

Supplementary Materials for

Old World megadroughts and pluvials during the Common Era

Edward R. Cook, Richard Seager, Yochanan Kushnir, Keith R. Briffa, Ulf Büntgen, David Frank, Paul J. Krusic, Willy Tegel, Gerard van der Schrier, Laia Andreu-Hayles, Mike Baillie, Claudia Baittinger, Niels Bleicher, Niels Bonde, David Brown, Marco Carrer, Richard Cooper, Katarina Čufar, Christoph Dittmar, Jan Esper, Carol Griggs, Björn Gunnarson, Björn Günther, Emilia Gutierrez, Kristof Haneca, Samuli Helama, Franz Herzig, Karl-Uwe Heussner, Jutta Hofmann, Pavel Janda, Raymond Kontic, Nesibe Köse, Tomáš Kyncl, Tom Levanič, Hans Linderholm, Sturt Manning, Thomas M. Melvin, Daniel Miles, Burkhard Neuwirth, Kurt Nicolussi, Paola Nola, Momchil Panayotov, Ionel Popa, Andreas Rothe, Kristina Seftigen, Andrea Seim, Helene Svarva, Miroslav Svoboda, Terje Thun, Mauri Timonen, Ramzi Touchan, Volodymyr Trotsiuk, Valerie Trouet, Felix Walder, Tomasz Ważny, Rob Wilson, Christian Zang

Published 6 November 2015, *Sci. Adv.* **1**, e1500561 (2015)
DOI: 10.1126/sciadv.1500561

The PDF file includes:

Introduction

OWDA as a scientific advancement over previous work

Gridded monthly scPDSI target field

OWDA tree-ring network

Climate sensitivity of OWDA tree-ring chronologies

Validation of OWDA tree-ring climate response

Augmenting tree-ring chronologies with historical tree-ring data

Standardizing OWDA tree-ring data for climate reconstruction

Standardizing OWDA historical/modern tree-ring data

Estimating low- to medium-frequency variance retention

Point-by-point regression

Comparisons with Pauling spring-summer precipitation reconstructions

Additional validation tests of the OWDA

References

Table S1. List of tree-ring chronologies used for producing the OWDA.

Fig. S1. Map of the OWDA domain showing 5414 half-degree grid points of JJA scPDSI (small black dots) and the 106 annual chronology tree-ring network (red and blue triangles).

Fig. S2. Maps, by decade (up to 1950), of the changing densities of precipitation stations (solid red dots) available for interpolation on the half-degree regular grid used to produce the CRU TS precipitation field (<http://badc.nerc.ac.uk>).

Fig. S3. Comparisons of calibration period (1928–1978) and validation period (1901–1927) scPDSI averages and their variances.

Fig. S4. Statistical properties of gridded summer scPDSI data over the 1928–1978 calibration period and tests of normality using a simple and robust test of normality based on joint use of skewness and kurtosis (69).

Fig. S5. CDL between 5414 grid points of summer scPDSI used for reconstruction over the OWDA domain.

Fig. S6. Summary maps of correlations between summer scPDSI and the tree-ring network over the 5414 grid points of the OWDA domain calculated for the 1928–1978 calibration period, using the PPR program in the same way that it was used to produce the OWDA reconstructions.

Fig. S7. Example of a historical/modern tree-ring chronology from northeastern France developed by the iterative procedure described in the text.

Fig. S8. Overlay plots of Tornetrask power spectra (frequencies from 0 to 0.1) for different detrending options before (RCS/SSD) and after (SF-RCS/SSD) the application of the signal-free method to the data using the same curve-fitting options: Opt 0—RCS detrending (designed to preserve the most low- to medium-frequency variance); Opt 1—negative exponential/linear detrending (monotonic nonincreasing, least flexible SSD option); Opt 2—cubic smoothing spline detrending based on the median segment length of the data (moderately data-adaptive, fixed intermediate flexibility); and Opt 3—the Friedman variable span smoother (locally adaptive, very flexible).

Fig. S9. Example of the two-stage SF-RCS method applied to the historical/modern *Quercus* species (QUSP) tree-ring data of northeastern France used as an example by Auer *et al.* (48).

Fig. S10. Additional low- to medium-frequency variance retained in the historical/modern tree-ring chronologies using the two-stage SF-RCS procedure.

Fig. S11. Calibration and validation statistical maps of the eight-member ensemble-average OWDA reconstructions.

Fig. S12. Correlations of OWDA JJA scPDSI reconstructions with Pauling spring-summer precipitation reconstructions primarily reconstructed from long instrumental and historical climate indices (18).

Fig. S13. Comparison of OWDA and Pauling maps for 1540 (“year-long unprecedented European heat and drought”) (21).

Fig. S14. Maps of exceptional droughts in Czech lands (20).

Fig. S15. Maps of the great European famine (22).

Fig. S16. OWDA mean and median maps for nine noteworthy 17th-century droughts over England and Wales: 1634, 1635, 1636, 1666, 1667, 1684, 1685, 1694, and 1695 (33).

Fig. S17. OWDA mean and median maps for eight noteworthy Ottoman Empire droughts: 1570, 1591, 1592, 1594, 1595, 1607, 1608, and 1610 (107).

Fig. S18. OWDA mean and median maps for 12 noteworthy pre-1450 historical droughts in England and Wales: 1084, 1129, 1136, 1222, 1242, 1252, 1263, 1272, 1284, 1288, 1305, and 1385 (*110*).

References (*45–110*)

“Old World Megadroughts and Pluvials During the Common Era”

Supplementary Materials

This **Supplementary Materials** contribution has the following sections:

1. *Introduction*
2. *OWDA as a scientific advance over previous work*
3. *Gridded monthly scPDSI target field*
4. *OWDA tree-ring network*
5. *Climate sensitivity of the OWDA tree-ring chronologies*
6. *Validation of the OWDA tree-ring climate response*
7. *Augmenting tree-ring chronologies with historical tree-ring data*
8. *Standardizing OWDA tree-ring data for climate reconstruction*
9. *Standardizing the OWDA historical/modern tree-ring data*
10. *Estimating low-to-medium frequency variance retention*
11. *Point-by-Point Regression*
12. *Comparisons with Pauling spring-summer precipitation reconstructions*
13. *Additional validation tests of the OWDA*
14. *References*

One table and eighteen figures accompany these sections and are referred to by them.

1. Introduction: We have developed the ‘Old World Drought Atlas’ (OWDA), a summer season history of droughts and pluvials reconstructed from tree rings spanning the past two millennia for Europe, North Africa, and the Middle East. The geographic limits of OWDA domain are shown in **Fig. SM1**. The drought metric used for reconstruction is the self-calibrating Palmer Drought Severity Index (scPDSI) (30), which has a more consistent range of relative variability in wetness and dryness across diverse climatic regions compared to the original PDSI (45). This product represents a major geographic expansion of drought atlas coverage over the Northern Hemisphere (NH) when added to the previously existing North American Drought Atlas (NADA) (8) and the Monsoon Asia Drought Atlas (MADA) (12). This Supplementary Materials (SM) section is intended to provide important additional details concerning the development of the OWDA over what is available in the main paper.

2. The OWDA as a scientific advance over previous work: Previous reconstructions of summer dryness and wetness over Europe have relied mostly on instrumental and historical climate records. The former provide nearly complete spatial coverage on a uniform grid for climatological analyses (46, 47), but consistent coverage is limited to approximately the last century. Europe also has abundant long instrumental (48, 49) and historical (13, 50) climate records that extend our understanding of drought and wetness back several centuries for selected locations. Using a reduced-space climate field reconstruction method, seasonal precipitation was reconstructed on a regular grid over Europe back to 1500 from a network of instrumental, historical, and natural climate archives, including a few tree-ring series in the Mediterranean region (18). This climate field reconstruction covers almost the entire OWDA domain and is thus the closest equivalent to the OWDA, but it does not extend back to medieval times, a period

that is especially interesting for climate change studies given the occurrence of megadroughts found in the NADA and MADA.

Another way to study past hydroclimatic variability is by its reconstruction from networks of long annual tree-ring chronologies, as performed for the NADA and MADA. In the Old World there are numerous examples of local or regional precipitation and drought reconstructions based on one or more long tree-ring chronologies (e.g., 29, 51-62), but spatial drought reconstructions of the kind provided by the OWDA are far less common (e.g., 44, 63, 64). The closest tree-ring equivalent to the OWDA is the spatial reconstruction of summer (April-September) scPDSI for the greater Mediterranean region (63), but it only covers a portion of the OWDA domain from 30°N to 50°N (cf. **Fig. SM1**) and like (18) only extends back to 1500 CE. Other spatial reconstructions of Old World drought, which do extend back to medieval times, have been made for northwestern Africa (64) and Fennoscandia (44), but they too only represent subset regions of the OWDA domain. Thus, in terms of both spatial coverage and reconstruction length, the OWDA provides a significantly more complete spatiotemporal history of summer hydroclimatic variability over the Old World for the Common Era.

3. Gridded monthly scPDSI target field: The OWDA instrumental target field used for reconstruction was extracted from a one-half degree gridded monthly scPDSI data set for global land areas (full data set available at <http://climexp.knmi.nl>), a total 5,414 grid points in this case for the OWDA domain (**Fig. SM1**). The global scPDSI data set was calculated from CRU-TS monthly temperature and precipitation data (36), updated to version CRU-TS 3.21 and covering the period 1901-2012 (<http://badc.nerc.ac.uk>). For calculating monthly scPDSI, the more physically realistic Penman-Monteith method of estimating potential evapotranspiration (PET) was used (along with a modification of the scPDSI algorithm to include a simple snowmelt hydrology model for winter snowpack dominated regions) (11), but exact details how PET is estimated has little impact on 20th century PDSI estimates (65) or on reconstructions of past PDSI from tree rings (35). The monthly scPDSI data extracted for the OWDA were seasonalized to produce summer season (June-July-August or JJA) average values for tree-ring based reconstruction at each grid point. This is the same season as reconstructed for the NADA and MADA, which is generally consistent with the growing season of NH trees.

The CRU TS 3.21 precipitation data used in the scPDSI calculations is considered to be a significant improvement over an earlier version of CRU precipitation data (66). However, it is still useful to investigate the density of precipitation stations used for interpolation for potential problems related to station coverage and its reduction back in time. **Fig. SM2** shows maps by decade (up to 1950) of the changing density of precipitation stations available for interpolation onto the CRU TS 3.21 precipitation field. Overall, station density is high and stable back in time for most areas of the OWDA domain. Only in Turkey and the Middle East is there an obvious loss of local precipitation data for interpolation prior to 1930. This could result in some loss of variance in the interpolated values in that region prior to 1930 due a “relaxation to climatology” (loss of variance towards the climatological mean) that is part of the CRU TS gridding methodology in large areas devoid of station data (67). The Shoemaker F-test for equal variances (68) was used to test for this possible effect in the summer scPDSI data by comparing the variances of the 1928-1978 calibration and 1901-1927 validation periods used for reconstruction and testing of the OWDA (see the *11. Point-by-Point Regression* section). The results of this test revealed statistically significant ($p < 0.01$) reductions in variance in the scPDSI data prior to 1928 over eastern Turkey, the Middle East (mostly Syria and Iraq), and fringe areas of the Sahara

Desert (**Fig. SM3**). This suggests a “relaxation to climatology” in those areas that could affect the reliability of validation tests of the OWDA reconstruction there.

A simple but powerful omnibus test of normality based on the joint use of skewness and kurtosis, the D’Agostino-Pearson K^2 statistic (69), was applied next to the summer scPDSI grid over the calibration period (1928-1978) when non-normality would have its biggest impact on the regression-based estimation of the statistical transfer functions used to reconstruct the OWDA. The maps of skewness (Skew) and kurtosis (Kurt-3) shown in **Fig. SM4** indicate a small tendency for the distribution of summer scPDSI to be positively skewed and with a fatter peak and thinner (platykurtic) tails than expected for normally distributed data. However, these data properties did not translate into statistically significant ($p < 0.05$) departures from normality over most of the OWDA domain as indicated by the D’Agostino-Pearson K^2 statistic (K2-test) and its probability (Prob). The summer scPDSI data used for reconstruction are for practical purposes normally distributed.

A final examination of summer scPDSI was conducted to determine the characteristic correlation decay length (CDL) defined by the e -folding level of correlation ($1/e$) between grid points. The pairwise correlations were binned and averaged by 100 km distance increments for each one-half degree latitude band. Doing so reduced the geographic bias in the number of pairwise correlations calculated as a function of latitude and diminished land area above $\sim 55^\circ\text{N}$ (**Fig. SM1**). The binned average correlations were also color coded by latitude to highlight how much latitudinal dependence there is in CDL. These results are shown in **Fig. SM5**, along with a 40% robust LOESS smooth fitted to the binned data to estimate the point where average correlation declines to $1/e$. The LOESS smooth indicates a CDL of 800 km with considerable latitudinal dependence. Lower latitudes typically have shorter CDLs compared to higher latitudes. This estimate of CDL is significantly longer than the 450 km value reported for monthly precipitation (70, 71). However, scPDSI also integrates temperature influences through PET, which may increase the CDL by some amount. But the biggest difference is probably the way in which CDL was calculated here from the binned data to reduce latitudinal bias. The 800 km value is also located around the north-south latitude center of the OWDA domain based on the color-coded latitude bins of average correlation (cf. **Figs. SM1** and **SM5**), which argues for it being a spatially reasonable estimate.

4. OWDA tree-ring network: The greatest challenge in creating the OWDA was the development of a tree-ring network that would span the past millennium or more over most of the domain. At the start of this project in 2011, publically available tree-ring data from the International Tree-Ring Data Bank (ITRDB) beginning before 1700 totaled 171 sites, but only 11 began before 1300. This was inadequate for reconstructing Old World hydroclimate back to medieval times or earlier. In order to reach this goal, tree-ring scientists working in various parts of the Old World kindly contributed many previously unavailable tree-ring collections to the OWDA project. Several of these data sets were based on the ring widths of living trees crossdated exactly with timbers sampled from historical (here meant to also mean “archeological”) structures (72, 73). The latter allowed many of the tree chronologies in the OWDA network to be extended back to the beginning of the Common Era. Other millennia-long chronologies, based on living trees and naturally preserved subfossil wood (74, 75), were compiled as well. See (75) for the use of both kinds of chronologies in the study of European climate variability over the Common Era and its impact on Old World cultural history.

Fig. SM1 shows the network of 106 annual tree-ring chronologies used, color coded by series length. See **Table SM1** for a more detailed list of these chronologies. It is a multi-species network comprised of oak (*Quercus robur*, *Q. petraea*), pine (*Pinus sylvestris*, *P. nigra*, *P. brutia*, *P. heldreichii*, *P. uncinata*, *P. cembra*, *P. laricio*, *P. leucodermis*), fir (*Abies alba*, *A. nordmanniana*), spruce (*Picea abies*), larch (*Larix decidua*), juniper (*Juniperus excelsa*, *J. phoenicea*), cedar (*Cedrus atlantica*, *C. libani*), beech (*Fagus sylvatica*), and yew (*Taxus bacata*). The benefit of using a multi-species tree-ring network for reconstructing climate in Europe has been documented in Europe (76) and elsewhere (77). Different tree species responding to the same basic growth-limiting factor (e.g., soil moisture deficit) may yield noisy or biased estimates of the modeled limiting factors in their ring widths related to phylogenetic (e.g., deciduous vs. evergreen) controls (43). Furthermore, site-level (e.g., site hydrology and soil moisture holding capacity) differences can also result in variation in the climate response. Such noise/biases can be reduced in climate reconstructions when multiple tree species and sites are used (77). They can also be corrected for to some extent by the use of time series modeling techniques that correct for differences in relative “redness” (i.e., positive autocorrelation) between tree rings and the climate variable being reconstructed (41,78), an aspect of climate reconstruction from tree rings not considered in a paper on spectral biases in tree-ring reconstructions of past climate (79).

5. Climate sensitivity of the OWDA tree-ring chronologies: In terms of the likely drought sensitivity of the OWDA tree-ring chronologies, even in the more mesic areas north of the Alps, recent published results indicate that summer moisture availability is *positively* correlated with radial growth in tree-ring chronologies distributed over most of the OWDA domain (43,80). These correlations between extensive multi-species data sets, some of which are also used in the OWDA, were found to be strongest for summer scPDSI. Similar results have been reported for oak, beech, and pine growing in central-west Germany (81). Some of those data are used here as well. These findings should not be surprising given the successful reconstructions of precipitation and drought described here over the Old World (29, 44, 51-64). In both (43) and (81), a significant carryover of climate information from the previous-year summer and fall in the tree rings was also noted. See also (82) for this carryover effect on extreme events recorded in the tree rings. It is standard practice to use this previous-year information to improve the point-wise reconstructions in drought atlases (41) and that was done here for the OWDA as well.

Although soil moisture availability during the growing season is likely to be the most common growth limiting factor in the OWDA tree-ring network, trees growing in cool, moist forested environments found at high elevations and high northern latitudes of the Old World can also be limited by growing season temperatures. In such cases, the correlations between tree rings and temperature can be strongly positive (43,80). This would appear to make it impossible to use such tree-ring series for drought reconstruction. However, it is still possible to use tree-ring chronologies having positive correlations with growing season temperature to indirectly reconstruct soil moisture availability through its *inverse* association with clear skies, warmer temperatures, and reduced precipitation. This results in tree growth being *negatively* correlated with scPDSI, a phenomenon most strongly found in the European Alps and the northern Fennoscandian and Russian portions of the OWDA domain. This negative correlation with scPDSI has been used here in chronologies based on both ring width and maximum latewood density (usually a stronger measure of summer warmth than ring width; 83) in an effort to provide reconstructions comparable in length to those produced farther south using tree-ring chronologies having positive correlations with scPDSI. The spatial reconstruction of

Fennoscandian drought described earlier (44) successfully used such tree-ring chronologies with negative correlations for the successful reconstruction of drought there.

6. Validation of the OWDA tree-ring climate response: Having reviewed what is known about the climate sensitivity and lagged responses of trees growing in the Old World, it is useful now to document these properties in the OWDA tree-ring network itself when correlated point-by-point with the summer scPDSI target field. This was done using the Point-by-Point Regression (PPR) method (41) identically to how the OWDA reconstructions themselves were produced. See the *Point-by-Point Regression* section for more details. Here we only show summary maps of correlations between summer scPDSI and the tree rings used for reconstruction at each grid point, calculated for the 1928-1978 calibration period and based on autoregressively modeled and prewhitened data (41,78). The correlations shown are both contemporaneous (year t scPDSI vs. year t tree rings) and lagged (year t scPDSI vs. year $t+1$ tree rings) to take into account the carryover effect of climate on tree growth. Results are condensed into a series of maps that show the spatial distributions of the median (Med), maximum negative (MaxN), and maximum positive (MaxP) correlations for year t and year $t+1$. These summaries are necessary condensations due to the the large number of correlations between scPDSI and tree rings, namely, 216,560 correlations over the 5,414 point grid, which resulted from a minimum of 40 year t and year $t+1$ correlations per grid point (as 20 or more chronologies were required to contribute to each grid point). The MaxN and MaxP correlations were also tested for significance and only those that passed the 2-tailed 95% level are plotted. **Fig. SM6** shows these maps.

The median correlation maps (**Fig. SM6** center) for year t (Med Rt) and $t+1$ (Med Rt+1) show the clear tendency for the OWDA tree rings to be positively correlated with summer scPDSI (more orange than blue), which is consistent with previous findings (43, 81). There are 44% (48%) more positive correlations for year t (year $t+1$) compared to negative correlations, not taking into account the statistical significance of the point-wise correlations. However, if only those correlations significant at the 2-tailed 95% level are considered, the difference in the frequency of positive vs. negative correlations is much greater: 277% (243%) more year t (year $t+1$) positive correlations compared to negative correlations. So there is little doubt that the dominant signal in the OWDA tree-ring network is related to growing season moisture availability and its carryover effect the following year. There are, however, some interesting geographic patterns associated with these results. As indicated earlier (43, 80), trees growing at high elevations and high latitudes tend to be more temperature limited and this appears to be the case in northern Fennoscandia, northern Russia, and the Adriatic region of the Mediterranean in year t . However, the $t+1$ correlations with the scPDSI in those same regions tend to be positive, thus largely compensating for the negative year t correlations.

The maximum negative (MaxN Rt, Rt+1) and maximum positive (MaxP Rt, Rt+1) correlation maps (**Fig. SM6**) show again the predominant influence of moisture stress on tree growth in the OWDA tree-ring network. The spatial patterns of significant MaxP correlation cover more of the OWDA domain than do the patterns of significant MaxN correlation. In addition, the MaxP patterns appear to be more stable from year t to year $t+1$ compared to the MaxN patterns.

7. Augmenting tree-ring chronologies with historical tree-ring data: Developing relatively homogeneous tree-ring chronologies for the OWDA network from a combination of modern and historical tree-ring data, extending back in some cases over 2,000 years, was particularly

challenging. We do not know where the trees were felled for construction in any given region, but the existence of a common signal between the modern and historical tree-ring samples implies that the historical data must have come from trees growing under reasonably similar site and growth limiting conditions in a reasonably similar climatic region. That being the case, we would like the historical tree-ring data to have statistical properties (e.g., measures of common signal strength; 84) consistent with the modern tree-ring data used for calibration with scPDSI to provide some degree of inferred uniformitarian stability in the reconstructions back in time. This is a somewhat more stringent requirement than that used by dendroarcheologists for dating historical structures. For this reason, we re-evaluated the regional historical/modern tree-ring data sets received from dendroarcheologists to seek the best subset of historical data having a common signal strength equivalent to that in the modern data. Any systematic drift away from the modern signal strength benchmark might reflect an undesirable degradation of the chronology climate signal back in time.

To this end, the historical tree-ring data provided to this project were carefully screened to help insure that they formed a reasonably homogeneous dataset. This was done by first finding the most recent historical data subset that overlapped and correlated most strongly with the inner portion of the modern tree-ring data set. The quality control computer program COFECHA (85) was used for this purpose. Those historical series identified by COFECHA to correlate at a level consistent with the modern data were added to the modern ensemble to produce an initial historical/modern tree-ring data set. Doing so both increased the chronology sample size per year (replication) in the overlap period and extended the modern chronology length back in time. This augmented historical/modern chronology was then compared to the remaining unused historical tree-ring data in COFECHA and the next subset of historical data that correlated most strongly with the previous combined data set were identified. This newly identified subset of historical data was then added to the initial historical/modern tree-ring data set to further replicate and extend the chronology back in time. This process was repeated until no additional historical data could be added to the historical/modern tree-ring data set with a level of coherence consistent with the previous additions. When possible the mixed historical/modern tree-ring data were also updated using “random” collections of ring-width data from the same tree species obtained from sawmills and other tree-ring sites to mimic the vague source locations of the historical tree-ring data from past constructions (37). This method of updating assumes that the random collections are from the same general tree growth class (e.g., dominant) as that used in past building constructions. This assumption is unlikely to be true in a strict sense. Thus, there is likely to be some level of growth rate bias in the updated ring widths from random collections (cf. 86). Even so, random collections as described in (37) will still reduce the potential for site-level bias in the updates of the historical/modern tree-ring chronologies, which is a good thing.

The procedure adopted here to develop the historical/modern tree-ring chronologies for climate reconstruction was labor intensive because, in some cases, several thousand tree-ring series from a single region had to be evaluated for inclusion in the final tree-ring chronology used for producing the OWDA, with a significant subset of data sometimes not being used. It must be clearly stated here that the *exclusion* of any historical tree-ring data from the final tree-ring chronologies used in the OWDA *does not* imply that those excluded series were incorrectly dated by the contributing dendroarcheologist. Rather, they simply did not have the same common signal strength characteristics as that found in the modern data as ascertained by COFECHA. This is a generally more stringent chronology development criterion compared to

that necessary to date historical structures (87, 88). Even so, there is no guarantee that the screened historical/modern data have not drifted in some way (e.g., site location, ecology, climatic signal) from the modern/historical overlap standard in earlier centuries because most of the historical data used do not directly overlap in time with the modern data. This is an issue that cannot be easily investigated given the data and metadata in hand, but the use of many tree-ring chronologies in each grid point reconstruction of the OWDA (see *11. Point-by-Point Regression*) should offer some protection from this possible problem.

Fig. SM7 shows an example of a historical/modern oak (*Quercus* sp.) tree-ring chronology from northeastern France (37) developed following the procedures just described. The methods of data processing used here are the same as those used in COFECHA to evaluate the high-frequency strength of crossdating between series: 32 year spline detrending followed by autoregressive modeling to produce ‘white noise’ residuals for the testing of crossdating.

The first plot (**Fig. SM7A**) shows the mean ring-width chronology of best crossdated series produced by the iterative crossdating procedure. There are 1,844 series in this chronology, which covers the period 511 BCE to 2011 CE, but it displays erratic behavior over time due to changing sample sizes and sources of wood and, therefore, cannot be used for climate reconstruction as is.

The second plot (**Fig. SM7B**) shows the mean chronology of the COFECHA processed and crossdated tree-ring data. This chronology is much more homogeneous over time compared to the one based on the ring-width measurements (**Fig. SM7A**), but only high-frequency variance used for determining crossdating has been retained without consideration for the preservation of low-to-medium frequency (LMF) variance (timescales of decades to centuries in duration) that might be attributed to climate. As such, this chronology cannot be used for climate reconstruction either.

The third plot (**Fig. SM7C**) shows the large changes in sample size per year (replication) that is a common feature of dendroarcheological chronologies. The large changes in replication reflect the number of buildings that were constructed during different time periods and the variable survival rate of those structures over time. Note the “pinch point” in replication in the 7th century, similar to the reduction of the number of tree fellings recorded in an oak tree-ring collection from central Europe during the social turmoil associated with the post-Roman ‘Migration Period’ of the 4th-5th centuries (75).

The fourth plot (**Fig. SM7D**) shows the strength of crossdating as defined by the average correlation between series (RBAR; 84) calculated for 51 year time periods with 25 year overlaps (running RBAR) (89). The most important feature of the running RBAR plot is its stability over time from the modern end based on only living trees back to the beginning of the chronology based on only historical tree-ring data. The obvious anomaly in this plot is the narrow spike in high RBAR in the 7th century when replication is conspicuously low. This spike in average correlation is very likely caused by multiple measurements taken from timbers from the same tree. More independent samples from the 7th century are needed to fix this weakness, but it is limited to a very narrow time period relative to the rest of the chronology and, therefore, should not be considered a serious problem, especially when viewed in the context of multiple chronologies being used for reconstruction at each grid point.

The fifth plot (**Fig. SM7E**) shows the Expressed Population Signal (EPS; 84) estimated from the running RBARs and changing sample sizes over time. EPS provides an estimate of how closely the finite sample size chronology matches the theoretical population chronology based conceptually on an infinite sample size. An EPS of 1.0 is a perfect match between the sample

chronology and its theoretical population chronology, which can never happen in practice because of sampling variability and measurement error. The red line representing an EPS of 0.85 is a 'rule of thumb' test for determining the useful length of the sample chronology. This is a desirable target, but "... in any particular case the chosen threshold will depend on the user's subjective evaluation of accuracy needs" (84). A corollary to this statement is that even if the 0.85 threshold is used, it is not absolutely necessary that EPS always stays above 0.85 over the usable length of the chronology. Minor transient dips of EPS below 0.85 may be considered acceptable in order to use as much of the chronology length as possible. The chronology from northeastern France has a running EPS that exceeds 0.85 for all but a few brief periods. Except for the anomalous spike in RBAR in the 7th century there is little reason to question its crossdating stability and uniform signal strength over the Common Era.

This example shows how the strength of crossdating in the historical/modern tree-ring chronologies can be maintained back in time in a reasonably homogeneous way through the iterative crossdating procedure described here. All of the historical/modern data sets used in the OWDA tree-ring network were developed this way and have comparable levels of crossdating stability. However, this procedure does not provide us with a chronology that is useful for climate reconstruction because all of the LMF variance in the tree-ring series, some undoubtedly due to climate, has been removed by the iterative crossdating procedure. In the next section, we will describe how the tree-ring chronologies used for OWDA reconstruction were developed, including a novel two-stage method of tree-ring detrending for processing the mixed historical/modern data sets.

8. Standardizing OWDA tree-ring data for climate reconstruction: Standardization of tree-ring measurement series is a fundamental step in preparing those data for climate reconstruction (90, 91). It involves the removal of biological (tree age/size dependent) growth trends considered to be unrelated to climate and the transformation of the detrended series into a set of dimensionless indices that are suitable for averaging into the final tree-ring chronology used for climate reconstruction. But the act of detrending itself can also inadvertently remove a significant amount of climate related LMF variance from multi-centennial to millennial length tree-ring series. Generically, much of this problem relates to what is called the "segment length curse" (SLC; 40), which basically states that it is impossible to retain time scales of variability in the data longer than the tree-ring series being detrended. It makes no difference how stiff the fitted growth curve used for detrending is. Unsurprisingly, the loss of LMF variance after detrending can become much more severe as the flexibility of the fitted detrending curves increases, e.g., through the use of cubic smoothing splines (92) or Friedman variable span smoothers (93). Regional curve standardization (RCS) was introduced (94) as a way to effectively break the SLC by preserving LMF variance in excess of the lengths of the series being detrended. The fundamental rationale and mechanics of RCS are well described in (94,95) and will not be repeated here, but this technique does have its limitations. RCS is best suited for standardizing large homogeneous remnant/modern wood such as those from the Austrian Alps (74) and northern Fennoscandia (94). However, RCS can be very difficult to apply to tree-ring series from only modern (living) trees because of the presence of "differing contemporaneous growth-rate bias" and "modern sample bias" between trees (95), which can impart spurious trends to the resulting RCS chronology. See also (86) for other examples of potential bias in forest growth rate studies.

For tree-ring data sets based exclusively on living trees, the SLC can be extremely difficult to break with any confidence because RCS can rarely be applied. The process of detrending individual series will always remove some long-term climate driven variability on time-scales longer than the series' lengths. However, the negative impact of this problem can be ameliorated somewhat through use of the signal-free (SF) method of tree-ring standardization (38), which has the capacity to preserve resolvable LMF variance that is *common* to all series in the tree-ring data ensemble. The SF method used in combination with single-series detrending (SSD) is referred to here as SF-SSD. SF-SSD was developed to correct for trend distortion artifacts at the ends of series that might be caused by the fitted growth curves tracking a common signal due to climate, an obvious example being a common positive trend in tree growth in many upper elevational and latitudinal treeline locations over the 20th century associated with climate warming. See (38) for examples of the trend distortion problem and how it can be substantially ameliorated by use of SF-SSD. Perhaps less obvious is the way that SF-SSD can recover *common* LMF variance throughout the length of the final standardized tree-ring chronology even in the case where very flexible detrending curves have initially removed it. This property is illustrated here for Lake Tornetrask *Pinus sylvestris* data from northern Sweden (94). While the data used in (94) began in 500 CE, we use here an augmented data that extends back to the beginning of the Common Era. The RCS method has also been extended now to include the signal free method (SF-RCS; 39,96), so we will also compare LMF variance retention using SF-RCS.

SF-RCS based on a single RCS growth curve was applied to the Tornetrask data to provide a “best case” example of the preservation of LMF variance in the chronology (Opt 0). For comparison, we detrended the same data using three different methods of ‘single-series’ detrending with large differences in flexibility: Opt 1 - negative exponential/linear detrending (monotonic non-increasing, least flexible), Opt 2 - cubic smoothing spline detrending based on the median segment length of the data (moderately data adaptive, fixed intermediate flexibility), and Opt 3 - the Friedman variable span smoother (locally adaptive, very flexible). The loss of LMF variance due to the SLC and increasing flexibility of the detrending methods used increases dramatically from Opt 0 to Opt 3. **Fig. SM8** illustrates this in the power spectra of the Tornetrask chronologies before (STD) and after (SF-SSD) the signal-free method was applied to the data using the same curve fitting options. The increasing flexibility of the curve fit options from Opt 1 to Opt 3 has resulted, as expected, in progressively larger losses of LMF variance in the STD chronologies. After the application of SF-SSD, most of the lost medium frequency (decadal to century) variance caused by the use of increasingly flexible detrending curves from Opt 1 to Opt 3 has been recovered. This is a remarkable property of SF-SSD. Thus, SF-SSD can provide a degree of protection from the excessive loss of *common* medium frequency variance caused by flexible detrending methods, but it cannot replace SF-RCS for preserving longer time scale variance in excess of the series lengths being detrended.

9. Standardizing the OWDA historical/modern tree-ring data: All of the OWDA tree-ring chronologies based only on living trees have been detrended using SF-SSD. Most of these series were detrended with age-dependent smoothing splines (97), which can mimic the behavior of monotonic negative exponential curves, but in a somewhat more adaptive way. However, as **Fig. SM8** shows, the detrending method used should have relatively little impact on the preservation of medium frequency variance as long as SF method is used. The creation of the historical/modern tree-ring chronologies used for scPDSI reconstruction in the OWDA was more

challenging. To illustrate how this was done for each of these more complicated data sets in the OWDA tree-ring network, we continue with the historical/modern oak ring-width data from northeastern France as our example data set.

Given the historical/modern tree-ring data set developed for northeastern France (**Fig. SM7B**), we wish to apply SF-RCS (39, 96) to the data in order to preserve as much LMF variance as possible in the chronology due to climate. This is not easy to do with the screened historical/modern tree-ring data sets. Each of these collections was screened using strength of RBAR as the criterion for acceptance, but here only based upon the high-frequency variations in ring width related to year-to-year influences of climate on growth. Now we wish to preserve centennial or longer variations in growth in the historical/modern chronologies. This is much easier said than done because it places much greater demands on the data relating to differences in absolute growth rates, uncertain provenances of the wood used for construction, and uncertain criteria used to sample and aggregate living and historical datasets (86, 95). Massive sample replication involving thousands of overlapping series distributed evenly through time can mitigate these problems (98) through the averaging process and the Law of Large Numbers, but this cannot always be depended upon. There is also ‘modern sample bias’ (95) to deal with, which is not necessarily minimized by replication and the averaging process alone because of the lack of random overlaps of tree-ring series at the modern end of the chronology. Potential bias in updating tree-ring chronologies using RCS (99) must also be kept in mind. In addition, sample replication can be highly variable over time (e.g., **Fig. SM7C**) as it depends on the frequency of construction and survival of historical structures available for sampling. During times of cultural crisis and upheaval, like during the ‘Migration Period’ in early post-Roman times (75), replication can plummet.

With these issues in mind, a 2-stage SF-RCS procedure was devised for this project to correct for inhomogeneities in the historical/modern RCS chronologies not removed by single-curve RCS detrending (94). This 2-stage procedure differs from that based on multiple-RCS curves (39, 96) to correct for differential growth rate effects and is not intended to be a replacement for that method. However, as will be shown, the 2-stage procedure can correct for the most egregious homogeneity problems caused by differential growth rate effects between the modern and historical data, and also due to the highly variable sources of wood in the historical portion of the data set, while at the same time preserving centennial time scale variability thought to be driven by climate.

We use again the historical/modern oak tree-ring data set from northeastern France described in 7. *Augmenting tree-ring chronologies with historical tree-ring data* and shown in **Fig. SM7**. **Fig. SM9A-B** shows the results of applying single-curve SF-RCS to the data. This is ‘Stage 1’ of the 2-stage SF-RCS procedure. A single mean biological growth curve (the RCS curve) has been estimated without any pith-offset adjustments (39, 96) because none were available. This RCS curve is shown in **Fig. SM9A** before and after two iterations of signal-free detrending were applied to the data. There is very little difference between the before-and-after RCS biological growth curves (**Fig. SM9A**) and the resulting chronologies (**Fig. SM9B**). Note the enormous similarity between the RCS chronologies and the simple mean of the raw data (**Fig. SM7A**) as well. In this case, the estimated RCS curves have not corrected the inhomogeneities in the raw data at LMF time scales, the most obvious period centered around 600 CE. This is primarily due to systematic (non-random in time) growth-rate differences that a single RCS curve is unable to account for, which is reflected in the means of the tree-ring indices of the individual segments after detrending by the single RCS curve. Of course, some of these ‘non-

random in time' growth rate differences may be due to climate, but the properties of the historical/modern data sets make it likely that these differences are mostly non-climatic in origin.

'Stage 2' of the 2-stage SF-RCS procedure is designed to reduce differential growth rate biases by removing the means of the individual RCS-detrended segments using an iterated signal-free method (**Fig. SM9C**). In so doing, the most obvious inhomogeneities are removed as indicated in the 'before' (blue) and 'after' (red) SF-RCS chronologies (**Fig. SM9D**). The differences between the 'before' and 'after' chronologies are very localized and most apparent in the poorly replicated 600 CE period and in the post-1800 CE period when modern sample bias is likely to be a significant problem. The SF-iterated effects of the Stage-2 procedure are most easily seen in the mean chronology smoothed curves (**Fig. SM9E**) from the beginning Stage-1 input (blue) to final Stage-2 SF-iterated output (red). The loss of spectral power in the final SF-RCS chronology from Stage-1 to Stage-2 is apparent, especially at periods >100 years. It is not clear how much of this lost LMF variance is due to climate, but it is almost certain that much of the added LMF variance in the Stage-1 chronology is due to non-climatic effects.

In the next section, the issue of LMF variance retention using a 2-stage SF-RCS procedure will be investigated. In particular, we seek to determine the degree to which we are doing better than what might be expected using the SF-SSD method based on very conservative detrending.

10. Estimating low-to-medium frequency variance retention: Retention of LMF variance in the overall OWDA tree-ring network after the use of signal-free methods is difficult to quantify precisely. **Table SM1** provides the median segment lengths (MSL) for the OWDA tree-ring chronologies. They range from 69 to 475 years with a median across all sites of 187 years. This finding suggests that we should be able to reconstruct scPDSI variability at time scales up to ~150 years over most of the OWDA domain using signal-free methods. In fact, it is highly likely that we are doing much better than that. Out of 106 chronologies, 63 are based on mostly modern (living) trees with a median sample size of 101 series per chronology and a MSL of 240 years. After standardization using SF-SSD (SF-RCS was not considered appropriate here), the maximum LMF variance retained in these chronologies should have a median value of ~200 years based on the results shown in **Fig. SM8**. With a MSL of 595 years (**Table SM1**), the retained LMF variance is close to the practical (~1/3N) resolvability limit of low-frequency signals in classical spectral analysis.

The remaining 43 chronologies, based on a combination of modern and historical or remnant wood samples, have a median sample size of 678 series and a MSL of 119 years. If these chronologies were standardized using SF-SSD, the maximum retained LMF variance would have a median value of ~100 years. Given that the median length of these chronologies is 1205 years (**Table SM1**), this level of retained LMF variance is unacceptably low. Consequently, the chronologies made up of historical/modern tree-ring data were standardized using the 2-stage SF-RCS procedure described above. This was practical to do because the overall sample size per OWDA data set (**Table SM1**) greatly exceeds the minimum of 50 that has been recommended for calculating a single RCS curve (39). However, the sample size in any given year can vary considerably (e.g., **Fig. SM7C**) and this will affect the signal strength of chronologies over time. But because OWDA chronologies are not used in isolation for reconstructing the scPDSI at any given grid point (see *Point-by-Point Regression*), the changing sample size and associated signal strength issue should not seriously affect the reconstructions except perhaps in cases where the earliest portions of the nested grid point reconstructions are based on only one chronology.

The recovery of additional LMF variance using 2-stage SF-RCS is illustrated again for the northeastern France tree-ring data set, but is expressed in a different way than in **Fig. SM9F**. **Fig. SM10A** shows the ratio of the chronology power spectra (SF-RCS/SF-SSD) plotted for periods from DC (infinite period) to 25 years. The largest effect of SF-RCS is indicated (red hatching) for periods longer than the median segment length of 108 years. The SF-RCS chronology has 5.52 times as much total spectral power over periods >100 years compared to the SF-SSD chronology based on age-dependent spline detrending. This we call the ‘LMF-Ratio’. An LMF-Ratio>1 means that 2-stage SF-RCS is preserving more LMF variance than SF-SSD for the given single-series detrending option used.

A summary plot of the LMF-Ratio for all 43 historical/modern chronologies as a function of MSL is shown in **Fig. SM10B**. There is a curvilinear relationship that is reasonably well approximated by a power function ($r=0.47$, $p<0.01$), which implies a log-log relationship between MSL and LMF-Ratio. This plot indicates that historical/modern chronologies benefit most from 2-stage SF-RCS when MSL is less than 150 years. The mean (median) LMF-Ratio for the 33 chronologies that satisfy this limit (76.7% of the total) is 4.8 (3.7). From these results it is evident that the 2-stage SF-RCS method generally preserves more LMF variance than the SF-SSD method and more than should be expected given the SLC alone (cf. **Fig. SM8**). It is unclear how different the results might be using the multiple-RCS method (39), but the much simpler 2-stage SF-RCS method appears to work acceptably well in preserving LMF variance in excess of SLC in the large majority of historical/modern tree-ring chronologies and also in correcting for obvious growth rate biases in those data sets (see 9. *Standardizing the OWDA historical/modern tree-ring data*).

11. Point-by-Point Regression: The Point-by-Point Regression (PPR) method was used to reconstruct the OWDA sequentially at each of its 5414 grid points from the network of 106 tree-ring chronologies (**Fig. SM1**). PPR is based on a principal components regression (PCR) procedure, and its basic mechanics and usefulness as a climate field reconstruction method are described in detail in (41). Each grid point PCR model used to produce the OWDA was estimated uniquely from an overlapping subset of tree-ring chronologies that fell within a given distance (search radius) from the grid point in question. A 1000 km search radius was used here (reference circle in **Fig. SM1**), which is a small increase over the 800 km average e-folding distance of the instrumental gridded data (**Fig. SM5**) and was considered necessary due to the irregular nature of the tree-ring network and its sparse coverage in the eastern portion of the domain.

A minimum of 20 tree-ring chronologies was required to produce each grid point reconstruction. When this minimum number could not be found using the initial 1000 km search radius, it was increased by 50 km increments until at ≥ 20 chronologies were found (cf. 41). Over the OWDA domain, the number of grid points that did not require a search radius enlargement to find at least 20 chronologies was 1725 out of 5414 or 31.8%, and the geographic centroid of these grid points falls almost exactly at the center of the plotted circle in **Fig. SM1**. However, the overall geographic (North-South-East-West) coverage of this region is much larger than the example search radius implies, as indicated by the long-dash rectangle. Beyond the limits of this rectangle all grid points required a dynamically expanded search radius to find at least 20 tree-ring chronologies. The added domain coverage of an expanded search radius of 1500 km (3755 grid points or 69.4% of the total) is shown by the short-dash rectangle. Most of the OWDA domain is now covered, and 1500 km still falls well within the north-south range of variability of

the correlation decay length of the scPDSI field itself (**Fig. SM5**). It also falls within the range of common tree-ring signature years (conspicuous wide or narrow rings) found between tree-ring sites across Europe (100, 101), which are likely due to climate teleconnections associated with large-scale atmospheric circulation patterns.

At this stage a calibration period was chosen to estimate each PCR model for reconstruction of scPDSI. Principal components analysis requires a common time period for estimating the covariance matrix and this was dictated by the earliest last year of the tree-ring chronologies used. In **Table SM1** that earliest last year is 1979. Taking into account a one-year lag response of tree growth to climate as described earlier (43, 81) (see also **Fig. SM6**), this means the last year of the calibration period in common for all chronologies is 1978. Coincidentally, that is the same last year used for calibration in producing the NADA (8), so the same calibration period was chosen for the OWDA as for the NADA: 1928-1978, a total of 51 years. The scPDSI available prior to 1928 was withheld from the calibration exercise to test the tree-ring estimates of scPDSI for skill. Thus, the validation period used here is 1901-1927, a total of 27 years.

As described in (41), and applied to the development of the other drought atlases (8, 12), the tree-ring and instrumental drought data used for calibration were prewhitened using autoregressive (AR) models fit to the data to correct for differences in autocorrelation between them (41, 78). The AR models were estimated from the calibration period data and applied back in time over the lengths of the tree-ring and scPDSI data not used for calibration. Doing so produces approximate ‘white noise’ reconstructions that can then be ‘reddened’ by adding the AR persistence in the instrumental data to them over the full lengths of the reconstructions.

Since (41), PPR has been modified to produce nested grid point reconstructions (102, 103) to provide the longest records of past climate possible given the lengths of the tree-ring series used as predictors at each grid point. The OWDA reconstructions also had the lost variance due to regression over the calibration period added back to the estimates. Doing so results in reconstructions that are extremely similar to those produced by the ‘composite-plus-scale’ (CPS) method (35) and also allow for the reconstructions to be seamlessly updated from 1979 to 2012 with instrumental scPDSI data. The PPR method has also been modified to allow for the differential weighting of each tree-ring predictor by some power of its correlation with the predictand variable (12, 42), in this case scPDSI. In so doing, an 8-member OWDA ensemble was produced. The ensemble members were then averaged and recalibrated and revalidated directly against the instrumental data and the average correlation between ensemble members at each grid point was calculated.

As described, PPR was applied to the reconstruction of gridded summer scPDSI over the 5414 point OWDA domain. The calibration period statistics, prefaced with a “C”, reported here are the coefficient of determination or R^2 ($CRSQ$) and cross-validation reduction of error ($CVRE$). The latter is a “leave-one-out” procedure analogous to R^2 based on Allen’s $PRESS$ statistic (104) and is a more conservative measure of explained variance than $CRSQ$. The validation period statistics reported here (prefaced with a “V”) are the Pearson correlation coefficient squared ($VRSQ$), the reduction of error statistic (VRE), and the coefficient of efficiency (VCE), all of which are measures of explained variance between actual and estimated values when positive. See (41, 42) for details. Note that these validation statistics (sign considered for $VRSQ$) can go negative as well, which indicates some lack of skill in the reconstructed values. As a rule of thumb (because no theoretical significance tests are available for VRE or VCE), if $VRE > 0$ or $VCE > 0$ then the reconstruction has some skill in excess of the calibration or verification period

means, respectively, of the instrumental data (41). In the case of VRSQ, values significant at the 95% level based on positive correlations only ($p < 0.05$, 1-tailed) were used as indications of significant reconstruction skill. The final statistic reported is the RBAR (84) between the eight ensemble members at each grid point. This statistic is a measure of common signal strength between the ensemble members.

Fig. SM11 shows maps of the calibration and validation statistics for the 8-member ensemble-average OWDA reconstructions. All statistics are in units of fractional variance. Only those validation statistics that passed (VRSQ with $p < 0.10$ 1-tailed, VRE and VCE > 0) are plotted. The number of grid points plotted (max=5414) and their median values as follows: CRSQ, N=5414, Median=0.328; CVRE, N=5414, Median=0.271; VRSQ, N=3057, Median=0.198; VRE, N=2495, Median=0.161; VCE, N=1529, Median=0.146; RBAR, N=5414, Median=0.820). The region of poorest validation in general is in Russia, which is also the region with the poorest local tree-ring coverage.

The calibration and validation period statistics are somewhat modest, but very likely underestimate the true skill and usefulness of the OWDA reconstructions back in time. This was illustrated in the main paper for six exceptional hydroclimatic events known from historical documents or instrumental records not used for calibration. Here we provide much more evidence for reconstruction skill in the OWDA, first by the direct quantitative comparison of the OWDA JJA scPDSI reconstructions with spring-summer (SPSU) precipitation reconstructed from long instrumental records, historical climate indices, and natural climate archives (18). We then qualitatively compare our reconstructions to reports of unusual hydroclimatic variability from historical climate records that are completely independent of the OWDA estimates and extend back to medieval times.

12. Comparisons with Pauling spring-summer precipitation reconstructions: Gridded reconstructions of seasonal precipitation have been produced for Europe back to 1500 from a network of instrumental, historical, and natural climate archives, including a few tree-ring series in the Mediterranean region (18). These we call the ‘Pauling’ reconstructions, after the lead author of that published result. The Pauling reconstructions are based on the same one-half degree grid as the OWDA and the domain covered by the OWDA covers most of the Pauling domain. This allows for nearly perfect one-to-one comparisons of the OWDA and Pauling grid point reconstructions to be made. The closest analog to JJA scPDSI is Pauling spring-summer (SPSU) precipitation total because scPDSI reflects antecedent springtime moisture conditions in its summer estimates through autoregressive persistence added to the monthly estimates (45).

Point-wise Pearson correlations of OWDA scPDSI with Pauling SPSU precipitation were calculated by century up to 1899, for the entire 1500-1900 pre-calibration period of the Pauling reconstructions, and for the 1901-1978 period of OWDA tree-ring scPDSI versus Pauling SPSU instrumental data. These correlation maps are shown in **Fig. SM12**. Given that the OWDA and Pauling data are both statistical reconstructions, it would not be easy to determine which one might be more at fault before 1900 in areas where their point-wise correlations were weak. For that reason, only correlations ≥ 0.30 (100 years, $p < 0.01$, 1-tailed, no correction for autocorrelation) are shown to emphasize the regions where we are confident that the two reconstructions truly match each other in a statistically meaningful way.

The correlation maps windowed by century and over the 1500-1900 full reconstruction period indicate considerable agreement between OWDA scPDSI and Pauling SPSU over most of central Europe, England, and Wales, with correlations often exceeding 0.50 ($p < 10^{-6}$). This is the

core region of predictor data used for both the OWDA (**Fig. SM1**) and Pauling reconstructions (18) (see their Figure 1). There are also significant levels of correlation over Turkey and Morocco. However, these results may be biased by the use of some common tree-ring data there in both the OWDA and Pauling reconstructions. The comparison of OWDA with the Pauling 1901-78 instrumental data shows excellent agreement, but this is undoubtedly biased by the calibration exercise used to produce the OWDA. The strong agreement between the OWDA and Pauling reconstructions in parts of central Europe (e.g., Germany, also in the pre-instrumental period) where the validation statistics indicated caution is warranted supports our inference that the reconstruction skill is underestimated.

One last comparison between the OWDA and Pauling reconstructions is shown in **Fig. SM13** for the “unprecedented” hot and dry year of 1540 in central Europe (21). The two maps are remarkably similar in terms of their spatial features of dryness and wetness over most of the domain common to both spatial reconstructions. Extreme summer drought and SPSU rainfall deficits are indicated in central Europe as expected (21). The most anomalous differences between the maps are in the British Isles and Turkey, the latter being a bit odd given some of the common tree-ring data used there. The Pauling map is also much smoother because of the reduced-space method of reconstruction used (18).

Overall, the comparisons of the OWDA JJA scPDSI and Pauling SPSU precipitation reconstructions have produced results that support the overall validity of the OWDA in central Europe and over much of the British Isles. Elsewhere it is not possible to determine the level of skill in the OWDA because the level of skill in the Pauling SPSU reconstructions is largely unknown. In the next section, the OWDA will be compared historical records of droughts and pluvials over Europe.

13. Additional validation tests of the OWDA: OWDA maps are provided for years of historically documented droughts and pluvials in Europe. Individual yearly maps are shown for some of the better-documented hydroclimatic events. Mean, median, and 95% significant level maps are provided in cases where more than six years are being evaluated or in cases where the uncertainty of the individual yearly historical accounts is high. See (13, 50, 105, 106) for discussions of the precautions needed in interpreting early historical records as accurate reflections of the weather because of issues related in part to ‘source validation’ (105).

Fig. SM14 shows maps of exceptional Czech lands droughts (20). The 1540 and 1616 droughts are described in the main paper. Besides those two droughts, the 1590, 1718, and 1719 maps show the remaining Brázdil “selected outstanding drought events”. The 1473 drought may be less well constrained by documentary evidence, but is listed in (20) as a JJA drought covering a large portion of the Czech lands (“Streams dried up. Forest fires. Poor harvest of cereals and other crops.”). All of these droughts are expressed well over the Czech lands and their full spatial extent over the Old World is revealed by the OWDA.

Fig. SM15 shows maps of the great European famine (22) due to a pluvial from 1314-1316. The wettest 1315 year is described in the main OWDA paper, but as noted in (22) 1314 was really the beginning year of three successive severe to extreme wet years in central Europe that contributed to subsequent crop failures and famine. By 1317 this pluvial had ended with a return to near-normal moisture conditions over central Europe. The years of excessive wetness are well expressed in the OWDA over the regions that were known from historical documents to be most strongly affected.

Fig. SM16 shows OWDA mean and median maps for nine noteworthy 17th century droughts over England and Wales: 1634, 1635, 1636, 1666, 1667, 1684, 1685, 1694, 1695 (33; see *Evidence for drought pre-1800* on page 23 of that report). The >95% map shows the region of drought that exceeds the 1-tailed 95% confidence level. The collective occurrence of these droughts is expressed well in the OWDA and their impacts are shown to have extended across the English Channel into France and Germany as well.

Fig. SM17 shows OWDA mean and median maps for eight noteworthy Ottoman Empire droughts: 1570, 1591, 1592, 1594, 1595, 1607, 1608, 1610 (107). These drought dates are based in part on previous tree-ring reconstructions (51, 52, 108, 109), which detracts somewhat from the independence of this comparison with the OWDA. Even so, the >95% map shows the region of drought over Turkey and into Greece that exceeds the 1-tailed 95% confidence level. Thus, evidence for Ottoman Empire droughts in the OWDA is highly consistent with previous reports.

Fig. SM18 shows OWDA mean and median maps for twelve noteworthy pre-1450 historical droughts in England and Wales: 1084, 1129, 1136, 1222, 1242, 1252, 1263, 1272, 1284, 1288, 1305, 1385 (110). The >95% map shows the region of drought that exceeds the 1-tailed 95% confidence level. These drought years were extracted from the table “An Index of More Noteworthy Meteorological Events”, where they are listed as summers that were both hot and dry (110). Year 1095 not used because OWDA shows wet, but conflicting chronicles in (110) suggest the occurrence of *drought in spring* and *excessive summer rains*. So this year appears to be unreliable. Nonetheless, the maps shown indicate that the OWDA has skill extending back to medieval times in England.

Collectively, these comparisons of the OWDA to documented historical hydroclimatic events in Europe leave little doubt that the Old World Drought Atlas has sufficient skill for use in climatological and historical analyses back at least to medieval times and probably earlier.

Table SM1. List of tree-ring chronologies used for producing the OWDA. The country (CNTRY) codes follow the ISO 3166-1 alpha-3 list (<http://unstats.un.org/unsd/tradekb/Knowledgebase/Country-Code>). The tree species codes (SPEC) are from <http://web.utk.edu/~grissino/downloads/species.pdf>. Location coordinates (LAT and LON) are in decimal degrees and are approximate depending on the regional nature of the samples used in each chronology. IFY, ILY, YRS are the first, last, and total years of each chronology. KIND is the tree-ring metric: TRW is total ring width and MXD is maximum latewood density. NSER is the total number of tree-ring series in each chronology and MSL is the median segment length of those series. DET is the method of detrending applied: SF-SSD is the signal-free single-series detrending method and SF-RCS is the 2-stage signal-free regional curve standardization method. LMFR is the low-to-medium frequency ratio of the total spectral power of the SF-RCS and SF-SSD chronologies at periods >100 years. There are no LMFR values for the SF-SSD chronologies. See the text for details.

NUM	CNTRY	SPEC	LAT	LON	IFY	ILY	YRS	KIND	NSER	MSL	DET	LMFR
1	ALB	PINI	40.25	21.00	1174	2010	837	TRW	311	167	SF-SSD	
2	ALB	PIHE	41.80	20.23	617	2008	1392	TRW	302	309	SF-RCS	1.74
3	ALB	FASY	40.62	20.78	1395	2007	613	TRW	97	147	SF-SSD	
4	AUT	PICE	47.50	11.17	0	2003	2004	TRW	1338	167	SF-RCS	1.24
5	BGR	PINI	41.60	24.57	1593	2009	417	TRW	53	282	SF-SSD	
6	BGR	PIHE	41.70	23.50	1143	2009	867	TRW	111	444	SF-SSD	
7	CHE	PCAB	47.00	7.50	1100	2012	913	TRW	1083	76	SF-RCS	7.98
8	CHE	ABAL	47.55	7.63	809	2012	1204	TRW	1393	109	SF-RCS	1.28
9	CHE	PISY	47.17	7.37	1341	1984	644	TRW	517	101	SF-RCS	6.02
10	CHE	PCAB	47.00	10.00	1079	2008	930	TRW	858	94	SF-RCS	9.69
11	CHE	QUPE	46.35	8.60	1562	2002	441	TRW	43	221	SF-SSD	
12	CHE	QUSP	47.08	8.00	755	2012	1258	TRW	1677	113	SF-RCS	1.57
13	CHE	LADE	46.33	10.00	800	2004	1205	TRW	1806	226	SF-RCS	0.72
14	CYP	PINI	34.92	32.87	1306	2006	701	TRW	238	261	SF-SSD	
15	CZE	ABAL	49.77	14.83	994	1996	1003	TRW	922	117	SF-RCS	1.10
16	DEU	PCAB	49.08	12.17	1456	2001	546	TRW	678	75	SF-RCS	5.73
17	DEU	QUSP	51.50	9.50	996	2005	1010	TRW	953	138	SF-RCS	8.31
18	DEU	QUSP	52.00	12.50	0	2010	2011	TRW	1392	110	SF-RCS	3.82
19	DEU	PCAB	49.08	12.17	1028	2001	974	TRW	2019	71	SF-RCS	3.62
20	DEU	QUSP	48.18	9.12	0	2010	2011	TRW	1141	106	SF-RCS	3.91
21	DEU	QUSP	54.00	9.67	501	1994	1494	TRW	1816	126	SF-RCS	3.43
22	DEU	ABAL	49.08	12.17	1010	1996	987	TRW	533	116	SF-RCS	1.40
23	DEU	QUSP	48.78	11.43	0	2010	2011	TRW	1921	121	SF-RCS	3.87
24	DNK	QUSP	56.50	9.37	487	1987	1501	TRW	597	135	SF-RCS	5.11
25	DZA	CEAT	36.00	5.00	912	2006	1095	TRW	391	240	SF-SSD	
26	DZA	PINI	36.45	4.10	1569	2005	437	TRW	39	137	SF-SSD	
27	ESP	PINI	40.00	-3.00	1300	2008	709	TRW	396	222	SF-SSD	
28	ESP	PIUN	42.53	1.42	1377	2005	629	TRW	42	210	SF-SSD	
29	ESP	PINI	37.82	-2.97	1125	2002	878	TRW	68	392	SF-SSD	
30	ESP	PIUN	42.53	1.42	1269	2003	735	TRW	57	450	SF-SSD	
31	ESP	PISY	42.00	-2.50	1511	2008	498	TRW	568	198	SF-SSD	
32	ESP	PIUN	42.65	1.00	924	2005	1082	TRW	235	202	SF-RCS	2.88
33	FIN	PISY	62.00	28.00	535	2002	1468	TRW	2228	147	SF-RCS	1.50
34	FIN	PISY	69.00	25.00	0	2004	2005	TRW	554	184	SF-RCS	1.00
35	FIN	PISY	61.50	28.50	535	1993	1459	TRW	558	152	SF-RCS	2.66
36	FIN	PISY	68.22	24.05	0	2006	2007	TRW	459	163	SF-RCS	0.60
37	FIN	PISY	68.22	24.05	0	2006	2007	MXD	603	134	SF-SSD	
38	FRA	PINI	42.07	9.20	1518	1980	463	TRW	30	191	SF-SSD	
39	FRA	QURO	48.45	2.68	651	2011	1361	TRW	598	113	SF-SSD	
40	FRA	ABAL	49.00	5.50	962	1995	1034	TRW	1009	111	SF-RCS	1.06
41	FRA	LADE	44.05	7.48	946	2007	1062	TRW	236	287	SF-SSD	
42	FRA	QUSP	49.00	5.50	0	2011	2012	TRW	1844	108	SF-RCS	5.52
43	GBR	QURO	52.00	-1.42	663	2009	1347	TRW	1830	98	SF-RCS	14.52
44	GBR	QURO	52.67	-1.00	880	2009	1130	TRW	707	119	SF-RCS	22.74
45	GBR	QURO	52.33	-1.20	406	2009	1604	TRW	2831	110	SF-RCS	5.46
46	GEO	ABAL	41.62	42.52	1676	2006	331	TRW	10	239	SF-SSD	
47	GEO	TABA	42.00	45.17	1526	1980	455	TRW	19	316	SF-SSD	
48	GRC	PIHE	40.32	20.88	1496	2010	515	TRW	16	335	SF-SSD	
49	GRC	PIHE	39.90	21.17	1291	2010	720	TRW	31	361	SF-SSD	
50	GRC	QUSP	41.50	28.25	1081	1985	905	TRW	455	106	SF-RCS	2.38

Table SM1 (continued). List of tree-ring chronologies used for producing the OWDA. See above for headings guide.

NUM	CNTRY	SPEC	LAT	LON	IFY	ILY	YRS	KIND	NSER	MSL	DET	LMFR
51	GRC	PIHE	39.98	21.85	1255	2008	754	TRW	23	445	SF-SSD	
52	GRC	PIHE	39.88	21.18	1225	2012	788	TRW	38	416	SF-SSD	
53	GRC	PINI	39.90	21.17	1232	2003	772	TRW	44	475	SF-SSD	
54	GRC	PIHE	21.02	40.10	575	2012	1438	TRW	83	328	SF-SSD	
55	GRC	QUSP	41.50	24.00	1170	1979	810	TRW	258	100	SF-RCS	2.32
56	IRL	QUSP	53.75	-7.75	0	2008	2009	TRW	1221	148	SF-RCS	3.70
57	ITA	PIHE	39.90	16.23	1227	2011	785	MXD	86	364	SF-SSD	
58	ITA	LADE	46.20	10.28	783	2008	1226	TRW	620	272	SF-RCS	1.13
59	ITA	PILE	39.90	16.23	1183	2011	829	TRW	148	330	SF-SSD	
60	ITA	PICE	45.90	9.57	961	2005	1045	TRW	769	202	SF-SSD	
61	ITA	PILA	39.38	16.57	1636	1999	364	TRW	35	215	SF-SSD	
62	JOR	JUPH	30.63	35.50	1469	1995	527	TRW	17	296	SF-SSD	
63	LBN	CELI	34.23	36.03	1382	2002	621	TRW	65	268	SF-SSD	
64	MAR	CEAT	34.00	-5.00	1291	2004	714	TRW	296	229	SF-SSD	
65	MAR	CEAT	35.03	-5.08	883	2005	1123	TRW	601	356	SF-SSD	
66	NLD	QUSP	51.97	6.00	624	2009	1386	TRW	615	115	SF-RCS	3.31
67	NOR	PISY	69.08	17.22	1403	1997	595	TRW	66	250	SF-SSD	
68	NOR	PISY	60.25	9.03	1131	2011	881	TRW	740	136	SF-RCS	1.11
69	NOR	PISY	68.80	15.73	945	1994	1050	TRW	70	260	SF-SSD	
70	NOR	PISY	61.33	7.00	1192	2007	816	TRW	199	210	SF-SSD	
71	POL	PISY	53.50	19.67	1126	2008	883	TRW	329	106	SF-RCS	5.73
72	POL	QURO	53.95	14.50	1554	1986	433	TRW	23	322	SF-SSD	
73	POL	PCAB	49.23	20.08	1647	2000	354	TRW	81	107	SF-SSD	
74	POL	QURO	53.50	18.67	685	1985	1301	TRW	669	123	SF-RCS	3.38
75	ROM	PCAB	47.45	25.47	1631	2010	380	TRW	153	236	SF-SSD	
76	ROM	PICE	47.25	25.25	994	2005	1012	TRW	408	216	SF-SSD	
77	ROM	PCAB	47.12	25.25	1637	2010	374	TRW	101	235	SF-SSD	
78	ROM	PCAB	45.63	24.68	1593	2010	418	TRW	232	221	SF-SSD	
79	RUS	PISY	68.42	35.28	1577	1997	421	TRW	49	174	SF-SSD	
80	RUS	PISY	64.75	31.00	1578	2002	425	TRW	84	265	SF-SSD	
81	RUS	LASI	67.42	70.00	0	1996	1997	TRW	252	148	SF-SSD	
82	RUS	PISY	66.27	30.80	1471	2002	532	TRW	172	277	SF-SSD	
83	SVK	ABAL	48.67	19.70	1109	1998	890	TRW	242	69	SF-SSD	
84	SVK	LADE	49.15	19.93	991	2011	1021	TRW	484	132	SF-RCS	1.27
85	SVN	QUSP	45.80	15.00	1442	2003	562	TRW	205	98	SF-SSD	
86	SWE	PISY	57.82	18.68	1127	2011	885	TRW	128	134	SF-RCS	7.47
87	SWE	PISY	60.45	14.73	1450	2011	562	TRW	99	177	SF-SSD	
88	SWE	PISY	63.50	15.50	908	2008	1101	TRW	262	172	SF-RCS	2.55
89	SWE	PISY	59.37	17.75	1470	2011	542	TRW	303	134	SF-SSD	
90	SWE	QUSP	56.50	13.50	720	1991	1272	TRW	404	100	SF-RCS	6.68
91	SWE	PISY	68.20	19.80	0	2004	2005	TRW	649	185	SF-RCS	2.64
92	TUR	JUEX	36.65	30.08	1032	2006	975	TRW	255	275	SF-SSD	
93	TUR	CELI	36.75	32.17	1551	2003	453	TRW	48	211	SF-SSD	
94	TUR	QUSP	40.50	29.50	1081	1985	905	TRW	197	119	SF-RCS	1.62
95	TUR	QUSP	40.50	32.00	1081	2004	924	TRW	360	143	SF-RCS	2.93
96	TUR	CELI	36.67	31.25	1449	2001	553	TRW	100	262	SF-SSD	
97	TUR	ABNO	41.30	41.90	1498	2007	510	TRW	45	323	SF-SSD	
98	TUR	PINI	37.33	32.33	1444	2010	567	TRW	160	284	SF-SSD	
99	TUR	PISY	40.50	42.33	1613	2001	389	TRW	306	189	SF-SSD	
100	TUR	PINI	41.00	32.00	1427	2008	582	TRW	146	287	SF-SSD	
101	TUR	JUEX	37.23	34.52	1235	2001	767	TRW	70	247	SF-SSD	
102	TUR	PINI	37.17	29.83	1159	2009	851	TRW	345	242	SF-SSD	
103	UKR	PCAB	48.15	24.53	1607	2012	406	TRW	274	252	SF-SSD	
104	UKR	FASY	48.27	23.62	1559	2009	451	TRW	171	107	SF-SSD	
105	UKR	PCAB	48.15	24.52	1676	2003	328	TRW	67	152	SF-SSD	
106	UKR	PCAB	48.60	23.92	1651	2012	362	TRW	207	229	SF-SSD	

OWDA scPDSI Target Field and Tree-Ring Network

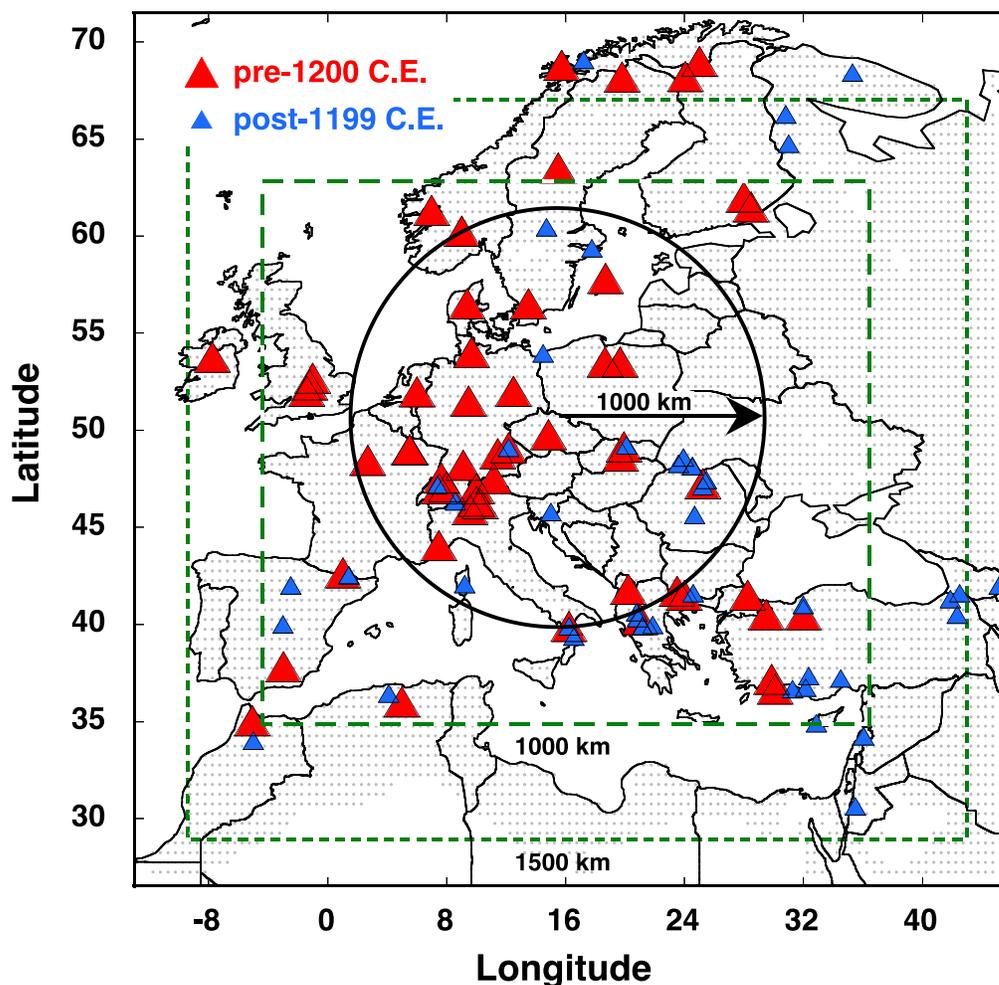


Fig. S1. Map of the OWDA domain showing 5414 half-degree grid points of JJA scPDSI (small black dots) and the 106 annual chronology tree-ring network (red and blue triangles). The tree-ring network covers most of the OWDA domain reasonably well except for the easternmost portion comprised mostly of Russia from which no tree-ring chronologies were available. The tree-ring chronologies have been coded for those beginning before 1200 C.E. (large red triangles) and those beginning after 1199 C.E. (small blue triangles) to provide a rough sense of how the time spans of the chronologies are spatially distributed. The circle shows an example 1000 km search radius used for locating a minimum of 20 tree-ring chronologies to reconstruct each scPDSI grid point. The long dashed rectangle shows the general (N-S-E-W) geographic limits of the domain for which a 1000 km search radius was sufficient to find a minimum of 20 chronologies. The short dashed rectangle show the limits of the additional domain area covered when the search radius was dynamically expanded out to 1500 km. See the text for more information.

CRU TS 3.2 Precipitation Station Coverage By Decade

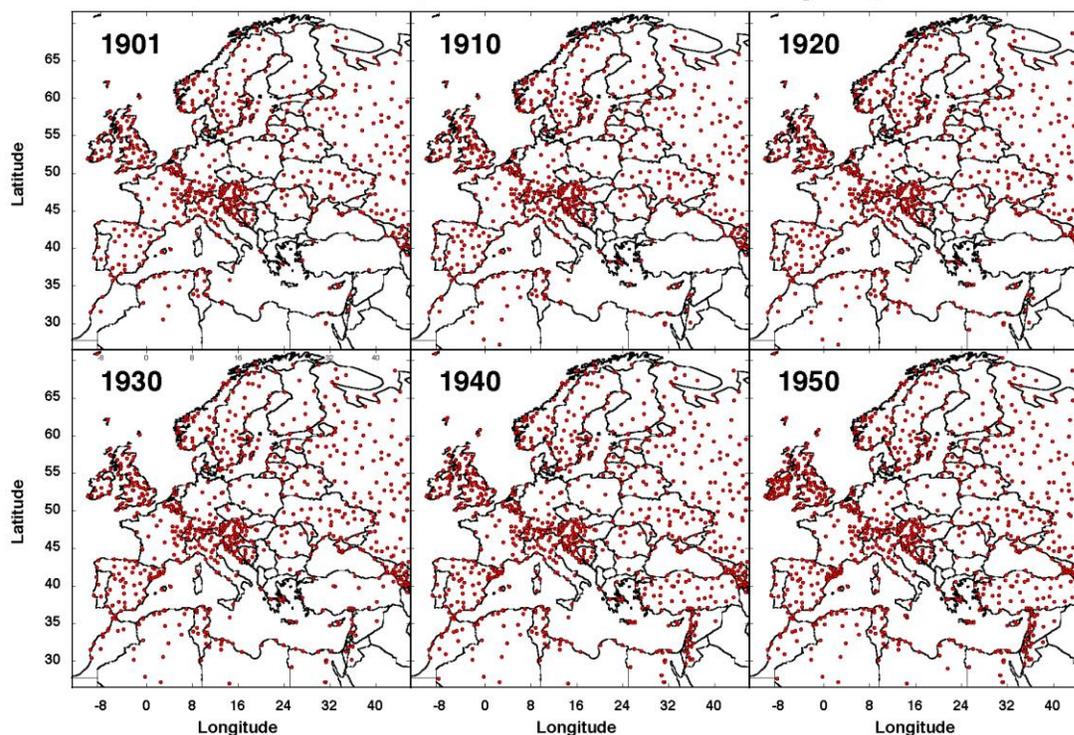


Fig. S2. Maps, by decade (up to 1950), of the changing densities of precipitation stations (solid red dots) available for interpolation on the half-degree regular grid used to produce the CRU TS precipitation field (<http://badc.nerc.ac.uk>). The density of stations is high and stable over most of the OWDA domain. Only in Turkey, the Middle East, and parts of North Africa are there obvious losses of local precipitation data for interpolation prior to 1930.

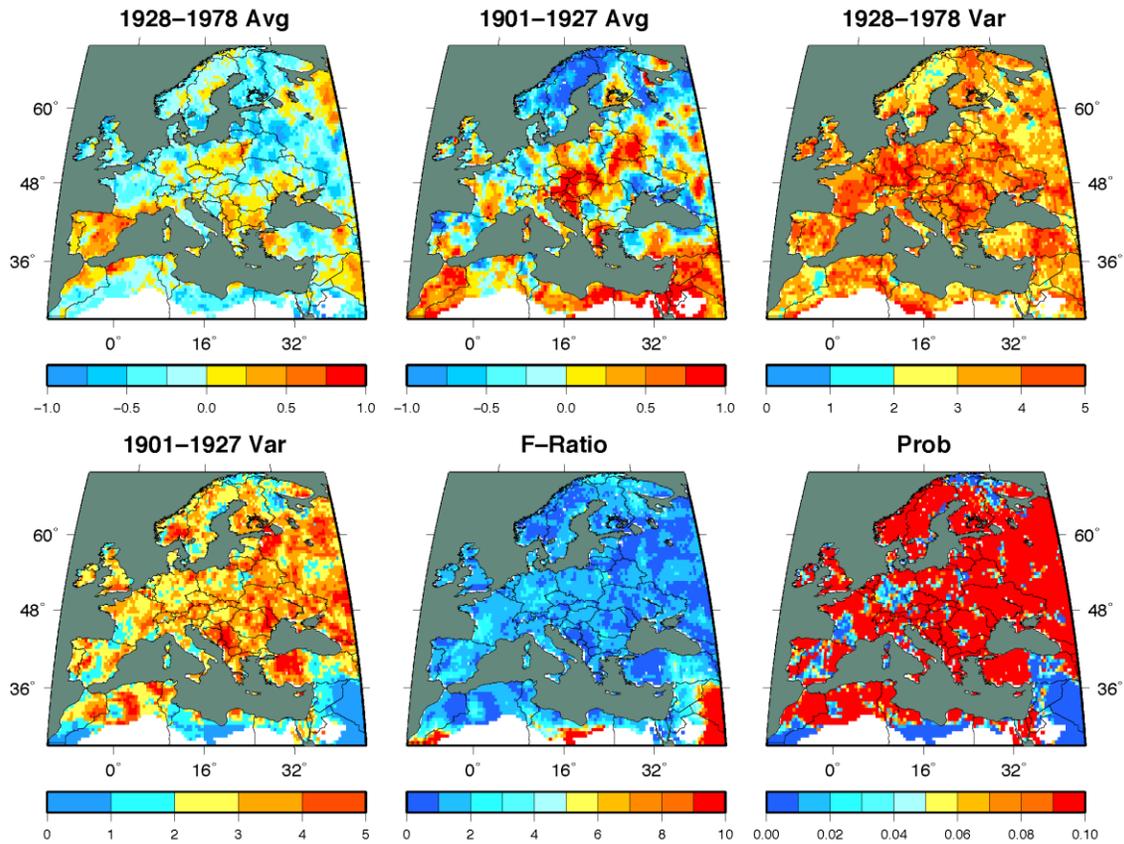


Fig. S3. Comparisons of calibration period (1928–1978) and validation period (1901–1927) scPDSI averages and their variances. The primary purpose here is to test for equal variances using the Shoemaker F-test (34) to determine the likelihood that there has been a significant loss of variance (‘relaxation to climatology’) in the scPDSI values back in time due to reduced precipitation station data as indicated in **Fig. SM2** for Turkey and the Middle East before 1930. The F-ratios indicate a statistically significant ($p < 0.01$) reduction of variance (red areas) in the validation period scPDSI values in eastern Turkey, the Middle East, and areas along the northern fringe of the Sahara Desert, which is highly consistent with the loss of stations back in time shown in **Fig. SM2**.

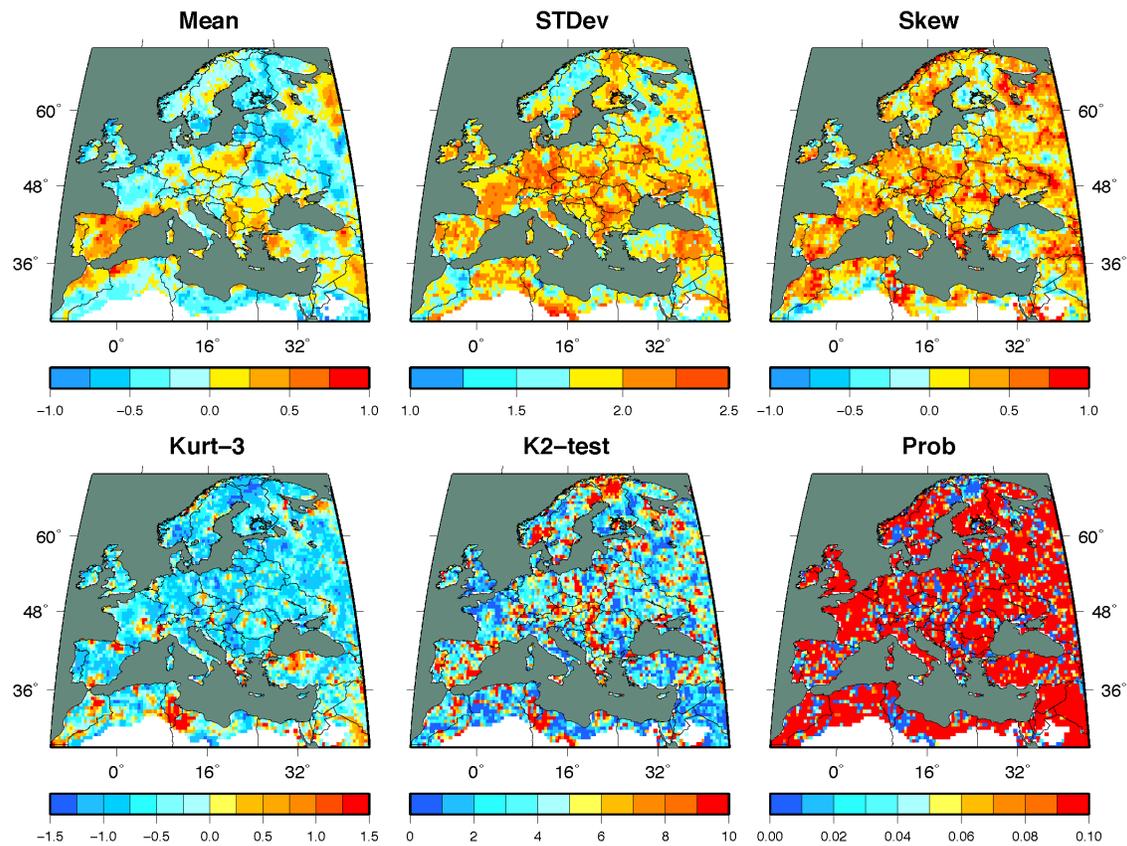


Fig. S4. Statistical properties of gridded summer scPDSI data over the 1928–1978 calibration period and tests of normality using a simple and robust test of normality based on the joint use of skewness and kurtosis (69). The expected value of 3.0 for kurtosis of the normal distribution has been subtracted (Kurt-3) The scPDSI data show a tendency for longer right-hand tails (Skew>0), but also thinner (platykurtic) tails in general (Kurt-3.0<0).

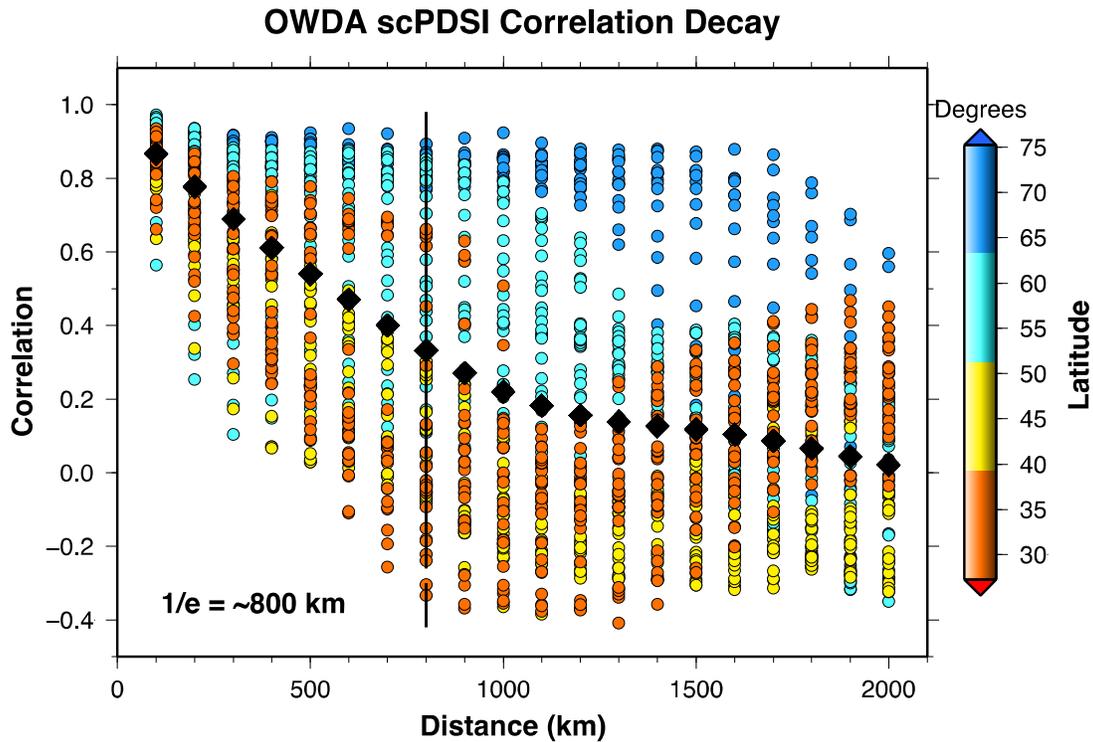


Fig. S5. CDL between 5414 grid points of summer scPDSI used for reconstruction over the OWDA domain. The pairwise correlations have been binned and averaged by 100 km distance increments and by one-half degree latitude band out to a maximum of 2000 km. A 40% robust LOESS smooth has been used to estimate the overall shape of the correlation decay function and at what distance it decays to $1/e$, the e -folding distance. The LOESS smooth indicates that the average e -folding distance is ~ 800 km, but there is a clear latitudinal dependence in it, with the e -folding distance increasing from south to north.

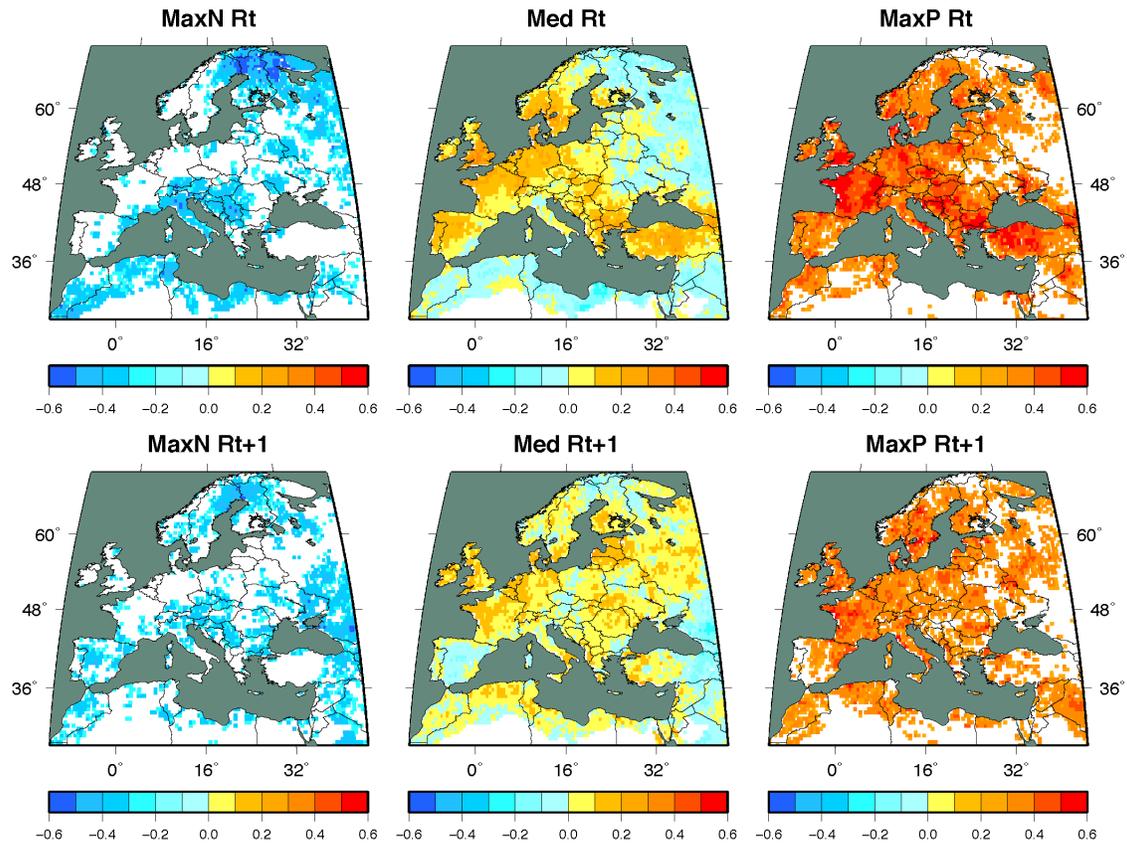


Fig. S6. Summary maps of correlations between summer scPDSI and the tree-ring network over the 5414 grid points of the OWDA domain calculated for the 1928–1978 calibration period, using the PPR program in the same way that it was used to produce the OWDA reconstructions. The correlations shown are both contemporaneous (year t scPDSI vs. year t tree rings - **Rt**) and lagged (year t scPDSI vs. year $t+1$ tree rings - **Rt+1**) to take into account the carryover effect of climate on tree growth. Due to the large number of correlations between scPDSI and tree rings produced by PPR at each grid point, the results are condensed into a series of maps that show the spatial distributions of the median (Med), maximum negative (MaxN), and maximum positive (MaxP) correlations for year t and year $t+1$. The color-coded grid points in MaxN and MaxP correlation maps are statistically significant at the 2-tailed 95% level. White areas are either not significant ($p>0.05$) or missing in the Sahara Desert region.

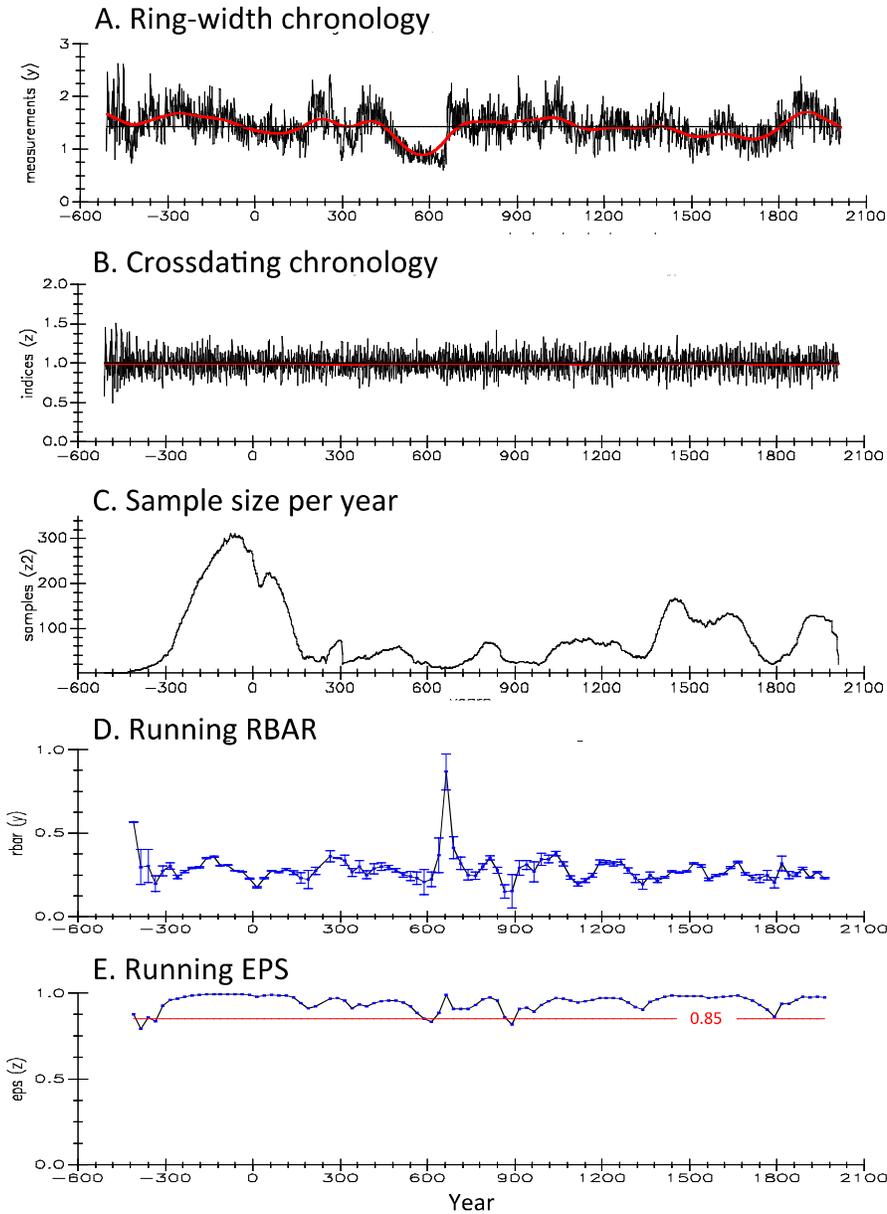


Fig. S7. Example of a historical/modern tree-ring chronology from northeastern France developed by the iterative procedure described in the text. The mean ring-width chronology shown in (A) illustrates the inhomogeneous properties of the raw measurements over time that render them unsuitable for climate reconstruction without further processing. The crossdating chronology (B) was developed in the same way as done in COFECHA for testing for common signal. It is much more homogeneous over time, as the iterative crossdating procedure is designed to produce, but cannot be used for climate reconstruction. The remaining plots provide information on changing sample size and the strength of crossdating. See the text for details.

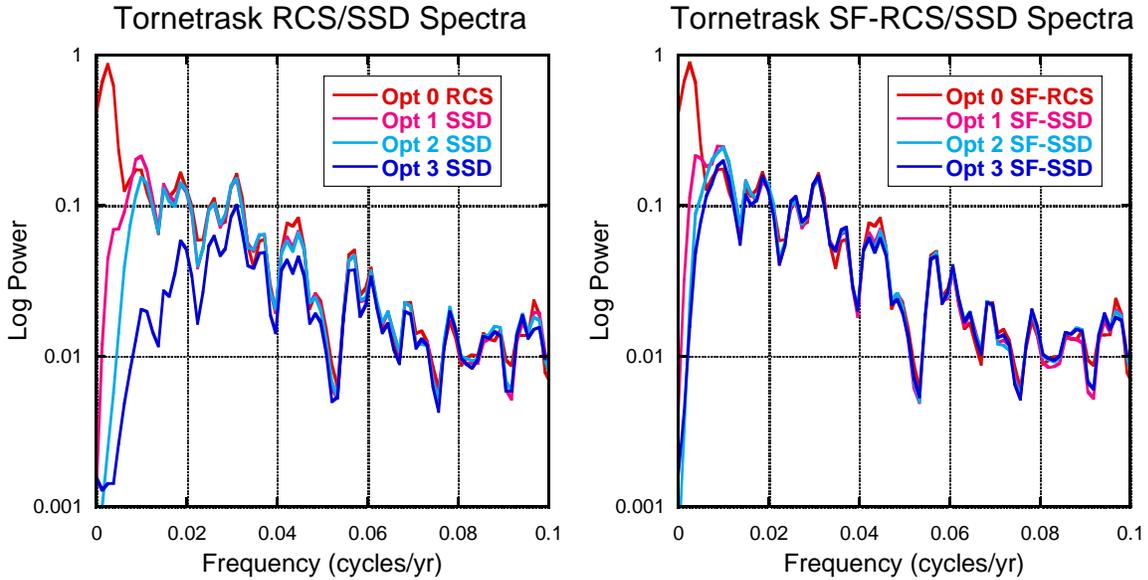
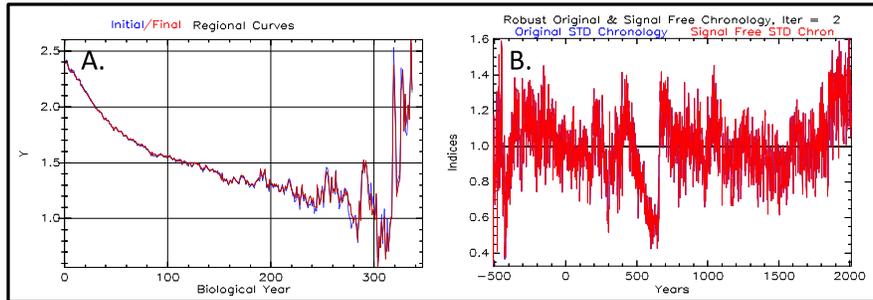


Fig. S8. Overlay plots of Tornetrask power spectra (frequencies from 0 to 0.1) for different detrending options before (RCS/SSD) and after (SF-RCS/SSD) the application of the signal-free method to the data using the same curve-fitting options: Opt 0—RCS detrending (designed to preserve the most low- to medium-frequency variance); Opt 1—negative exponential/linear detrending (monotonic nonincreasing, least flexible SSD option); Opt 2—cubic smoothing spline detrending based on the median segment length of the data (moderately data adaptive, fixed intermediate flexibility); and Opt 3—the Friedman variable span smoother (locally adaptive, very flexible). Opt 0 using the RCS detrending method preserves the most LMF variance in the data as expected and there is virtually no difference between the RCS and SF-RCS spectra in this case. The increasing flexibility of the curve fit options from Opt 1 to Opt 3 has resulted, also as expected, in progressively larger losses of LMF variance in the STD chronologies. SF-SSD has recovered most of the lost common medium frequency (decadal to century) variance caused by the increasingly flexible detrending curves from Opt 1 to Opt 3. Thus, SF-SSD can provide a degree of protection from the loss of medium frequency variance caused by flexible detrending methods.

2-Stage SF-RCS Applied to Northeastern France Historical/Modern QUSP Tree-Ring Data

Stage 1



Stage 2

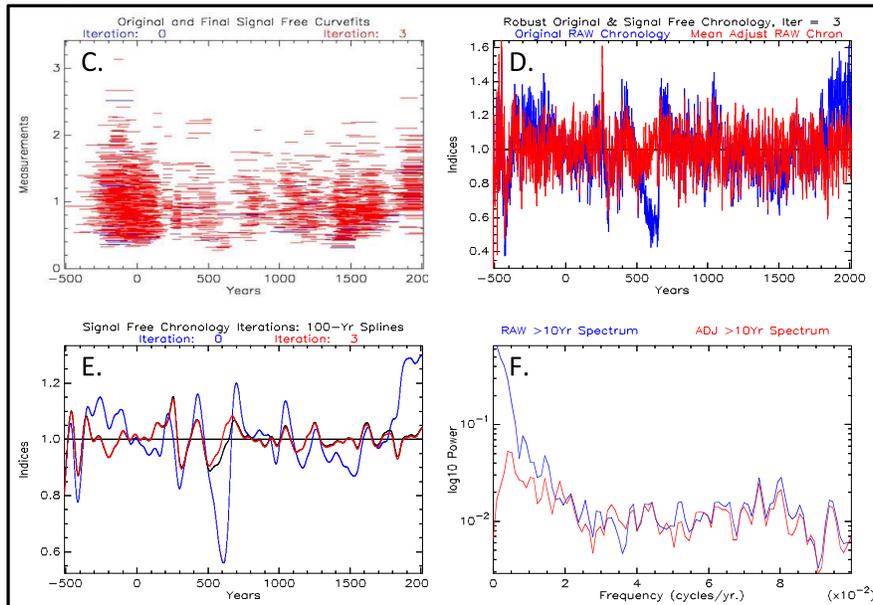


Fig. S9. An example of the two-stage SF-RCS method applied to the historical/modern *Quercus* species (QUSP) tree-ring data of northeastern France used as an example by Auer *et al.* (48). *Augmenting tree-ring chronologies with historical tree-ring data* and shown in Fig. SM7. **Fig. 9A-B** shows the results of applying single-curve SF-RCS to the data. This is ‘Stage 1’ of the 2-stage SF-RCS procedure. The mean biological growth curves in (A) are virtually identical before and after two iterations of signal-free detrending are applied. Consequently, there is virtually no difference between the before-and-after RCS chronologies shown (B). **Fig. 9C-F** show the effects of the ‘Stage 2’ adjustment to the SF-RCS chronology that results from the SF-iterated removal of the mean differences (C) between the tree-ring index series after Stage-1 SF-RCS detrending. The resulting Stage-2 SF-RCS chronology (red series in D) shows the large, but localized, effects of this mean adjustment process. The SF-iterated effects of this process are most easily seen in the mean chronology smoothed curves (E) from the beginning Stage-1 input (blue) to final Stage-2 SF-iterated output (red). The loss of spectral power in the chronology from Stage-1 to Stage-2 (F) is apparent, especially at periods >100 years. See the text for details.

Additional LMF Variance Retained (>100 years) in OWDA
Historical/Modern Tree-Ring Chronologies Using 2-Stage SF-RCS

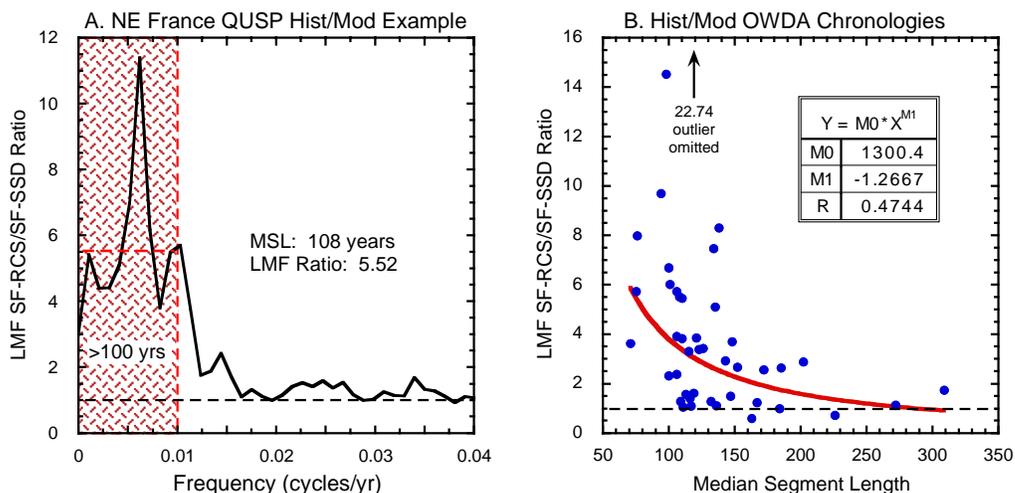


Fig. S10. Additional low- to medium-frequency variance retained in the historical/modern tree-ring chronologies using the two-stage SF-RCS procedure. The ratio of the power spectra (SF-RCS/SF-SSD) for the northeastern France QUSP data set is shown in **(A)**, plotted for periods from DC to 25 years. The largest effect of SF-RCS is indicated (red hatching) for periods longer than the median segment length of 108 years. The SF-RCS chronology has 5.52 times as much variance on average over periods >100 years compared to the SF-SSD chronology. This we call the 'LMF-Ratio'. An LMF-Ratio > 1 means that 2-stage SF-RCS is preserving more variance than SF-SSD for the given detrending option used. A summary plot of the LMF-Ratio for all 43 historical/modern chronologies as a function of MSL is shown in **(B)**. There is a curvilinear relationship that is reasonably approximated by a power function, which implies a log-log relationship between MSL and LMF-Ratio. (One extreme LMF-Ratio value of 22.74 was not used to calculate the fit.) This plot indicates that the historical/modern chronologies benefit most from 2-stage SF-RCS when MSL is less than 150 years.

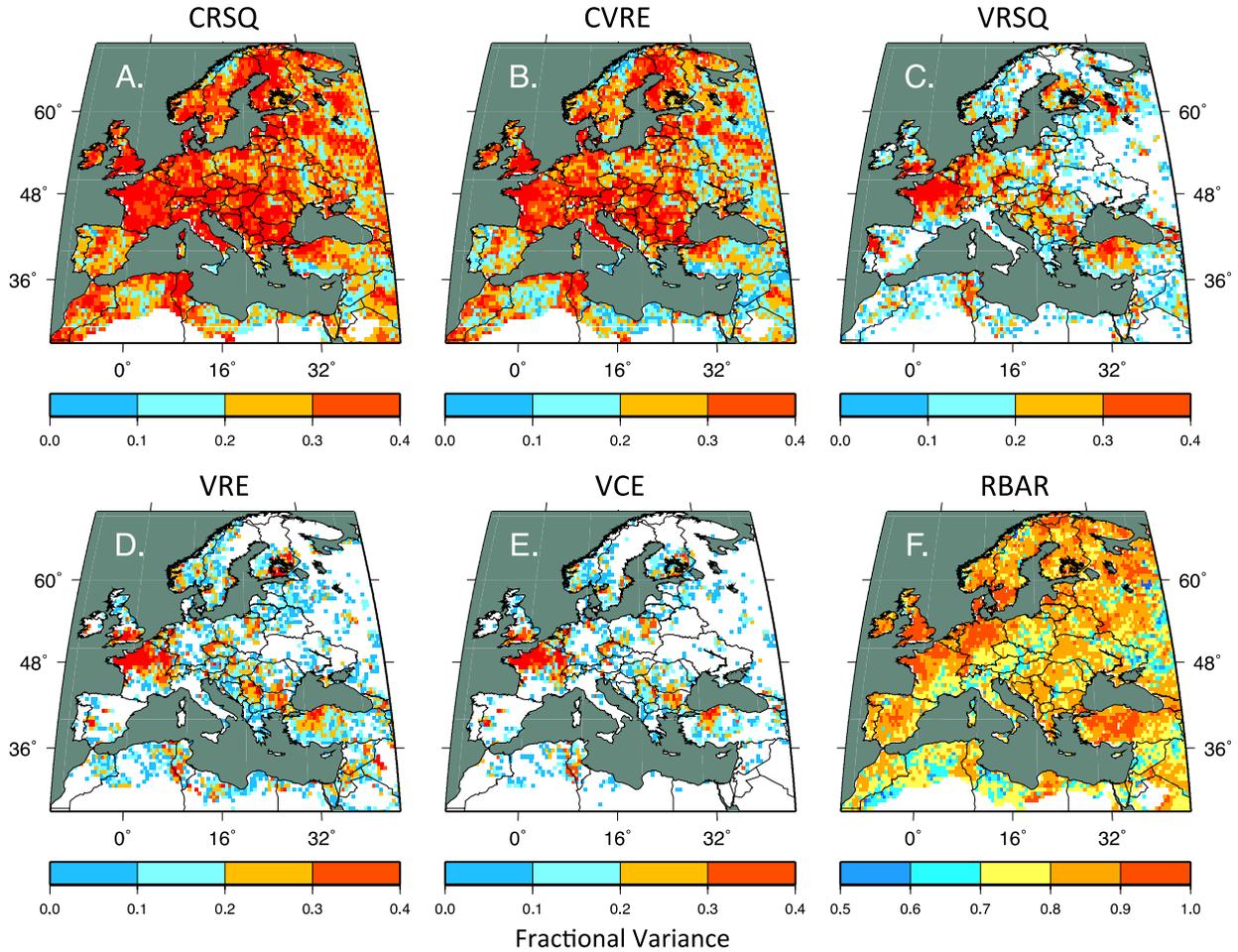


Fig. S11. Calibration and validation statistics maps of the eight-member ensemble-average OWDA reconstructions. All statistics are in units of fractional variance. Only those validation grid points that passed (VRSQ with $p < 0.10$ 1-tailed, VRE and VCE > 0) are plotted. The number of grid points plotted (max=5414) and their median values are given. (A) Calibration period regression coefficient of determination (CRSQ, $N=5414$, Median=0.328), (B) Calibration period leave-one-out cross-validation reduction of error (CVRE, $N=5414$, Median=0.271), (C) Validation period Pearson correlation squared (VRSQ, $N=3057$, Median=0.198), (D) Validation period reduction of error (VRE, $N=2495$, Median=0.161), (E) Validation period coefficient of efficiency (VCE, $N=1529$, Median=0.146), (F) Average correlation between ensemble member reconstructions (RBAR, $N=5414$, Median=0.820).

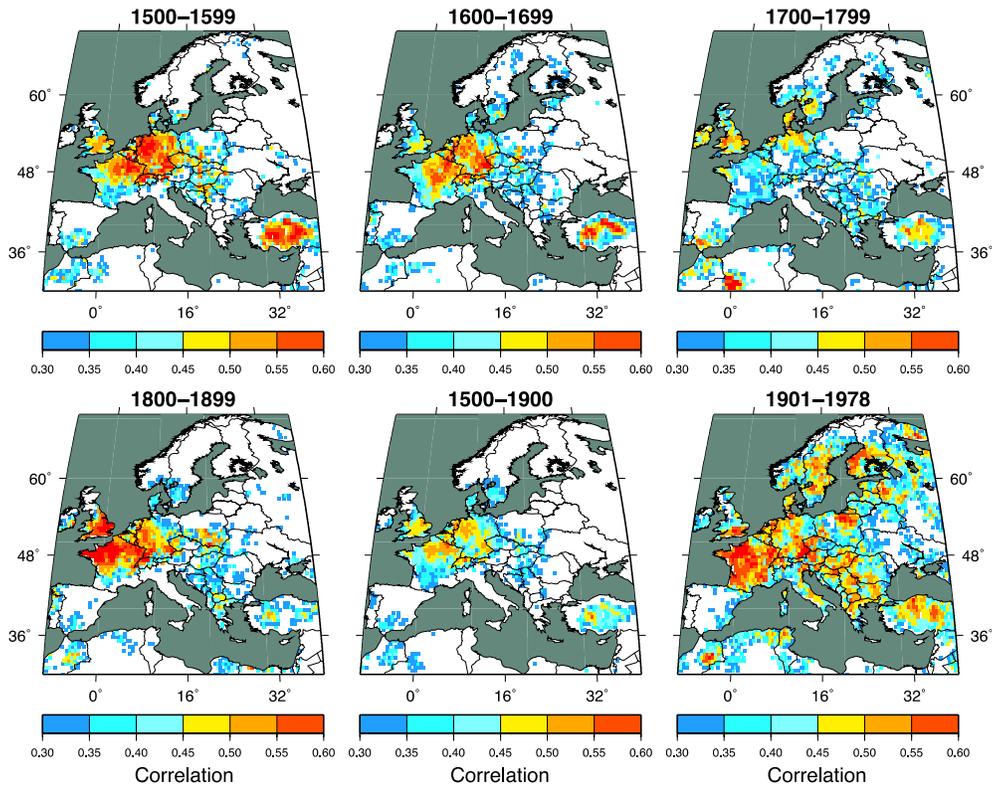


Fig. S12. Correlations of OWDA JJA scPDSI reconstructions with Pauling spring-summer precipitation reconstructions primarily reconstructed from long instrumental and historical climate indices (18). The data used are from <http://www.ncdc.noaa.gov/paleo/pubs/pauling2006/pauling2006.html>. A few tree-ring series in Morocco and Turkey were also used in the Pauling reconstructions, so there is a small degree of data overlap in the Mediterranean sector. The correlations are windowed by century and for the entire 1500-1900 pre-calibration interval of the Pauling reconstructions. The Pauling 1901-1978 data are instrumental.

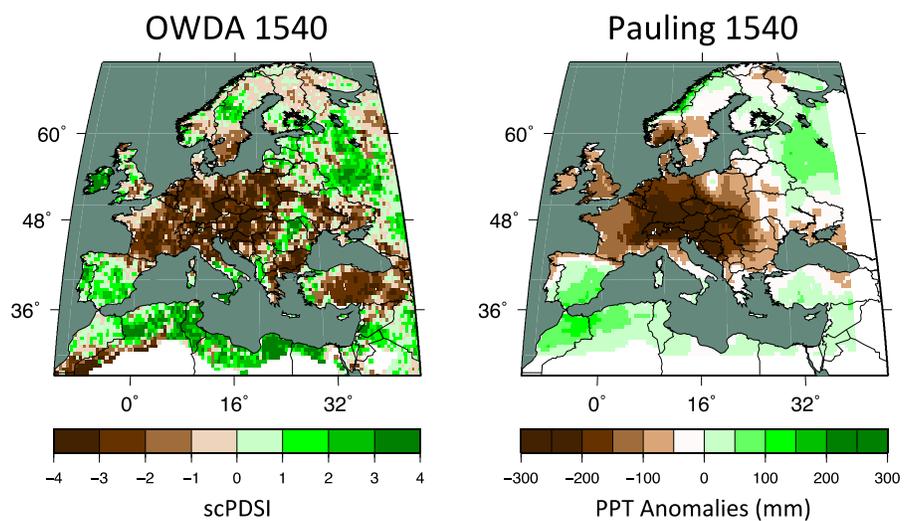


Fig. S13. Comparison of OWDA and Pauling maps for 1540, (“year-long unprecedented European heat and drought”) (21).

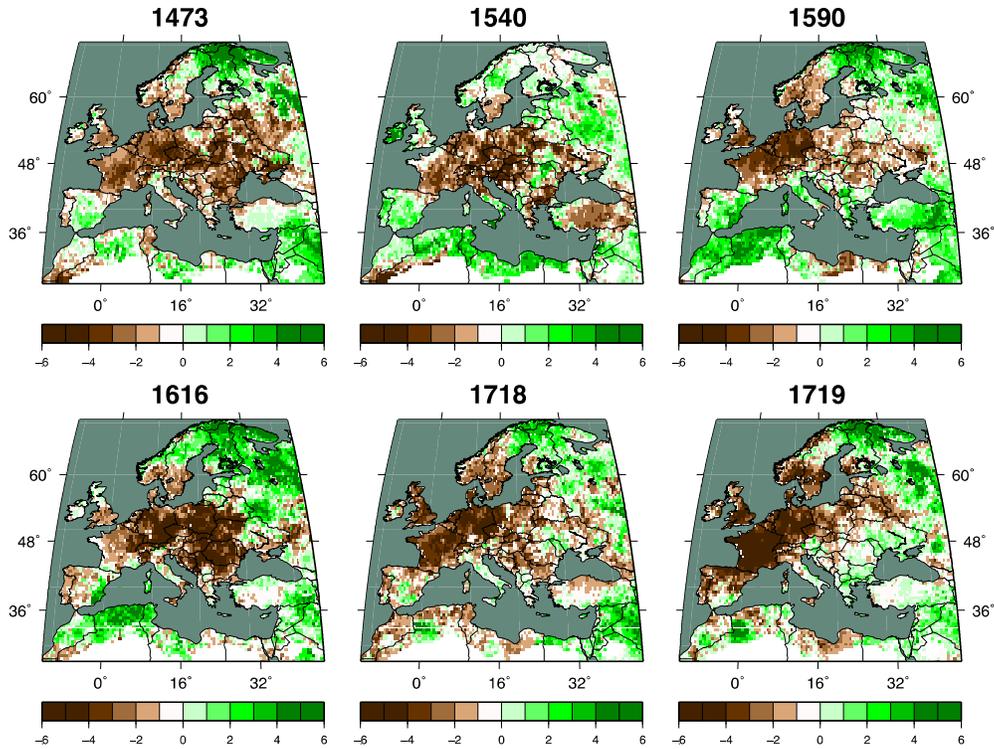


Fig. S14. Maps of exceptional droughts in Czech lands (20). Units are in scPDSI from extreme dry (<-4) to extreme wet ($>+4$). The 1540 and 1616 droughts are described in the main paper. Besides those two droughts, the 1590, 1718, and 1719 maps show the remaining Brázdil “selected outstanding drought events.” The 1473 drought may be less well constrained by documentary evidence, but is listed in that paper as a JJA drought covering a large portion of the Czech lands (“Streams dried up. Forest fires. Poor harvest of cereals and other crops.”). All of these droughts are expressed well over the Czech lands and their full spatial extent over the Old World is revealed by the OWDA.

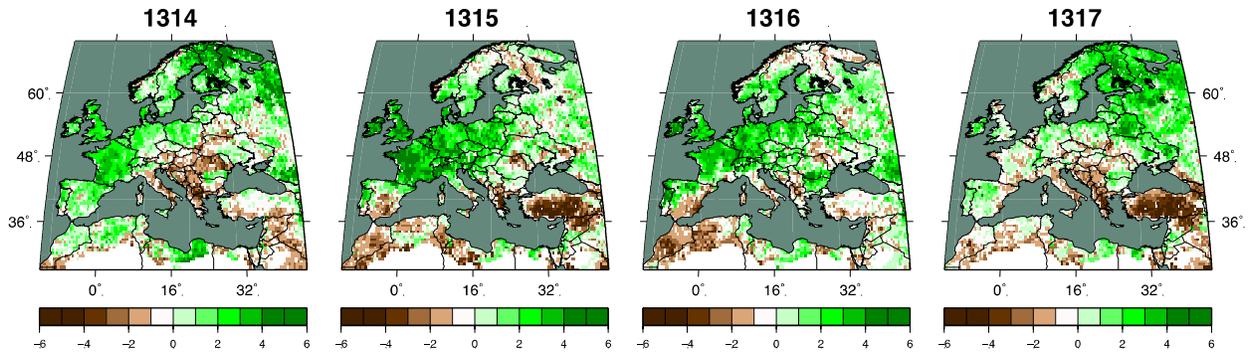


Fig. S15. Maps of the great European famine (22). Units are in scPDSI from extreme dry (<-4) to extreme wet ($>+4$). The wettest 1315 year is described in the main paper, but as noted 1314 was really the beginning year of three successive severe to extreme wet years in central Europe that contributed to crop failures and famine. By 1317 this pluvial had ended with a return to near-normal moisture conditions over central Europe.

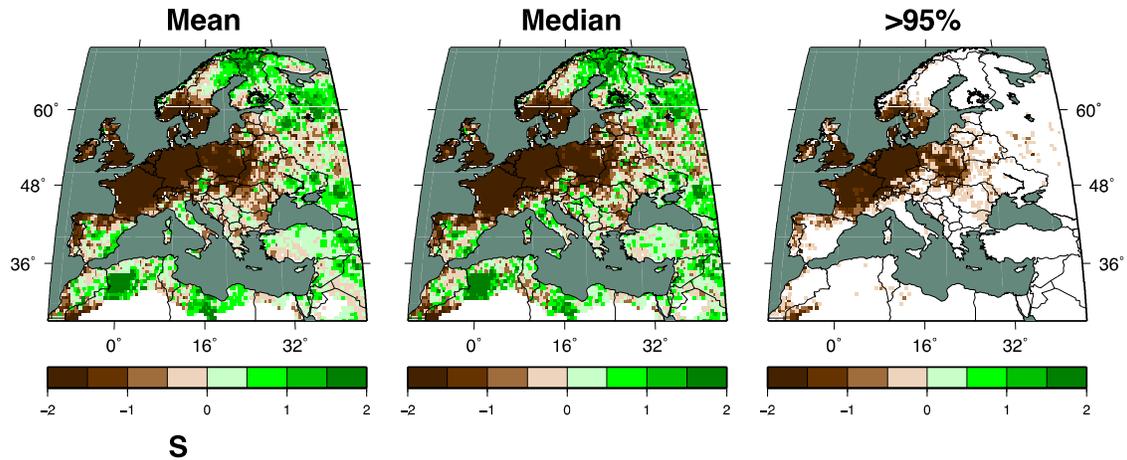


Fig. S16. OWDA mean and median maps for nine noteworthy 17th-century droughts over England and Wales: 1634, 1635, 1636, 1666, 1667, 1684, 1685, 1694, and 1695 (33). The >95% map shows the region of drought that exceeds the 1-tailed 95% confidence level. The impact of these droughts is estimated by the OWDA to have extended well across the English Channel into France and Germany.

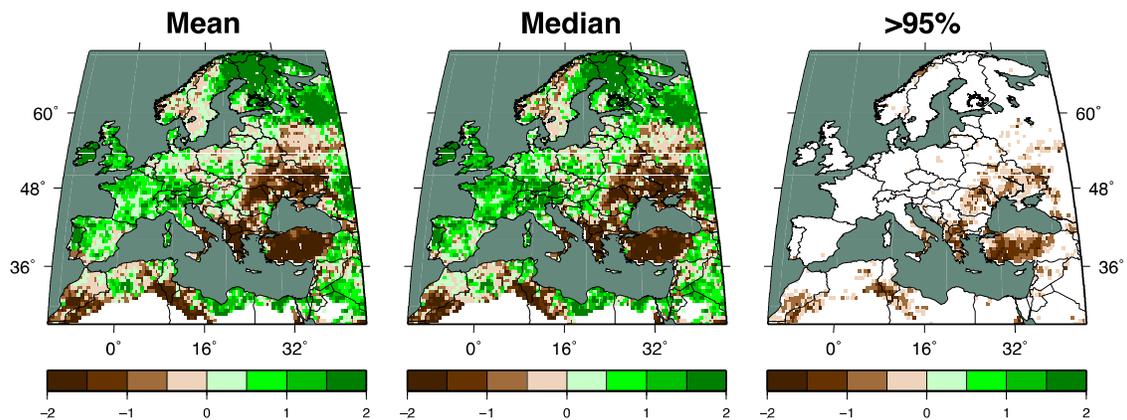


Fig. S17. OWDA mean and median maps for eight noteworthy Ottoman Empire droughts: 1570, 1591, 1592, 1594, 1595, 1607, 1608, and 1610 (107). The drought dates are based in part on previous tree-ring reconstructions (51, 52, 108, 109). The >95% map shows the region of drought that exceeds the 1-tailed 95% confidence level. Evidence for Ottoman Empire droughts in the OWDA is consistent with previous reports.

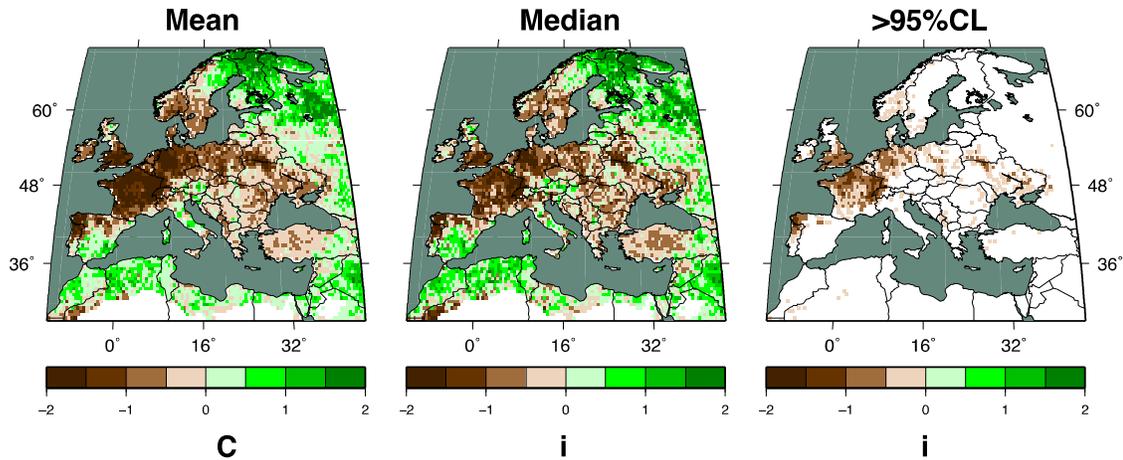


Fig. S18. OWDA mean and median maps for 12 noteworthy pre-1450 historical droughts in England and Wales: 1084, 1129, 1136, 1222, 1242, 1252, 1263, 1272, 1284, 1288, 1305, and 1385 (110). The >95% map shows the region of drought that exceeds the 1-tailed 95% confidence level. These drought years taken from the table “An Index of More Noteworthy Meteorological Events” and are listed there as summers that were both hot and dry. Year 1095 was not used because OWDA shows wet, but conflicting chronicles in (110) say both *drought in spring* and *excessive summer rains*. So this year appears to be unreliable.