site specific optimal season, to the CRU TS3.2 grid (Harris et al. 2014 - see Table 1 for grid specifications).



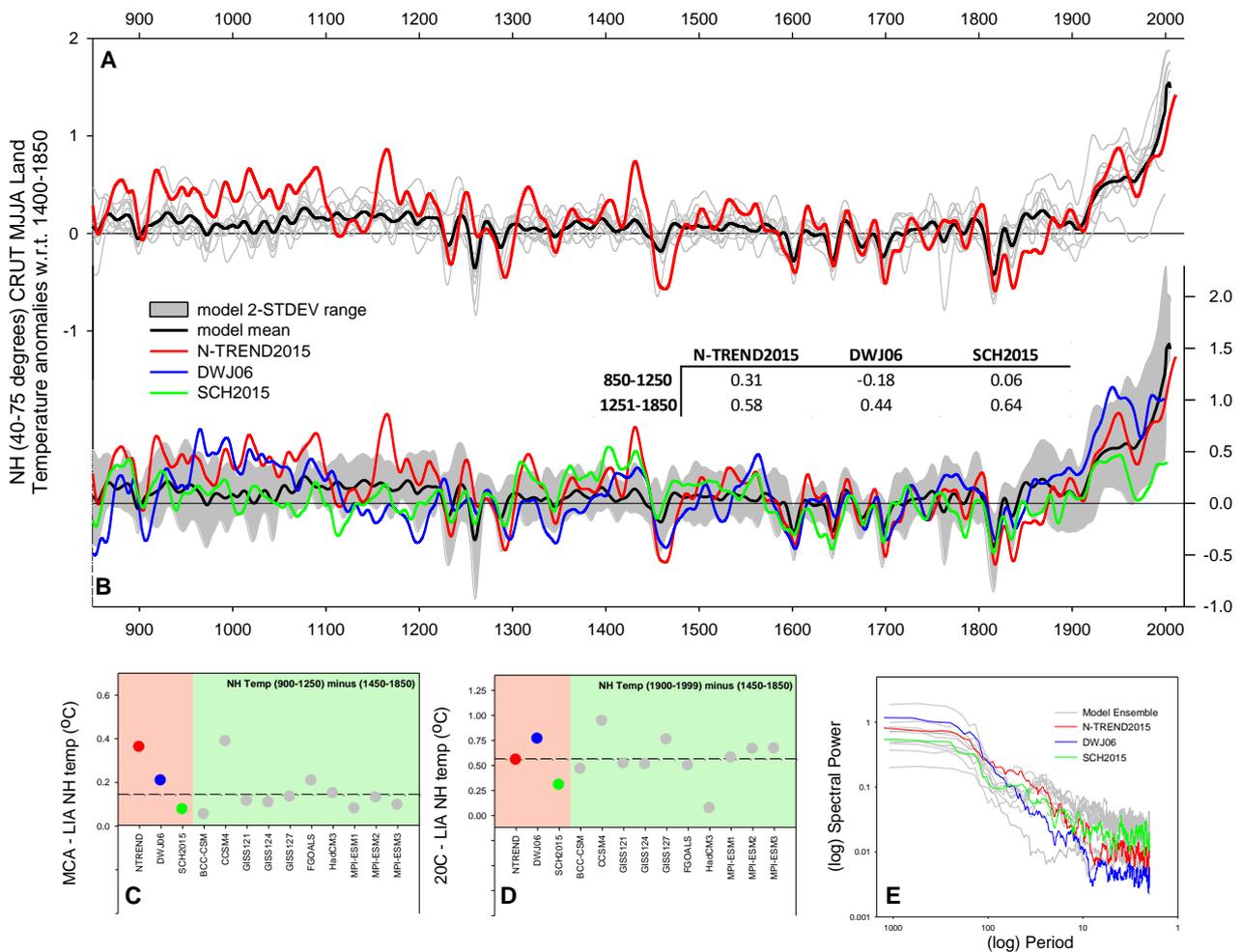
**Figure 2:** N-TREND2015 reconstruction and assessment metrics. **Upper panel:** TR series replication (A), unfiltered and filtered (20-year spline) mean May-August reconstruction with 2-sigma error bar for both 750-2011 (B+C) and 1850-2011 periods (D+E). Errors are derived from the RMSE of the full calibration period (1880-1988) residuals (post scaling), calculated for each nest. An additional error component is also incorporated to estimate the extra uncertainty of the early weaker replicated periods of each individual input series (see main text and Appendix Figure B2). **Lower panel:** F: 1880-1988 calibration  $r^2$  values for unfiltered and filtered (20-year spline) versions; G: 1880-1988 residual analysis – Durbin and Watson (DW) values and correlation of linear trend (Lin r) of residuals; H: Validation  $r^2$  values using the 3 validation periods 1880-1934, 1935-1988 and 1916-1952; I: As for H, but for RE; J: As for H, but for CE. See Appendix Table B1 for detail reconstruction metrics for each nest.



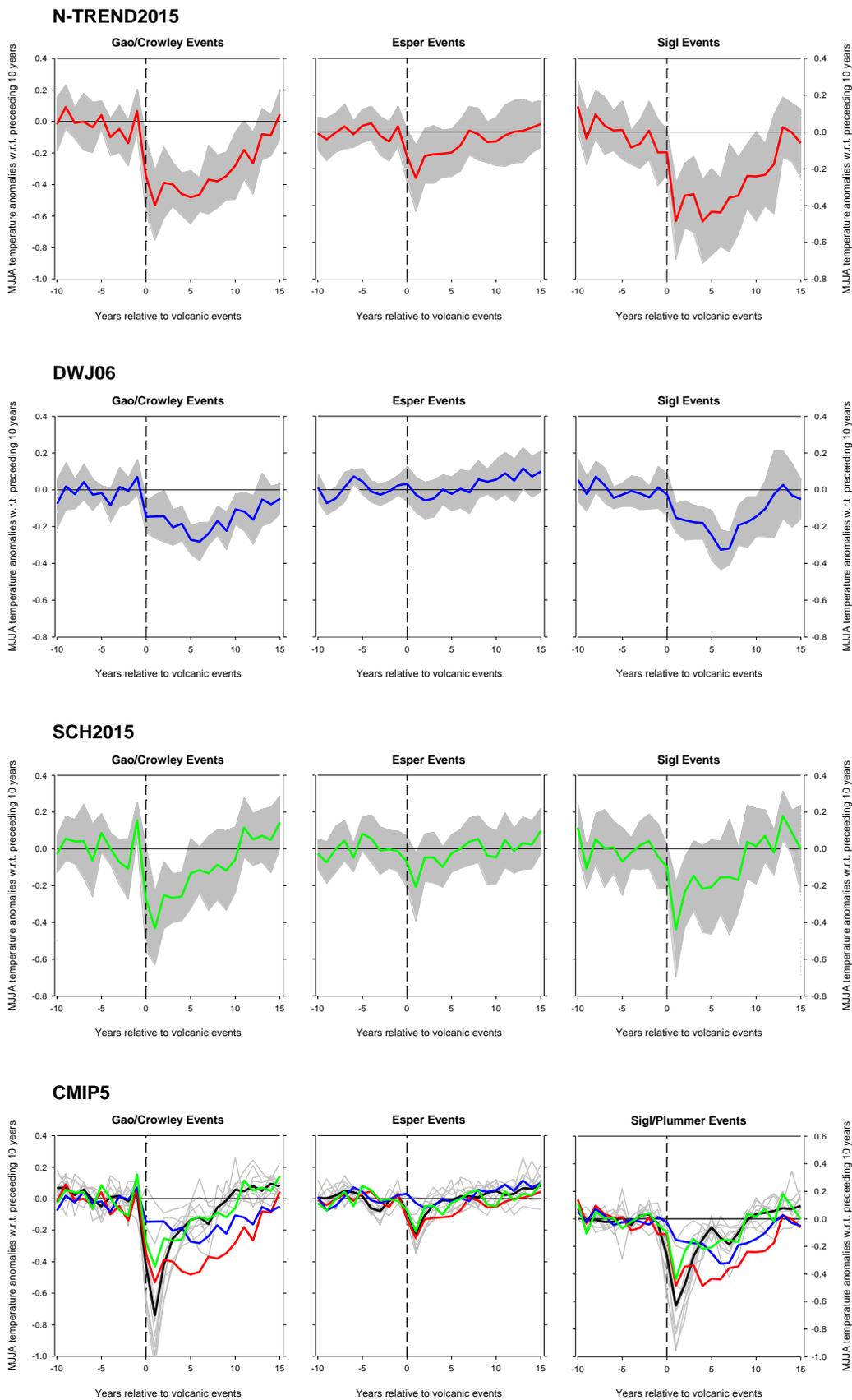
**Figure 3:** Tree-ring based summer temperature NH reconstructions. N-TREND2015, DWJ06 (D’Arrigo et al. 2006) and SCH2015 (Schneider et al. 2015). The series have been scaled to CRUTEM (40-75) MJJA land temperatures over the 1880-1988 period and expressed as anomalies w.r.t. 1400-1850. Smoothed function is a 20-year spline. Lower panel shows sliding 51-year correlation values between N-TREND2015 and the other two reconstructions. A correlation matrix (918-1999) is also shown for unfiltered and filtered (20-year spline - in grey) versions.



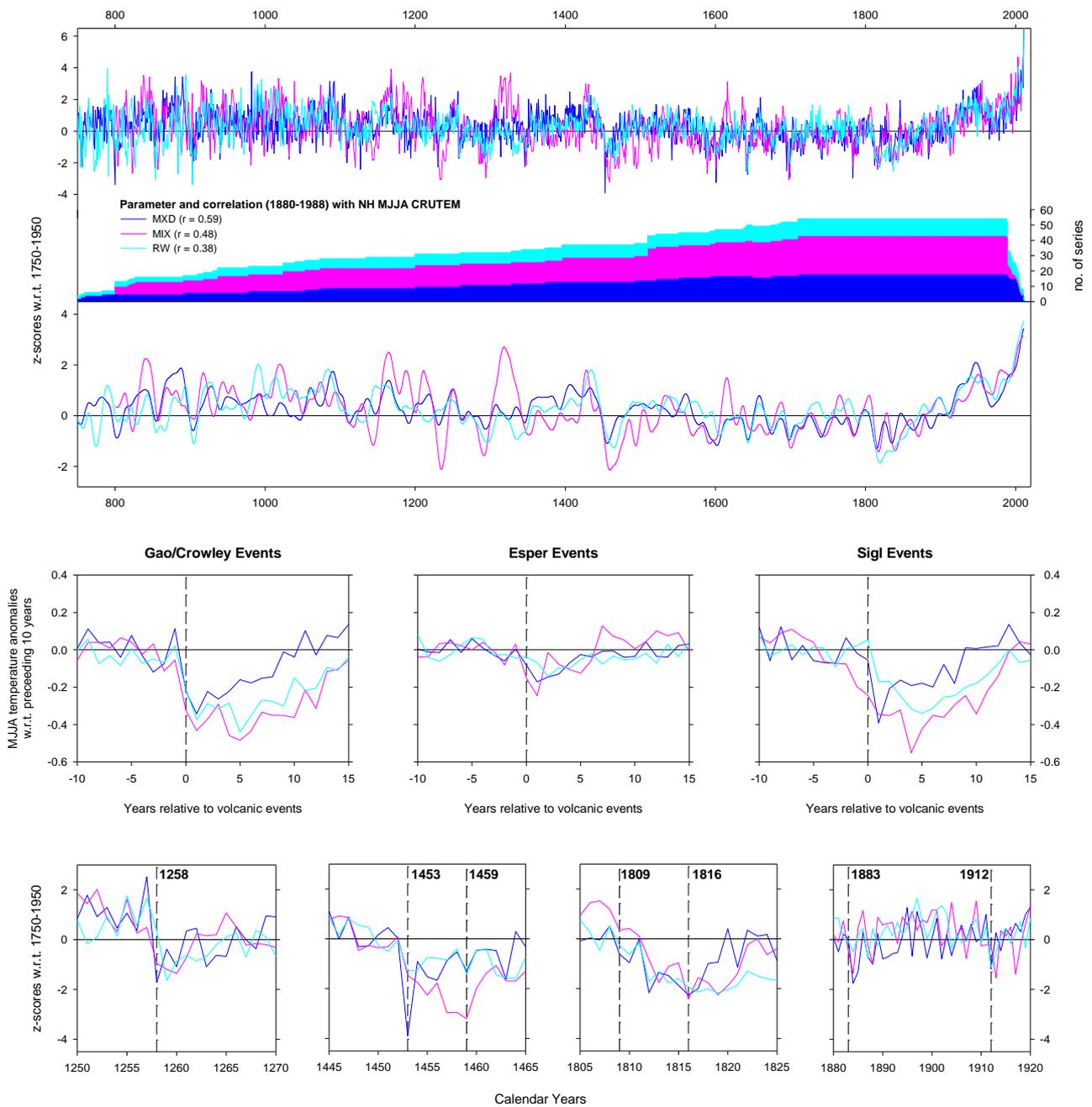
**Figure 4:** Comparison of TR based reconstructions (N-TREND2015, DWJ06, SCH2015 and FRK07) with select multi-proxy (MP) reconstructions (Amman and Wahl 2007; Hegerl et al. 2007; Ljungqvist 2010; Mann et al. 2008) published since 2007. All smoothed (20-year spline) series have been normalised to the 1400-1850 period to enhance differences in the Medieval and recent periods. The green-time series is the normalised average of AMWA07, HGL07, LJN07 and MANN08-EIV and the correlation shown is between this mean MP composite and the individual TR NH series over the 918-1992 period.



**Figure 5: Upper:** Comparisons of N-TREND with 10 CMIP5 models. Time series have been smoothed using a 20-year spline and transformed to anomalies w.r.t. 1400-1850. **A:** N-TREND versus the 10 CMIP5 models; **B:** N-TREND, DJW06 and SCH2015 compared to the 2-STDEV range of the 10 CMIP5 models. Inset table shows correlations between each TR NH series and the model mean for the two periods 850-1250 and 1251-1850; **C:** Temperature difference (°C) from subtracting the mean of 1450-1850 from the mean of 900-1250 (NB. NGOALS starts in 1000). Dashed line is the model mean of the difference between the two periods; **D:** Temperature difference (°C) from subtracting the mean of 1450-1850 from the mean of 1900-1999. Dashed line is the model mean of the difference between the two periods; **E:** Spectral Power plot of N-TREND2015, DJW06 and SCH2015 and the 10 CMIP5 models calculated over the 900-1999 period (1000-1999 for FGOALS).



**Figure 6:** Superposed epoch analysis for N-TREND2015, DWJ06, SCH2015 and the CMIP5 10 models assessing mean response to major volcanic events. Lower panels compare individual model runs (plus their mean) with each reconstruction. Table 3 lists the difference years extracted from Gao et al. 2008; Crowley and Unterman 2013; Esper et al. 2013; Sigl et al. 2013 for the three different lists.



**Figure 7: Upper:** Parameter specific NH composites using the N-TREND database. Method used was 2-quadrant (North American and Eurasia) weighting with variance stabilisation (see Appendix Figure C1); **Middle:** SEA (see Figure 6) using each parameter composite; **Lower:** Parameter composite comparison for select periods with significant volcanic events (see Table 3).

Tables and Figures - with track changes

Regional Grouping	Code	Lat	Long	Period Used	Parameter	Chronologies Type	Reference	Correlation Tmean (post 1950)
<b>NORTH AMERICA</b>								
Seward	NTR	65.11 to 65.22	162.18 to 162.27	1710-2001	MXD	STD	D'Arrigo et al. (2004)	JA: 0.47 (0.40)
Coastal Alaska	GOA	60.01 to 60.45	-149.31 to -141.42	800-2010	RW	RCS	Wiles et al. (2014)	Jan-Sep: 0.42 (0.52)
Wrangells	WRAX	60-65	-145.00 to -140.00	1593-1992	MXD	STD	Davi et al. (2003); used in Wilson et al. (2007)	JA: 0.61 (0.66)
Firth	FIRT	68.39	-141.38	1073 - 2002	MXD	RCS-SF	Anchukaitis et al. (2013)	JA: 0.58 (0.64)
Southern Yukon	YUS	59-62	-140 to -133	1684-2000	RW	STD	Youngblut and Luckman [2007]; used in Wilson et al. (2007)	JA: 0.58 (0.60)
Northern Yukon	YUN	65-70	-125 to -135	1638-1988	RW	STD	Szeicz and MacDonald [1995]; used in Wilson et al. (2007)	June: 0.44 (0.55)
Int. British Columbia	IBC	49.02-50.59	-121.43 to -117.03	1600-1995	RW/MXD/BI	STD/SF	Wilson et al. (2014)	JA: 0.68 (0.74)
Icefields	ICE	52-16	-117.19	918 - 1994	RW/MXD	RCS	Luckman and Wilson (2005); Used in D'Arrigo et al. (2006)	JA: 0.74 (0.79)
Idaho	IDA	40-45	-110 to -120	1135-1992	RW	STD	Biondi et al. [1999]; used in Wilson et al. (2007)	July: 0.47 (0.51)
Coppermine	COP	67.14	-115.55	1551 - 2003	MXD	RCS	Anchukaitis et al. (2013)	MJJA: 0.59 (0.60)
Thelon	THE	64.02	-103.52	1492 - 2004	MXD	RCS	Anchukaitis et al. (2013)	MJJA: 0.67 (0.73)
Quebec	QUEX	57.30	76.00	1373-1988	MXD	RCS	Schneider et al. (2015)	MJJA: 0.70 (0.79)
Quebec	QUEW	57.30	74.00	910-2011	RW	RCS	Gennaretti et al. (2014)	JA: 0.47 (0.57)
Northern Quebec	NUQ	55-60	-70 to -65	1642-2002	RW	STD	Payette (2007); used in Wilson et al. (2007)	JJ: 0.39 (0.35)
Labrador	LAB	56.33 to 57.58	-62.25 to -61.56	1710-1998	RW/MXD	STD/RCS	D'Arrigo et al. (2003 (RW) and 2013 (MXD))	JA: 0.53 (0.52)
<b>EURASIA</b>								
Scotland	SCOT	57.08	-3.44	1200-2010	RW/BI	STD/RCS	Rydval et al. (submitted)	JA: 0.75 (0.73)
Pyrenees	PYR	42 - 43	0 - 1	1260-2005	MXD	RCS	Dorado-Linan et al. (2012)	AS: 0.53 (0.75)
W Alps - Lentschental	ALPS			755-2004	MXD	RCS	Buentgen et al. (2006).	JJAS: 0.75 (0.73)
E Alps - Tyrol	TYR	47.30	12.30	1053-2003	MXD	RCS	Schneider et al. (2015)	JAS: 0.72 (0.72)
Jaemtland	JAEM	63.30	13.25	783-2011	MXD	RCS	Zhang et al. (2015)	AMJJA: 0.75 (0.65)
Tjeggelvas, Arjeplog, Ammarnäs composite	TAA	65.54 - 66.36	16.06 - 18.12	1200-2010	MXD	RCS	Linderholm et al. (2014)	MJJA: 0.81 (0.68)
North Fenno	Efmean	66-69	19-32	750-2010	MXD	RCS	Esper et al (2014) and Matskovsky and Helama (2014)	JA: 0.76 (0.73)
Forfjorddalen	FORF	68.47	15.43	978-2005	MXD	MXD	McCarroll et al. (2013)	JA: 0.71 (0.72)
Tatra	TAT	48-49	19-20	1040-2010	RW	RCS	Büntgen et al. (2013)	MJ: 0.45 (0.47)
Mt Olympus, Greece	MOG	40.09	22.37	1521-2010	MXD	RCS	Klesse et al. (2014)	JJA: 0.54 (0.53)
South Finland	SFIN	62.19.30	28.19.30	760-2000	MXD	RCS	Helama et al. (2014)	MJJA: 0.71 (0.63)
Khibiny (Kola)	KOL	67.38 - 67.50	33.13 - 34.15	821-2005	RW/BI	RCS/STD	McCarroll et al. (2013)	JA: 0.73 (0.75)
Polar Urals	POLx	66.51	65.40	891-2006	MXD	RCS	Schneider et al. (2015)	JA: 0.80 (0.83)
Yamal	YAM	67.32	69.54	750-2005	RW	RCS	Briffa et al. (2013)	July: 0.61 (0.51)
Asia Grid 1	ASGrid1	40.15 - 46.15	60.15 - 68.15	817-1989	MX	RCS/STD	Cook et al. (2012)	JJ: 0.51 (0.60)
Asia Grid 2	ASGrid2	40.15 - 46.15	70.15 - 78.15	827-1989	MX	RCS/STD	Cook et al. (2012)	JA: 0.26 (0.59)
Asia Grid 10	ASGrid10	48.15 - 54.15	60.15 - 68.15	937-1989	MX	RCS/STD	Cook et al. (2012)	JA: 0.33 (0.52)
Asia Grid 11	ASGrid11	48.15 - 54.15	70.15 - 78.15	937-1989	MX	RCS/STD	Cook et al. (2012)	MJJA: 0.51 (0.71)
Kyrgyzstan	KYR	41.36 to 42.11	75.09 to 78.11	1689-1995	RW/MXD	STD	Wilson et al. (2007)	JA: 0.70 (0.77)
Mangazeja	MAN	66.42	82.18	1328-1990	MXD	RCS	Schneider et al. (2015)	JA: 0.79 (0.69)
Asia Grid 3	ASGrid3	40.15 - 46.15	80.15 - 88.15	800-1989	MX	RCS/STD	Cook et al. (2012)	JA: 0.44 (0.69)
Asia Grid 12	ASGrid12	48.15 - 54.15	80.15 - 88.15	800-1989	MX	RCS/STD	Cook et al. (2012)	JA: 0.62 (0.84)
Altai MXD	ALT	50.00	88.00	750-2007	MXD	RCS	Schneider et al. (2015)	MJJA: 0.65 (0.72)
Asia Grid 4	ASGrid4	40.15 - 46.15	90.15 - 98.15	800-1989	MX	RCS/STD	Cook et al. (2012)	JJ: 0.60 (0.81)
Asia Grid 13	ASGrid13	48.15 - 54.15	90.15 - 98.15	1024-1989	MX	RCS/STD	Cook et al. (2012)	JA: 0.52 (0.80)
Mongolia	OZN	51.15	99.04	931-2005	RW	RCS	Davi et al. (2015)	JJ: 0.52 (0.62)
Taymir	TAY	72.01	102.00	755 - 1997	RW	RCS	Jacoby et al. (2000); Used in D'Arrigo et al. (2006)	JJ: 0.45 (0.45)
Asia Grid 5	ASGrid5	40.15 - 46.15	100.15 - 108.15	800-1989	MX	RCS/STD	Cook et al. (2012)	JA: 0.39 (0.59)
Asia Grid 14	ASGrid14	48.15 - 54.15	100.15 - 108.15	1396-1989	MX	RCS/STD	Cook et al. (2012)	JA: 0.58 (0.70)
Asia Grid 6	ASGrid6	40.15 - 46.15	110.15 - 118.15	800-1989	MX	RCS/STD	Cook et al. (2012)	JA: 0.25 (0.50)
Asia Grid 15	ASGrid15	48.15 - 54.15	110.15 - 118.15	1396-1989	MX	RCS/STD	Cook et al. (2012)	JA: 0.15 (0.46)
Asia Grid 7	ASGrid7	40.15 - 46.15	120.15 - 128.15	1024-1989	MX	RCS/STD	Cook et al. (2012)	MJJA: 0.30 (0.66)
Asia Grid 16	ASGrid16	48.15 - 54.15	120.15 - 128.15	1510-1989	MX	RCS/STD	Cook et al. (2012)	MJJA: 0.04 (0.41)
Asia Grid 8	ASGrid8	40.15 - 46.15	130.15 - 138.15	1510-1989	MX	RCS/STD	Cook et al. (2012)	JJ: 0.58 (0.68)
Asia Grid 17	ASGrid17	48.15 - 54.15	130.15 - 138.15	1510-1989	MX	RCS/STD	Cook et al. (2012)	MJJA: 0.41 (0.57)
Asia Grid 9	ASGrid9	40.15 - 46.15	140.15 - 148.15	1510-1989	MX	RCS/STD	Cook et al. (2012)	JJ: 0.67 (0.71)
Asia Grid 18	ASGrid18	48.15 - 54.15	140.15 - 148.15	1510-1989	MX	RCS/STD	Cook et al. (2012)	JJ: 0.52 (0.53)
North Japan	NJAP	43-51	142-145	1640-1993	RW/MXD	STD	D'Arrigo et al. (2014)	JJ: 0.65 (0.62)
Yakutia	YAK	67.27 to 70.33	142.37 to 150.17	1342 - 1994	RW	RCS	Hughes et al. (1999); Used in D'Arrigo et al. (2006)	June: 0.30 (0.62)

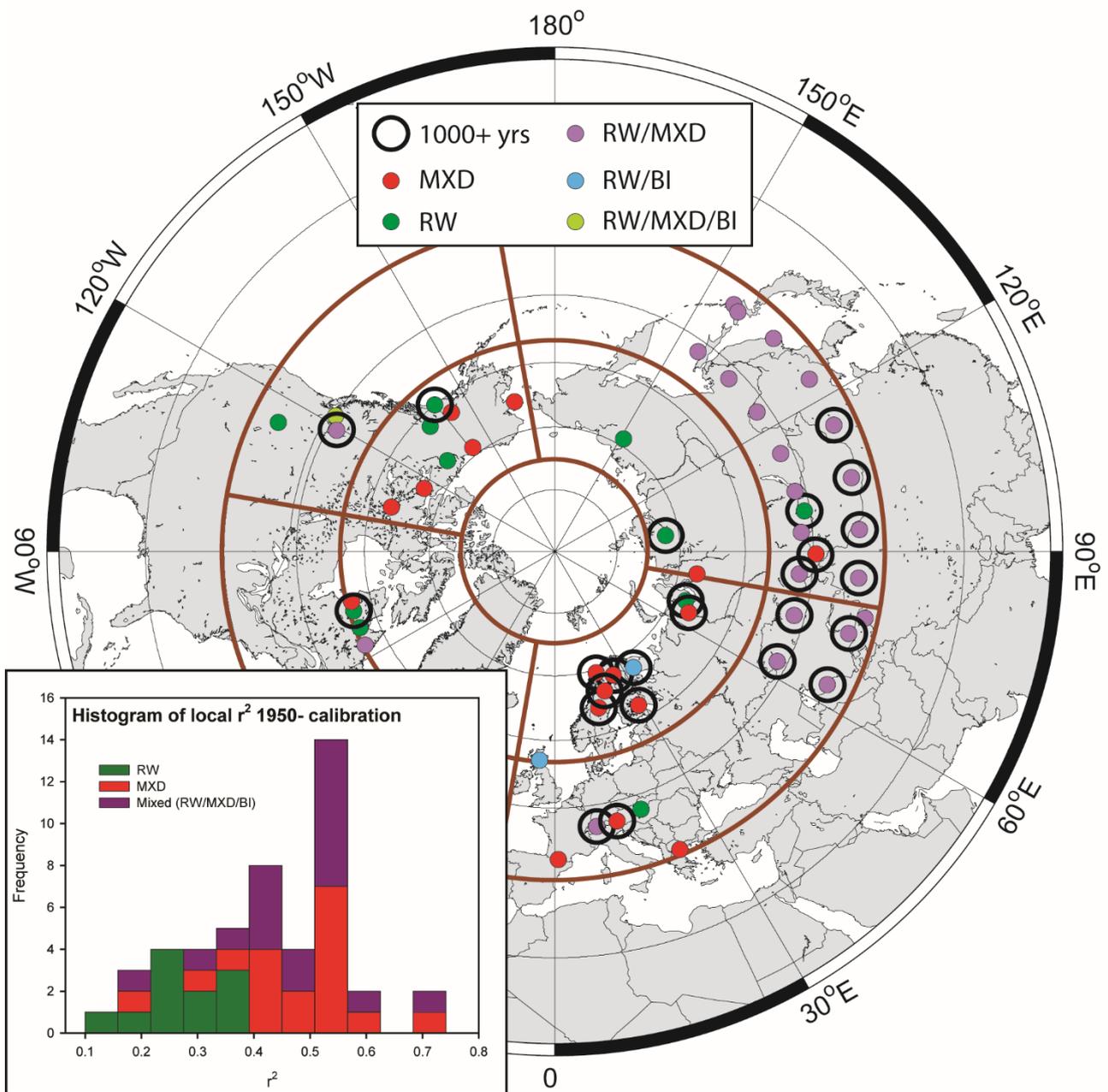
**Table 1:** Meta Information for tree-ring reconstructions used in the current N-TREND2015 network. See Appendix A [Table A1](#) for brief information on how each series was derived. Final column presents correlations for 1901-present (1950-present) for all series against the local 1 degree (or larger where relevant) CRUTS3.2 (Harris et al. 2014) mean temperature grid for the TR site's optimal summer season. Histogram of post 1950 correlation squared values is embedded in Figure 1.

Rank	Warm	Anomaly	$\pm 2\sigma$	Cool	Anomaly	$\pm 2\sigma$
1	1994-2003	0.34	0.25	1812-1821	-1.03	0.19
2	1946-1955	0.30	0.19	1832-1841	-1.00	0.19
3	1161-1170	0.27	0.23	1462-1471	-0.96	0.19
4	1980-1989	0.19	0.19	1695-1704	-0.94	0.19
5	1425-1434	0.17	0.19	1452-1461	-0.90	0.19

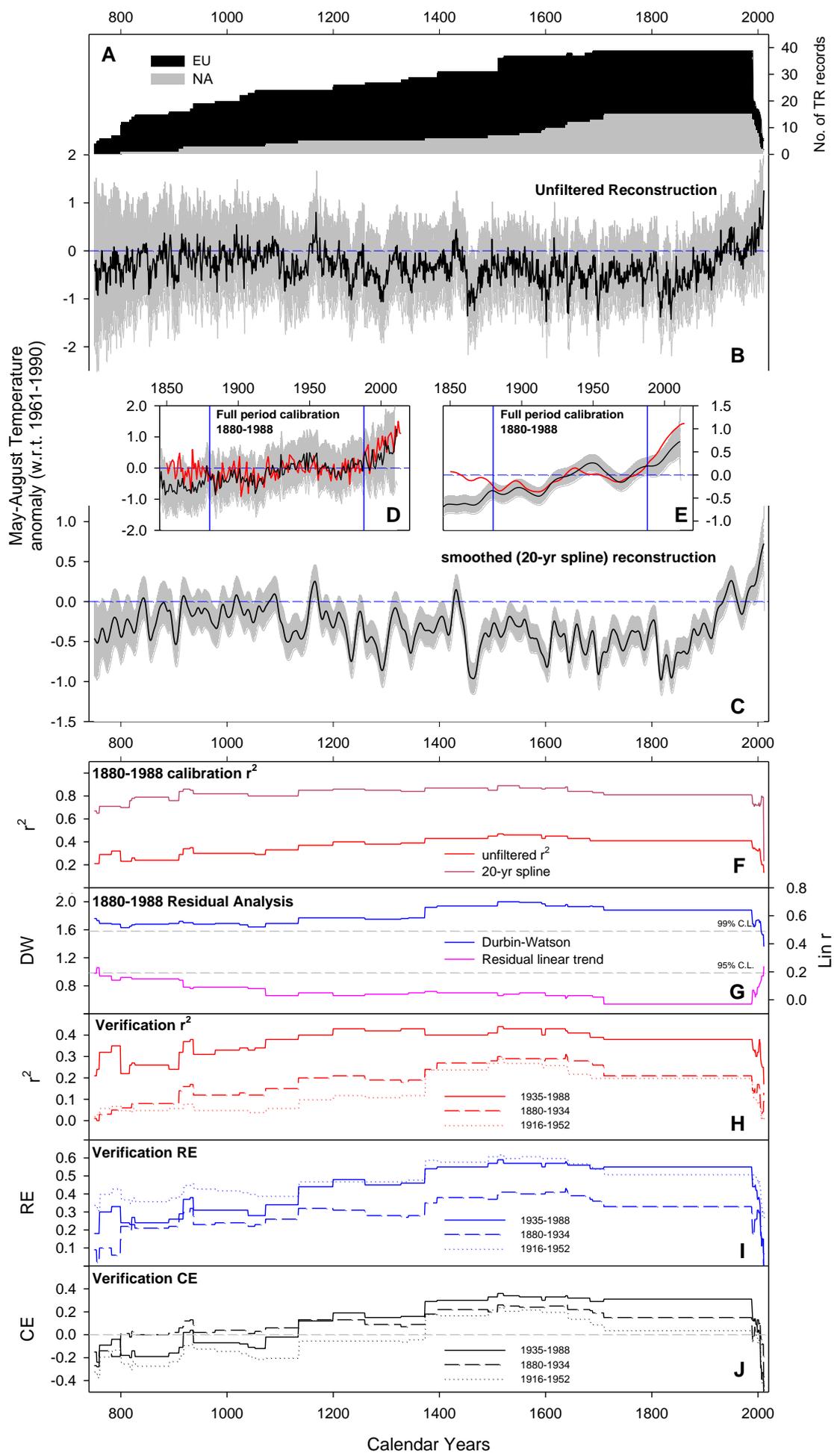
**Table 2:** [Five5](#) warmest and coldest non-overlapping decadal periods in N-TREND2015. Calculated over 918-2004. Temperature anomalies (plus 2-sigma error) shown with respect to 1961-1990.

Event	Gao/Crowley	Esper	Sigl
Unknown	1227		1229
Samalas (Indonesia)	1258		1257
Katla (Iceland)		1262	
Unknown	1284		1286
Unknown			1344
Oraefajokull (Iceland)		1362	
Kuwae? (Vanuatu)	1459		1458
Barbarbunga (Iceland)		1477	
Hekla (Iceland)			
Colima (Mexico)	1584		
Ruiz (Columbia)			1594
Huaynaputina (Peru)	1600	1600	1600
Katla (Iceland)		1625	
Furnas (Azores)		1630	
Komaga-Take (Japan)		1640	
Parker Peak (Philippines)	1641		
Usu (Japan)		1663	
Shikotsu (Japan)		1667	
Unknown			1694
Fuji (Japan)		1707	
Shikotsu (Japan)		1739	
Katla (Iceland)		1755	
Laki (Iceland)	1783		
Unknown	1809		1809
Tambora (Indonesia)	1815	1815	1815
Babuyan (Philippines)	1831		
Cosiguina (Nicaragua)	1835	1835	
Shiveluch (Russia)		1854	
Makian (Indonesia)			
Askja (Iceland)		1875	
Krakatau (Krakatau)		1883	
Santa Maria (Guatemala)		1902	
Ksudach (Russia)		1907	
Katmai (USA)		1912	
Colima (Mexico)		1913	
Bezymianny (Russia)		1956	
Pinatubo (Indonesia)			1991

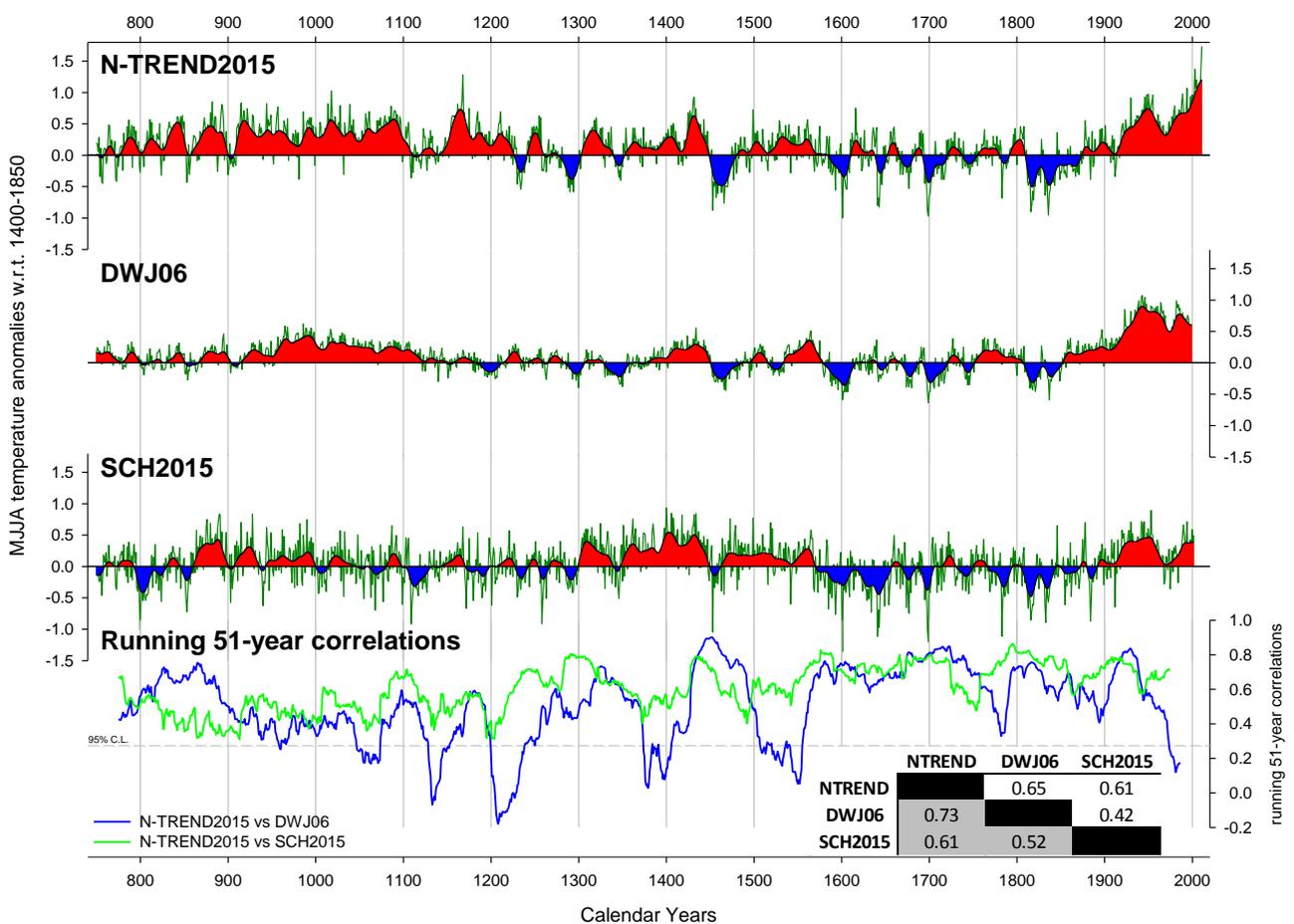
**Table 3:** Three different lists of significant volcanic events derived from: Gao/Crowley = the earliest year with common sulphate deposition values in both hemispheres  $> 15 \text{ kg km}^{-2}$  detailed in both Gao et al. (2008), Crowley and Unterman (2013); Esper = The large volcanic events list in Esper et al. (2013) which were identified via historical documents and had a Volcanic Explosivity Index  $\geq 5$ ; Sigl = volcanic events, detailed Sigl et al. (2013), with sulphate deposition values in both hemispheres  $> 15 \text{ kg km}^{-2}$ .



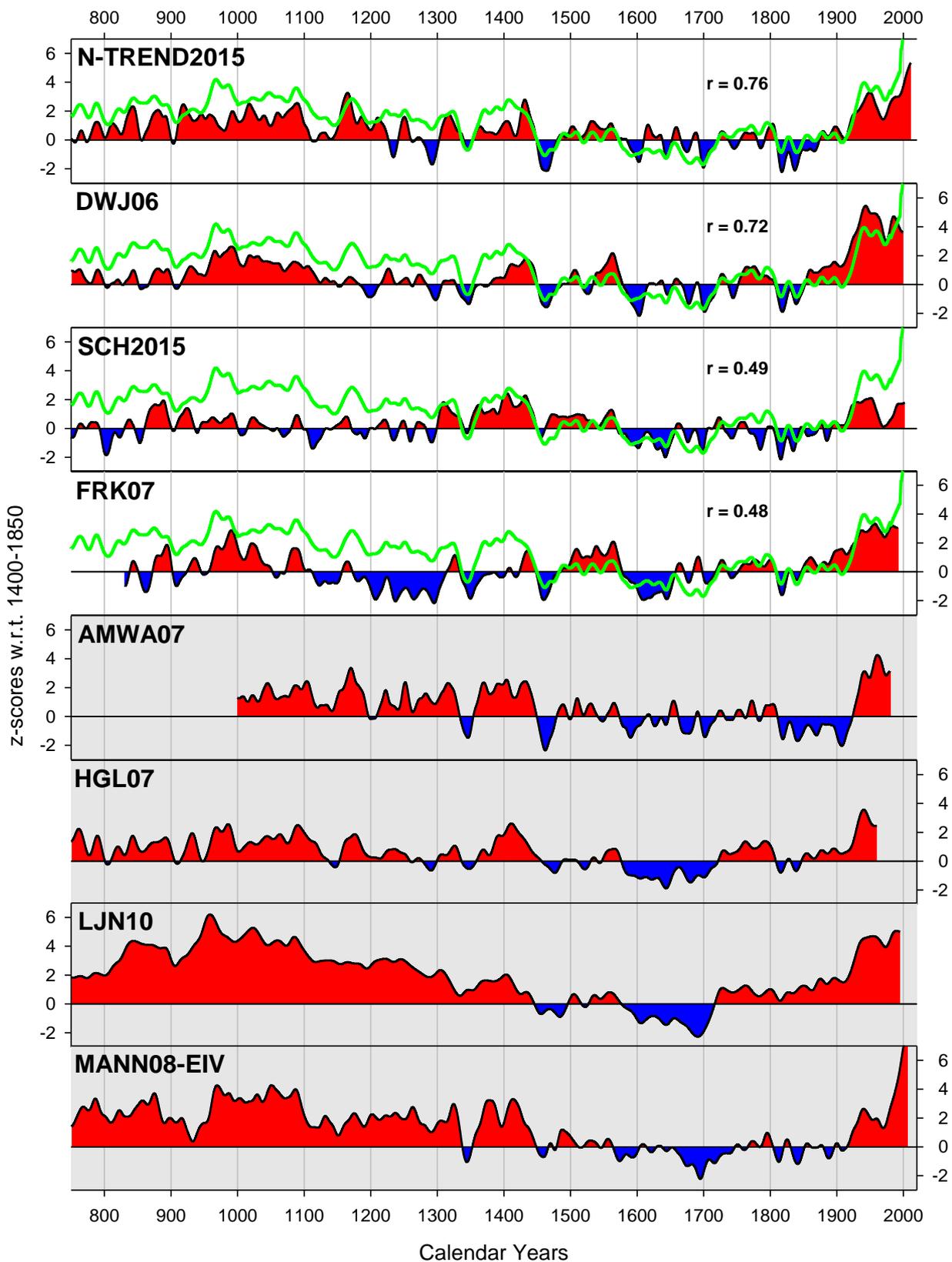
**Figure 1:** Location map of individual tree-ring reconstructions used in the current N-TREND2015 network. See Table 1 for site information and Appendix Table 1A for brief information on how each series was derived. Brown lines denote the high/low latitude and four longitude quadrants. Inset histogram shows distribution of post 1950  $r^2$  values for local calibration, for site specific optimal season, to the CRU TS3.2 grid (Harris et al. 2014 - see Table 1 for grid specifications).



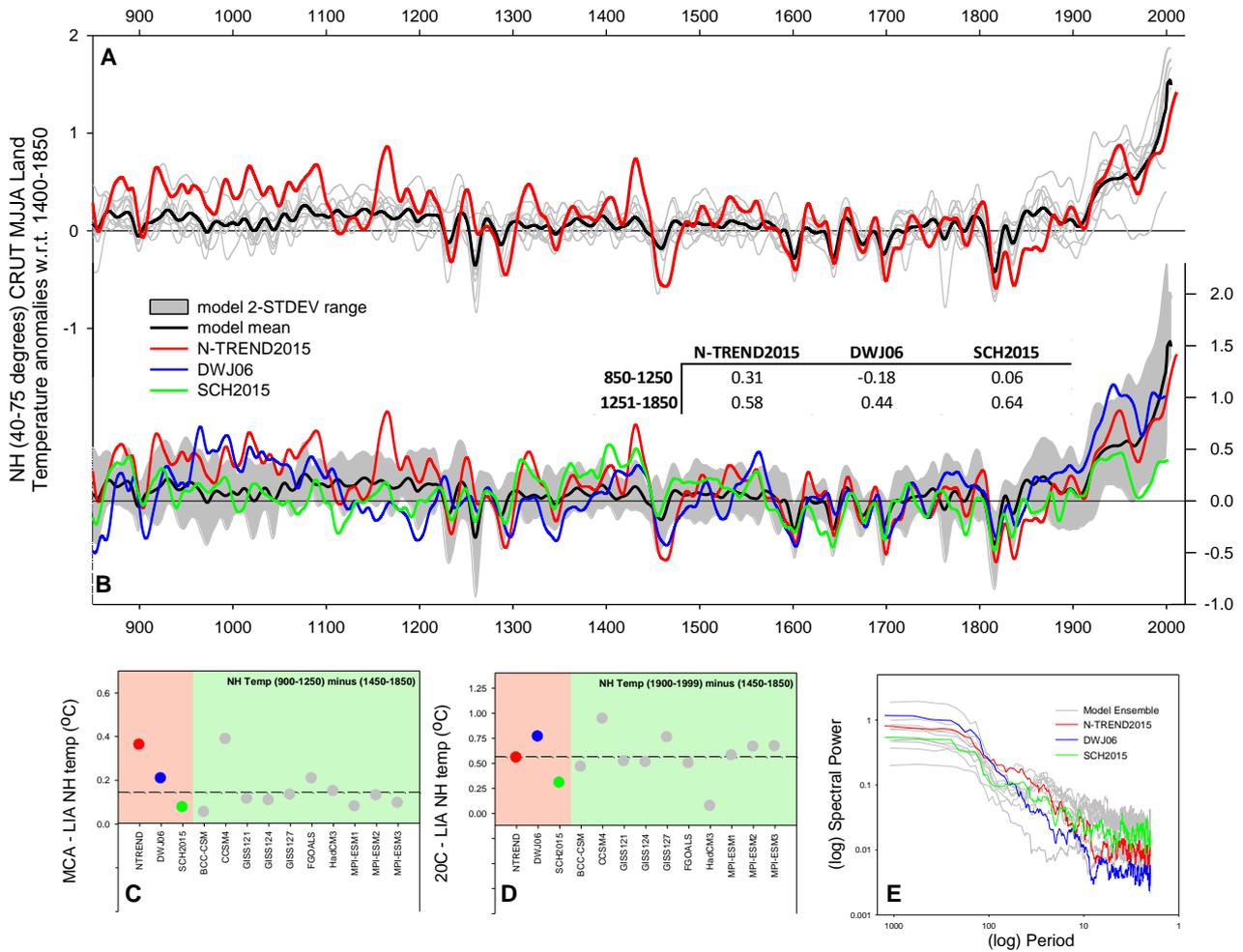
**Figure 2:** N-TREND2015 reconstruction and assessment metrics. **Upper panel:** TR series replication (A), unfiltered and filtered (20-year spline) mean May-August reconstruction with 2-sigma error bar for both 750-2011 (B+C) and 1850-2011 periods (D+E). Errors are derived from the RMSE of the full calibration period (1880-1988) residuals (post scaling), calculated for each nested. An additional error component is also incorporated to estimate the extra uncertainty of the early weaker replicated periods of each individual input series (see main text [and Appendix Figure B2](#)). **Lower panel:** F: 1880-1988 calibration  $r^2$  values for unfiltered and filtered (20-year spline) versions; G: 1880-1988 residual analysis – Durbin and Watson (DW) values and correlation of linear trend (Lin r) of residuals; H: Validation  $r^2$  values [for the using the](#) 3 validation periods 1880-1934, 1935-1988 and 1916-1952; I: As [for H](#), but for RE; J: As [for H](#), but for CE. See Appendix Table B1 for detail reconstruction metrics for each nest.



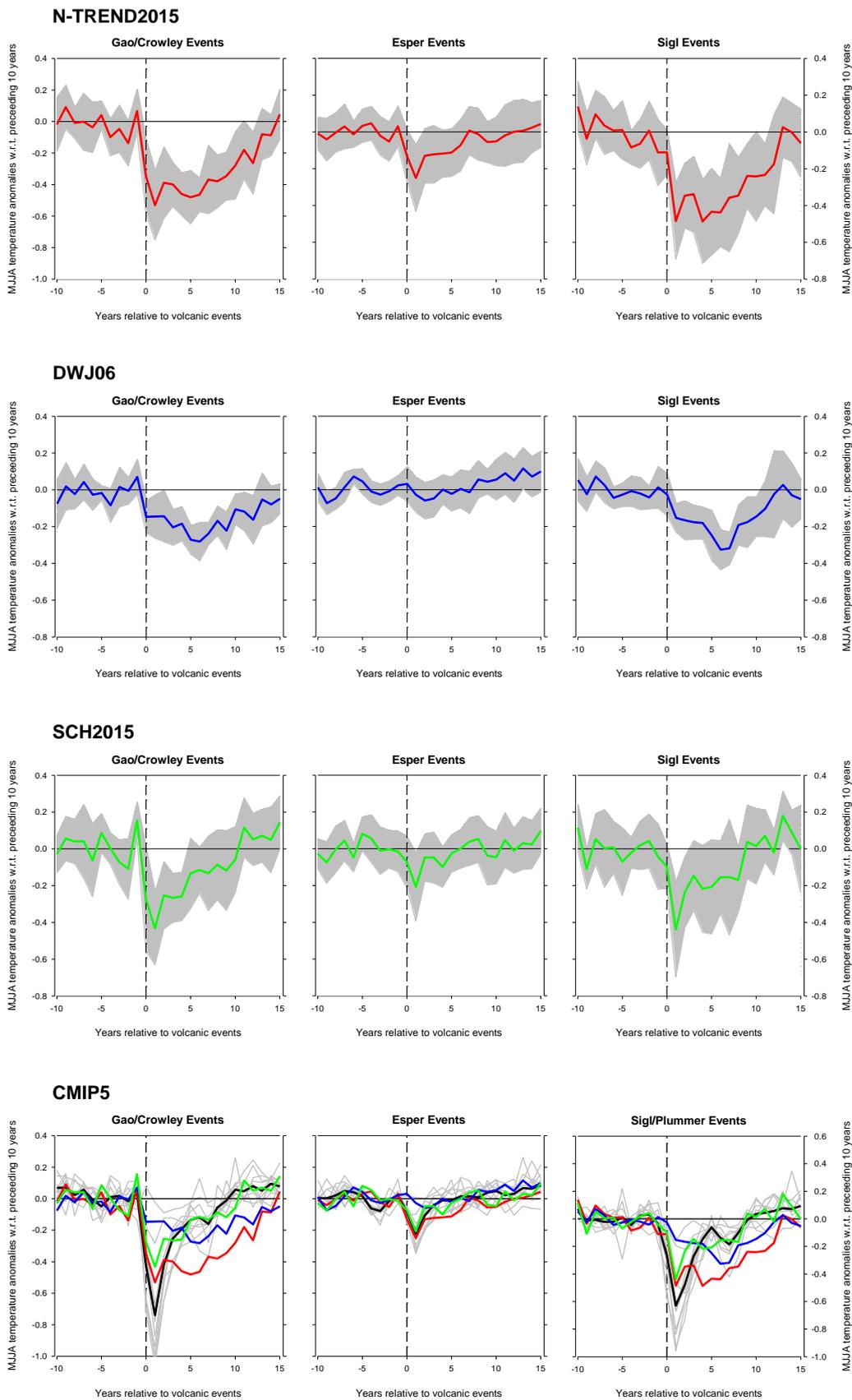
**Figure 3:** Tree-ring based summer temperature NH reconstructions. N-TREND2015, DWJ06 (D’Arrigo et al. 2006) and SCH2015 (Schneider et al. 2015). The series have been scaled to CRUTEM (40-75) MJJA land temperatures over the 1880-1988 period and expressed as anomalies w.r.t. 1400-1850. Smoothed function is a 20-year spline. Lower panel shows sliding 51-year correlation values between N-TREND2015 and the other two reconstructions. A correlation matrix (918-1999) is also shown for unfiltered and filtered (20-year spline - in grey) versions.



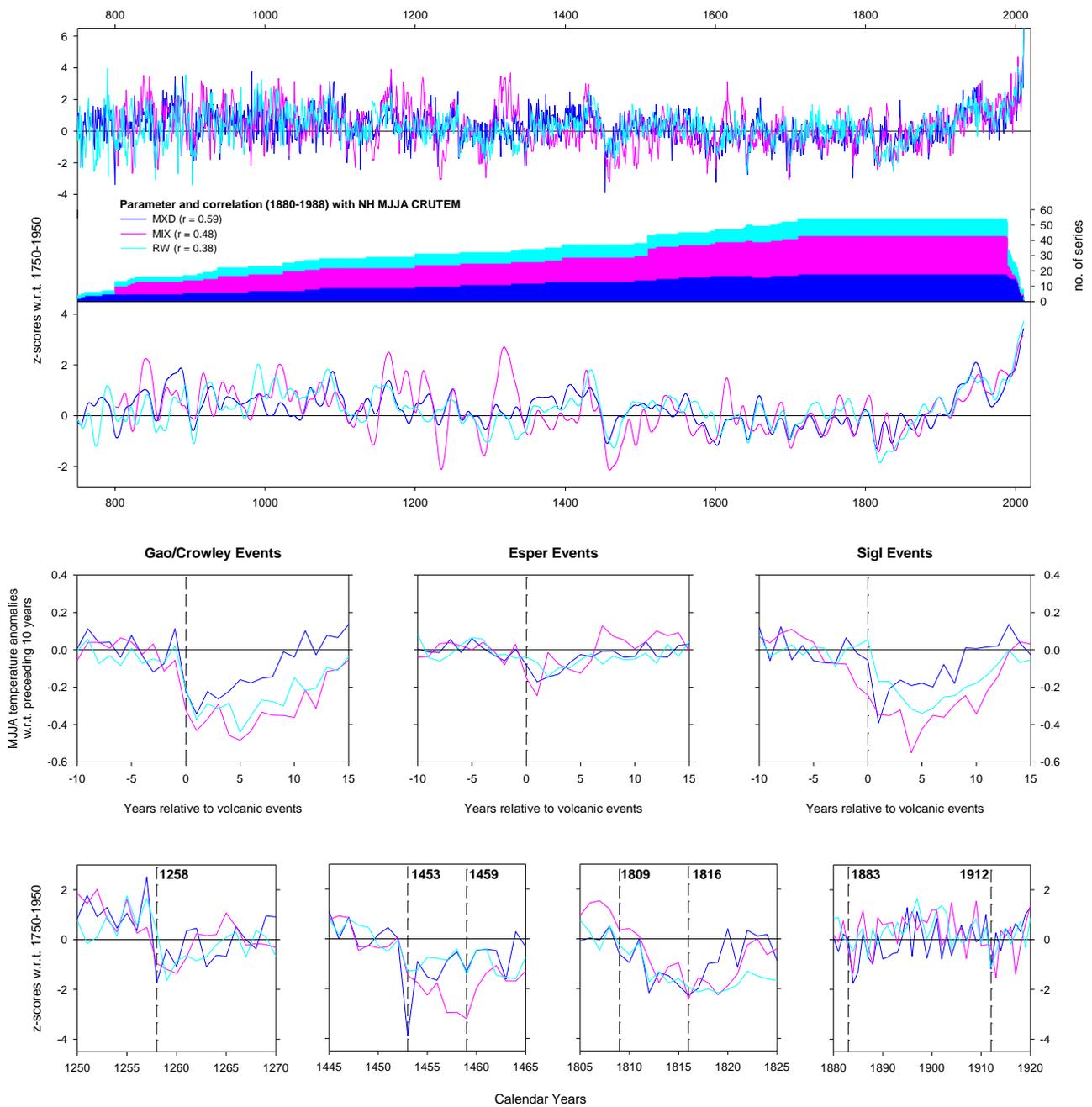
**Figure 4:** Comparison of TR based reconstructions (N-TREND2015, DWJ06, SCH2015 and FRK07) with select multi-proxy (MP) reconstructions (Amman and Wahl 2007; Hegerl et al. 2007; Ljungqvist 2010; Mann et al. 2008) published since 2007. All smoothed (20-year spline) series have been normalised to the 1400-1850 period to enhance differences in the Medieval and recent periods. The green-time series is the normalised average of AMWA07, HGL07, LJN07 and MANN08-EIV and the correlation shown is between this mean MP composite and the individual TR NH series over the 918-1992 period.



**Figure 5: Upper:** Comparisons of N-TREND with 10 CMIP5 models. Time series have been smoothed using a 20-year spline and transformed to anomalies w.r.t. 1400-1850. **A:** N-TREND versus the 10 CMIP5 models; **B:** N-TREND, DJW06 and SCH2015 compared to the 2-STDEV range of the 10 CMIP5 models. Inset table shows correlations between each TR NH series and the model mean for the two periods 850-1250 and 1251-1850; **C:** Temperature difference (°C) from subtracting the mean of 1450-1850 from the mean of 900-1250 (NB. NGOALS starts in 1000). Dashed line is the model mean of the transition difference between the two periods; **D:** Temperature difference (°C) from subtracting the mean of 1450-1850 from the mean of 1900-1999. Dashed line is the model mean of the difference between the two period transition; **E:** Spectral Power plot of N-TREND2015, DJW06 and SCH2015 and the 10 CMIP5 models calculated over the 900-1999 period (1000-1999 for FGOALS).



**Figure 6:** Superposed epoch analysis for N-TREND2015, DWJ06, SCH2015 and the CMIP5 10 models assessing mean response to major volcanic events. Lower panels compare individual model runs (plus their mean) with each reconstruction. Table 3 lists the difference years extracted from Gao et al. 2008; Crowley and Unterman 2013; Esper et al. 2013; Sigl et al. 2013 for the three different lists.



**Figure 7: Upper:** Parameter specific NH composites using the N-TREND database. Method used was 2-quadrant (North American and Eurasia) weighting with variance stabilisation (see Appendix Figure C1); **Middle:** SEA (see Figure 6) using each parameter composite; **Lower:** Parameter composite comparison for select periods with significant volcanic events (see Table 3).

## \*Highlights (for review)

- A new northern hemisphere (NH) tree-ring (TR) summer temperature reconstruction
- The most replicated (54 data-sets) TR NH millennial reconstruction to date
- A product of the **N**orthern Hemisphere **T**ree-Ring **N**etwork **D**evelopment (N-TREND) consortium
- The N-TREND2015 reconstruction is robust at all time-scales from 918 – 2004
- Medieval temperature estimates are warmer than expressed by CMIP5 models



## Detailed response

I have minor comments on some of the SEA analyses of the record of annual temperature responses to major volcanic eruptions. All eruptions prior to 1600 have minor differences in dating between the records and it appears that each of the three composite records evaluated uses a different combination of eruptions. Would it not have been better to restrict the comparison to the same well-dated eruptions for this analysis to provide completely comparable data sets?

***The Superposed Epoch Analysis (SEA) has been refined. We have now clarified the criteria for the identified volcanic events. Specifically from the Gao/Crowley and Sigl papers, we identify significant volcanic events with common sulphate deposition values in both hemispheres  $> 15 \text{ kg km}^{-2}$ . The Esper events were independently derived using historical documents for events with VEI values  $\geq 5$ . Please note that we have now removed reference to Plummer et al (2012). With the above refinement, the number of significant volcanic events for the Gao/Crowley and Sigl lists has been reduced so the SEA has been re-done. The overall results have not changed (Figure 6) although the Sigl based SEA is more in line with Gao/Crowley results.***

***The differences between these three event lists is exactly what we wanted to test. i.e. how sensitive is the SEA to the use of the different lists. We therefore feel that nothing would be gained from focussing only on common well dated set of events as this has been done elsewhere.***

Also it would be interesting to compare SEA for standardized RW and MXD data from the same site to demonstrate differences in the lagged response in these archives.

***This is a good point but is beyond the scope of the current paper which utilises published reconstructions for each location, some of which are a fusion of different TR parameters. However, a recent paper by Esper et al. undertakes this type of comparison:***

***Esper, J., Schneider, L., Smerdon, J., Schöne, B., Büntgen, U., 2015. Signals and memory in tree-ring width and density data. *Dendrochronologia*, 35, 62-70.***

In addition to the NH reconstruction this paper does include composite N American and Eurasian reconstructions in Appendix Figure 2C. It might be useful to include this figure in the main text with appropriate caveats about the relatively low replication unless it is planned to treat this more fully in the subsequent paper. This figure appears to be a major reason for the discussion of Atlantic variability on p 10.

***This is a fair comment but the focus of the paper is on the long term context of the northern hemisphere composite record. We purposely make only a brief reference to the North American vs Eurasian comparison (Appendix Figure C3) and minimal discussion w.r.t. the expression of the medieval warmth in NE North America and NW Europe (Appendix Figure C2). This latter feature will be discussed in more detail in the part II companion paper (Anchukaitis et al.) due to its spatial expression. However, another reason for not going into too much detail is the fact that there is only 1 record from NE North America which appears anonymously warm compared to other records and we feel at this time it is better to take a cautious approach and not over interpret this region until more data exist. We therefore make no changes.***

Page	Para	Line	
2	1	6	temperature

***No change made – did not find a problem with current wording.***

3	2	17	time scales
---	---	----	-------------

***So corrected***

3	2	21	data-sparse
---	---	----	-------------

***So corrected***



- 3 2 22 complement  
**So corrected**
- 3 4 4 tree rings  
**So corrected**
- 4 1 3-4 allow...improved understanding  
**So corrected**
- 4 3 13 to record.....events accurately,  
**So corrected**
- 4 3 15 delete final comma  
**So corrected**
- 4 4 2 temperature-sensitive (similar changes are needed throughout where compound adjectives used)  
**So corrected – throughout paper**
- 4 4 2 improve characterisation of (avoid split infinitive).  
**We have left with current wording as we feel it flows better**  
**“.....we focus entirely on temperature- sensitive tree-ring (TR) records to both reduce and better characterize many of the uncertainties detailed above.”**
- 5 1 5 -a predominantly  
**So corrected**
- 5 2 5 needs to be explicit- were the reconstructions used or the TR data on which the reconstructions were based?  
**We have clarified the wording in the text – new text is highlighted,**  
**“we utilise mostly published TR temperature reconstructions (or chronologies used in published reconstructions) that start prior to 1750”**  
**and have edited Appendix Table 1 to detail what N-TREND2015 series are published reconstructions or published chronologies. However, many reconstructions are simply re-scaled/calibrated chronologies so the differences are moot.**
- 5 2 10 includes TR records or TR based reconstructions? See above. I think the latter.  
**As above**
- 5 3 Therefore the paper accepts the earlier reconstructions without reprocessing?  
**Yes – see next comment – using the latest published version of these data assumes that the original authors have undertaken the most robust analysis possible.**
- 5 3 8 and therefore still retains any biases (see intro) due to differences in the way that these original data were processed.  
**The biases referred to in the introduction were related to the multi-proxy approach and the problems of mixing proxies of different seasons, temporal resolution and carrying dating precision. We make a clear statement that our strategy uses published chronologies and/or reconstructions that “explicitly incorporates the knowledge of the original authors as they [should have] derived the most robust reconstruction possible from the available data at that particular location”**
- 5 3 11 only three involve BI, perhaps a brief comment needed on its value?  
**The reference to BI in the main text has not been changed as it makes reference to the relevant methodological papers. However, in Appendix Table 1, for the IBC site – the first listed using BI – we have inserted the following clarifying sentence, “BI - Blue Intensity is a relatively new tree-ring parameter (McCarroll et al. 2002; Campbell et al. 2011; Björklund et al. 2014; Rydval et al. 2014) that utilises the intensity of the reflectance of blue light of the TR latewood to derive an estimate of lignin content in latewood cells. As lignin content is reflected by density, raw BI and MXD are inversely correlated and both calibrate similarly to summer temperatures.”**



5 3 last by 1000 or before 1000?

**Slight re-word of sentence to "The common period for all 54 records is 1710-1988; 38 records extend back to 1500 while 23 records extend back to at least 1000. Substantially more TR records are available for Eurasia (39) compared to North America (15) with a distinct imbalance of data at 1000 (20 vs. 3 – Figure 2a)."**

6 2 8-9 do we see these results?

**Yes – these results are detailed later in the paper and in Appendix table B1**

6 2 12 perhaps label calibration trials as A and B to clarify this point?

**We believe Appendix table B1 already clearly details the results. These are not trials – just co-occurring different approaches/periods for model validation.**

6 3 last reference to where these data are shown?

**This section details the reconstruction methodology, so reference to results comes later in the paper. The error estimates are detailed in Appendix table B1 and Appendix Figure B2.**

6 4 3 comma after and?

**So corrected**

Fig 2 caption 4 nest or nested period? Better explanation needed of the lower diagrams?

**Changed to "nest". Edits have been made to figure caption for clarification.**

9 2 N-TREND also shows a greater range of temperature variability than the other two reconstructions.

**Yes – agreed – but this is discussed later w.r.t. Figure 5.**

9 2 1 but the 1500s?

**Not clear what the reviewer is wanting us to say about the 1500s.**

9 2 15-17 but there are stronger deviations in the multidecadal variability between TREND and DWJ decadal variability

**Yes – see above comment – this is discussed w.r.t. Figure 5.**

10 1 App C2 should there be some measure of uncertainty with these regional records?

**As Part II focuses on the spatial reconstruction of temperature, we feel that this figure should stay within Appendix C. As these series are not specifically calibrated to appropriate quadrant instrumental mean series, it would not be possible to place uncertainty ranges upon them. The periods where the number of records is less than three is already highlighted – denoting these periods should be interpreted cautiously.**

10 1 6 end of the 12th century? But after a brief hiatus continues to the mid-14th century?

**We believe the current wording is valid**

10 1 11 where is MCA defined in this paper

**MCA has now been changed to Medieval throughout the paper.**

10 1 11 is being or has recently been?

**Changed.....has recently been....**

10 1 17 in which specific periods are these megadroughts? What about the potential influence of the Pacific here? It is strange there is no reference to possible differential effects between the influences of the Atlantic and the Pacific.

**All mention of megadroughts have now been removed from this paragraph and we now focus specifically on temperature.**

10 3 3 end of 10th or end of 11th?

**Taking ALL the different records into account, one must conclude the 10<sup>th</sup> century – no changed made to current wording.**

10 3 12-15 how many of these MP reconstructions have proxies with a winter or annual temperature variables?

**This is difficult to address in detail. This would require a detailed review of the original MP publications and data used. We believe that this detail would not change anything towards our current paper. The main points**



**still hold w.r.t. MP reconstructions – that they are a mix of proxies of varying season (albeit possibly weighted towards the summer season), include records with dating uncertainty and are not annually resolved. Our record is exactly calendar dated, has annual resolution and uses only summer sensitive TR proxy records.**

Fig 5 Caption 6 what is the transition between? Reference to text needed?

**Caption has been clarified: Dashed line is the model mean of the difference between the two periods.**

11 5 4 medieval period?

**So corrected**

12 1 6 disagreement of? Or between estimates of?

**Sentence has been clarified to, “In a recent appraisal of proxy and CMIP5 model output over the Last Glacial Maximum, mid-Holocene, and the recent period, a substantial disagreement between paleoclimate and modelled estimates was reported, although the models were marginally better at representing modern late Holocene climate (Harrison et al. 2015).”**

12 2 14 three lists and five sources?

**The Gao/Crowley information was combined to identify hemispherically relevant volcanic events. These data are used to drive volcanic forcing within the CMIP5 model runs. The Esper list of events is an independently derived list of significant volcanic events. We have now dropped reference to the Plummer paper and focus on the Sigl paper only as they identify significant hemispheric relevant volcanic events from two new ice cores in Greenland and Antarctic and provide a new set of independent refined dates to compliment the Gao/Crowley list.**

13 2 10 Table 3 gives 1912 as Katmai in Alaska

**So corrected**

13 3 3 This misses the point. They also developed a new improved version of the IPSL-CM5A-LR model for volcanic forcing which simulated less cooling from the eruptions than those the CMIP models used. It was development of this new simulation that reconciled the previous dissimilarities between the tree-ring and volcanic estimates of cooling.

**This final paragraph has been re-worded for clarification, “In a recently published new TR reconstruction of NH JJA mean temperatures, Stoffel et al. (2015) utilised 30 regional records (using both RW and MXD) to assess large-scale volcanic response expressed in their composite record by comparison to a modified version of the IPSL-CM5A-LR model (Schmidt et al. 2012). Their model version simulated less cooling and they stated that with respect to volcanic forcing their work had reconciled previously reported dissimilarities between climate simulations and TR based reconstructions. Although the “missing ring” hypothesis (Mann et al. 2012) is now thoroughly rejected (Anchukaitis et al. 2012; D’Arrigo et al. 2013; Esper et al. 2015), Stoffel et al. (2015) and the results detailed herein, indicate that there are still substantial uncertainties that need to be addressed before we fully understand the temporal and spatial dynamical response of the climate system to volcanic forcing as well as the response of specific TR parameters to volcanically forced cool summers.”**

13 5 1 represents

**So corrected**

14 1 5 outcome= reconstruction?

**We have changed to “record”**

14 1 6 statistical methods?

**So corrected**

14 1 7 delete commas

**So corrected**

14 2 1 how well?

**So corrected**

14 2 3 local gridded temperature data

**So corrected**



14 2 7 what high frequency bias- doesn't RW decrease high frequency variability - should this be the reduced high frequency bias?

**Sentence re-worded slightly for clarity, "Considering as well, the well-known lagged high frequency biases of RW data....."**

14 2 9 robust estimates of interannual variation This needs to be qualified and is in the following sentence. Delete material about MXD and BI and run on to particularly relevant clause. The comment re MXD and BI could come in a following sentence.

**We believe the minor sentence modification above improves the sense of the overall sentence.**

14 2 19 rephrase. This situation can further be potentially confused by

**So corrected**

14 3 3 Stoffel at

**So corrected**

14 3 6 millennial length?

**So corrected**

15 1 1 delete the

**So corrected**

15 1 4 exist

**So corrected**

15 1 6 BI is somewhat of a sleeper in this whole discussion as its potential to replace MXD is never specifically mentioned and only three series use BI data. This needs to be remedied earlier in the text .

**This is a good point. As stated above, we now make a specific reference in Appendix A Table A1 w.r.t. the IBC reconstruction about BI actually is. The relevant methodological based papers are already referred to within the text.**

16 Briffa et al 1998 for consistency include second initials

**2<sup>nd</sup> initials deleted from all references**

16 Cohen Letters

**So corrected**

19 Fritts Academic Press?

**So corrected**

20 Hegerl 2006 journal title in full

**So corrected**

20 Jones 1995 Letters

**So corrected**

21 Jungclaus journal title in full

**So corrected**

21 Landrum journal title in full

**So corrected**

22 Mann 1999 Letters

**So corrected**

23 Meko journal title in full

**So corrected**

24 Pongratz journal title in full

**So corrected**

24 Rohde journal title in full

**So corrected**

24 Rossi 2012 Caps in Journal title?

**So corrected**



24 Rossi 2013 Caps in Journal title?

**So corrected**

24 Schmidt journal title in full

**So corrected**

25 George out of Order (2 references)

**So corrected**

25 Steinhilber journal title in full

**So corrected**

25 Tingley duplicate authors

**So corrected**

26 Trouet Oscillation

**So corrected**

26 Wu are authors complete? Journal title in full

**So corrected**

28 Table 1 Font size is barely readable without enlargement

**Will await advice from journal if paper is accepted**

28 Table 2 Five?

**So corrected**

32 4 word(s) missing after nested?

**So corrected**

32 6-10 what precisely is calibrated here?

**We believe this is self-explanatory**

33-35 Figs 3-5 Figs 3 and 5 calibrated to 1450-1850 whereas Fig 4 is to 1000-1960. Is there a specific reason for this? What is the absolute difference between these two means?

**We agree there was an inconsistency here. Figure 4 has now been re-drafted and the z-scores calculated w.r.t. the 1400-1850 period. This period was chosen to enhance differences between the records for comparative purposes in the Medieval and recent periods.**

34 6 what is the transition?

**See earlier – caption text has been changed to clarify.**

36 3 lists the different years or shows the differences in the years? There are only 3 cases ( 1600, 1815, 1835) where all three records show the same date and seven events ( 1584, 1640, 1667, 1739, 1809, 1883, 1912) where two of the dates agree and seven ( all before 1600 except for Laki) where there are minor differences in the dates . So why use all of these rather than those events which are synchronous between the records? 22 of the events are only identified in a single record

**Please see very first comment – we have refined the SEA analysis. We believe that the reviewer has not quite understood what we were trying to do with the analysis. i.e. the choice of chosen events a is very important for such analyses – see Esper et al. (2013) for further clarification of this issue.**

A2 1 Büntgen also misspelt in Table 1

**So corrected**

**Supplementary Appendices**

[Click here to download Supplementary Data: Wilson et al - Appendices.docx](#)

**Supplementary Appendices - track changes**

[Click here to download Supplementary Data: Wilson et al - Appendices - track changes.docx](#)