Article

Fennoscandian tree-ring anatomy shows a warmer modern than medieval climate

🔴 Check for updates

Jesper Björklund[1,2✉], Kristina Seftigen[1,3], Markus Stoffel[4,5,6], Marina V. Fonti[1,2], Sven Kottlow[1], David C. Frank[7], Jan Esper[8,9], Patrick Fonti[1,2], Hugues Goosse[10], Håkan Grudd[11], Björn E. Gunnarson[12,13], Daniel Nievergelt[1,2], Elena Pellizzari[14], Marco Carrer[14,15] & Georg von Arx[1,2,15]

Earth system models and various climate proxy sources indicate global warming is unprecedented during at least the Common Era[1]. However, tree-ring proxies often estimate temperatures during the Medieval Climate Anomaly (950–1250 CE) that are similar to, or exceed, those recorded for the past century[2,3], in contrast to simulation experiments at regional scales[4]. This not only calls into question the reliability of models and proxies but also contributes to uncertainty in future climate projections[5]. Here we show that the current climate of the Fennoscandian Peninsula is substantially warmer than that of the medieval period. This highlights the dominant role of anthropogenic forcing in climate warming even at the regional scale, thereby reconciling inconsistencies between reconstructions and model simulations. We used an annually resolved 1,170-year-long tree-ring record that relies exclusively on tracheid anatomical measurements from *Pinus sylvestris* trees, providing high-fidelity measurements of instrumental temperature variability during the warm season. We therefore call for the construction of more such millennia-long records to further improve our understanding and reduce uncertainties around historical and future climate change at inter-regional and eventually global scales.

Over the past decades, considerable effort and debate have been devoted to how putative climate episodes, such as the Medieval Climate Anomaly (MCA; 950–1250 CE) or the Little Ice Age (1450–1850 CE) are characterized and relate to the Current Warm Period (CWP)[6–8]. Although the view that CWP is globally warmer than any other period during the Common Era is firmly consolidated[1], large uncertainties persist in models and proxies at regional scales[3,8–10], particularly before 1400 CE. In Europe, temperatures during the medieval period were found to be equal to, or perhaps warmer than, those during the twentieth century[3], leading to speculation that medieval warming was a critical factor for Norse settlements in Greenland[11] in 985 CE. However, an exceptionally warm MCA is difficult to reconcile with forced climate model simulations[4], thus suggesting that the unforced climate variability was larger and more persistent than indicated by current models. Reducing uncertainties, particularly during the MCA and the Medieval Quiet Period[12] (MQP; 725–1025 CE), is therefore urgently needed to further evaluate the sensitivity and internal variability of the climate system[13], not least to make future projections more robust[5]. Simulations and reconstructions of the MQP and MCA can provide important benchmarks of the amplitude of natural climate variability caused by limited volcanic forcing and relatively unperturbed solar forcing[12] for comparison with that of the anthropogenically forced CWP.

Tree-ring data are central to most empirical studies characterizing temperature changes during the Common Era[14] and are therefore often critically assessed for their ability to capture key characteristics of historical climate variations[15–21]. Debates are ongoing around the ability of the two most commonly used proxy sources—tree-ring width and maximum latewood density (MXD)—to intrinsically reflect climate change over the past millennia[15,16,22], to track the full amplitude of climate extremes[21,23–25] and to express spectrally unbiased climate transformations[17–19,24,26]. Because of recent progress in retrieving information from tree rings[27,28] we investigated whether some of these uncertainties could be reduced by directly analysing the dimensions of the wood cells (Fig. 1). This approach based on quantitative wood anatomy (QWA) offers remarkable measurement precision[29] and often substantially increases climate fidelity[30–33] when used for tree rings. However, tree-ring anatomy records that assess both the natural and anthropogenic ranges of climate variability over longer time periods and with sufficient robustness have been lacking.

[1]Swiss Federal Institute for Forest Snow and Landscape Research WSL, Birmensdorf, Switzerland. [2]Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland. [3]Regional Climate Group, Department of Earth Sciences, University of Gothenburg, Gothenburg, Sweden. [4]Climate Change Impacts and Risks in the Anthropocene (C-CIA), University of Geneva, Geneva, Switzerland. [5]Dendrolab.ch, Department of Earth Sciences, University of Geneva, Geneva, Switzerland. [6]Department F.-A. Forel for Environmental and Aquatic Sciences, University of Geneva, Geneva, Switzerland. [7]Laboratory of Tree-Ring Research, University of Arizona, Tucson, AZ, USA. [8]Department of Geography, Johannes Gutenberg University, Mainz, Germany. [9]Global Change Research Institute of the Czech Academy of Sciences (CzechGlobe), Brno, Czech Republic. [10]Earth and Life Institute, Université Catholique de Louvain (UCLouvain), Louvain-la-Neuve, Belgium. [11]Swedish Polar Research Secretariat, Abisko Scientific Research Station, Abisko, Sweden. [12]Department of Physical Geography, Stockholm University, Stockholm, Sweden. [13]Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden. [14]Department of Land, Environment, Agriculture and Forestry (TeSAF), University of Padua, Padua, Italy. [15]These authors contributed equally: Marco Carrer, Georg von Arx. ✉e-mail: jesper.bjoerklund@wsl.ch

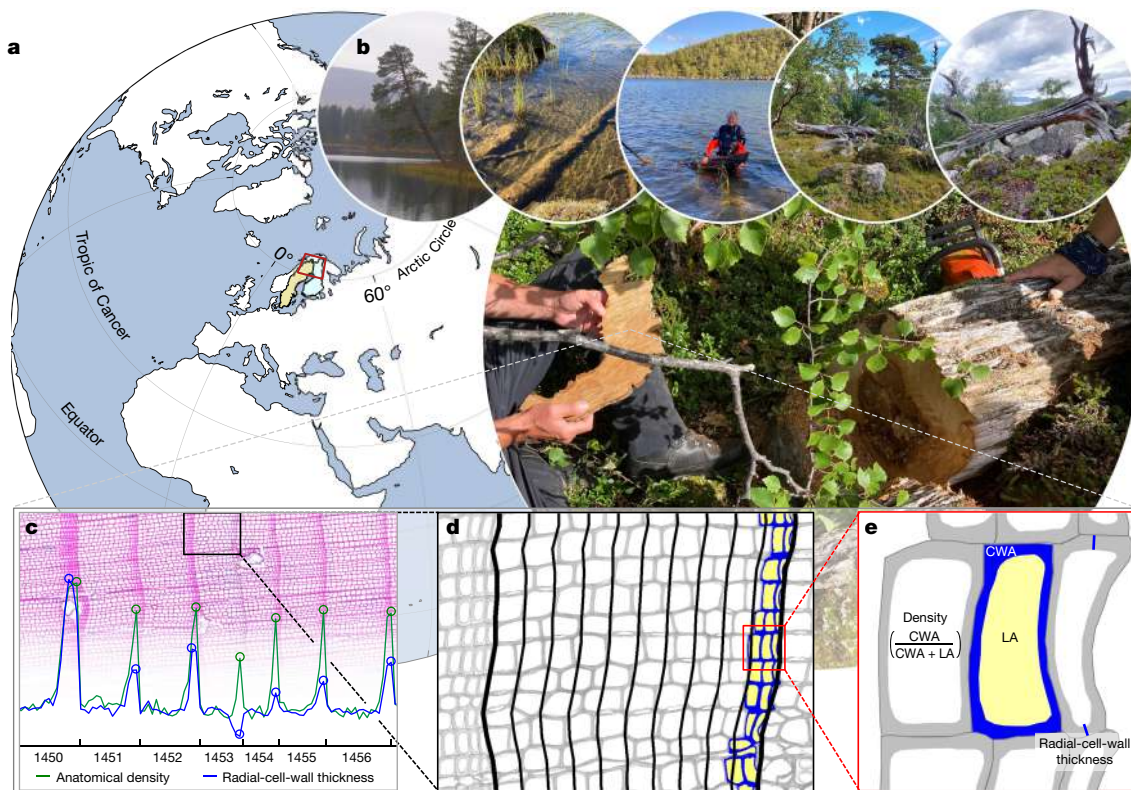**Fig. 1 | From Fennoscandian forests and lakes to QWA-based tree-ring data. a**, Map showing the general tree-sampling region (red polygon) where instrumental temperature data were obtained and for which climate model data were obtained (yellow, Sweden; blue, Finland). The map was created in the MATLAB computing environment using the M Map mapping package[67]. **b**, Wood sampling from living and dead trees. **c**, From the resulting high-resolution intra-ring profiles, annual maxima and minima, and annual latewood and earlywood values of radial-cell-wall thickness and anatomical density were extracted. **d**, Visualization of bands of aggregated QWA data (thick black lines, tree-ring boundaries; thin black lines, 30-μm bands within a tree ring), in which the coloured cells are included in the final band of aggregation in the latewood. **e**, Illustration of the various cell dimensions considered in this study. CWA, cell-wall area; LA, lumen area.

Here we develop a QWA-based 1,170-year-long reconstruction of warm-season temperatures (May, June, July and August; MJJA) using 188 living and dead Scots pine trees (*Pinus sylvestris*) from Fennoscandia (A-FEN). This region is disproportionally included in the millennium-length tree-ring datasets of the Arctic Circle and is thus important for understanding past climate in this sensitive area and elsewhere. Although these records constitute some of the most skilful examples of Common Era temperature reconstructions[20], climate variability and change during the MCA, LIA and CWP have repeatedly been revised and questioned[2,16,34–37]. We therefore subject the A-FEN to an array of tests, including comparisons with relevant reconstructions[1,2,9,10,37–39] and climate model simulations within the Coupled Model Intercomparison Project Phase 5 (CMIP5) and Palaeoclimate Modelling Intercomparison Project Phase 3 (PMIP3) (hereafter, CMIP5–PMIP3)[40–52]. Relying on the A-FEN, we provide a strong long-term background for recent warming by providing critical insights into climate dynamics in the region, which in turn will increase confidence in model projections[52].

Improved proxy calibration

The tree-ring anatomy data in A-FEN was randomly resampled into an ensemble of 100 datasets (each member with a maximum replication of 15 trees per year), explaining 71–79% of the variance in regional MJJA temperatures (Fig. 2a). The total A-FEN replication never falls below 20 trees per year and the average pairwise correlation between trees ($\bar{R}$)[53] is high, thus suggesting a stable reconstruction for 850–1849 CE (Fig. 2b). A-FEN tightly tracks the current increasing temperature trend, captures around 80% of both cold and warm extremes (Extended Data Fig. 1) and shows a similar first-order autocorrelation as do the MJJA temperatures (Extended Data Fig. 2). Results remain nearly the same regardless of whether idiosyncratic signal processing[54] (known as standardization or detrending in the tree-ring literature) is used (Extended Data Fig. 3). This means that the A-FEN dataset is essentially free of artificial climate-signal distortion that can occur when attempting to remove non-climate variability[15].

Although the climatic imprint in A-FEN is similar to the corresponding MXD-based reconstruction[9] (Fennoscandian X-ray-based data; X-FEN), A-FEN is significantly stronger (Fig. 2a) and consistently performs better across inter-annual and decadal timescales (Extended Data Fig. 2). A-FEN is potentially less susceptible to decreasing reconstruction quality than is X-FEN. This is because X-FEN $\bar{R}$ and replication calculated over the calibration period (1850–2019 CE) substantially exceed corresponding statistics for the reconstruction period (850–1849 CE) (Fig. 2b). Moreover, A-FEN is not appreciably biased by extremely narrow rings (Extended Data Fig. 4). For instance, there was widespread damage to pine-tree canopies across Fennoscandia[55] during the winter and spring of 1902–1903, leading to reduced photosynthetic capacity and persistent reductions in wood-cell divisions[56] and thus ring widths[57]. This episode is important because it represents a temporary collapse of climate being the dominant control over ring width. Nevertheless, A-FEN continues to track MJJA temperatures, showing that the collapse affected only cell divisions (ring width) but not cell dimensions. By contrast, X-FEN systematically underestimates summer temperatures during this period (Fig. 2a), suggesting that the X-ray measurements are affected by extremely narrow rings and thus contain more biological noise. X-FEN is also less independent of
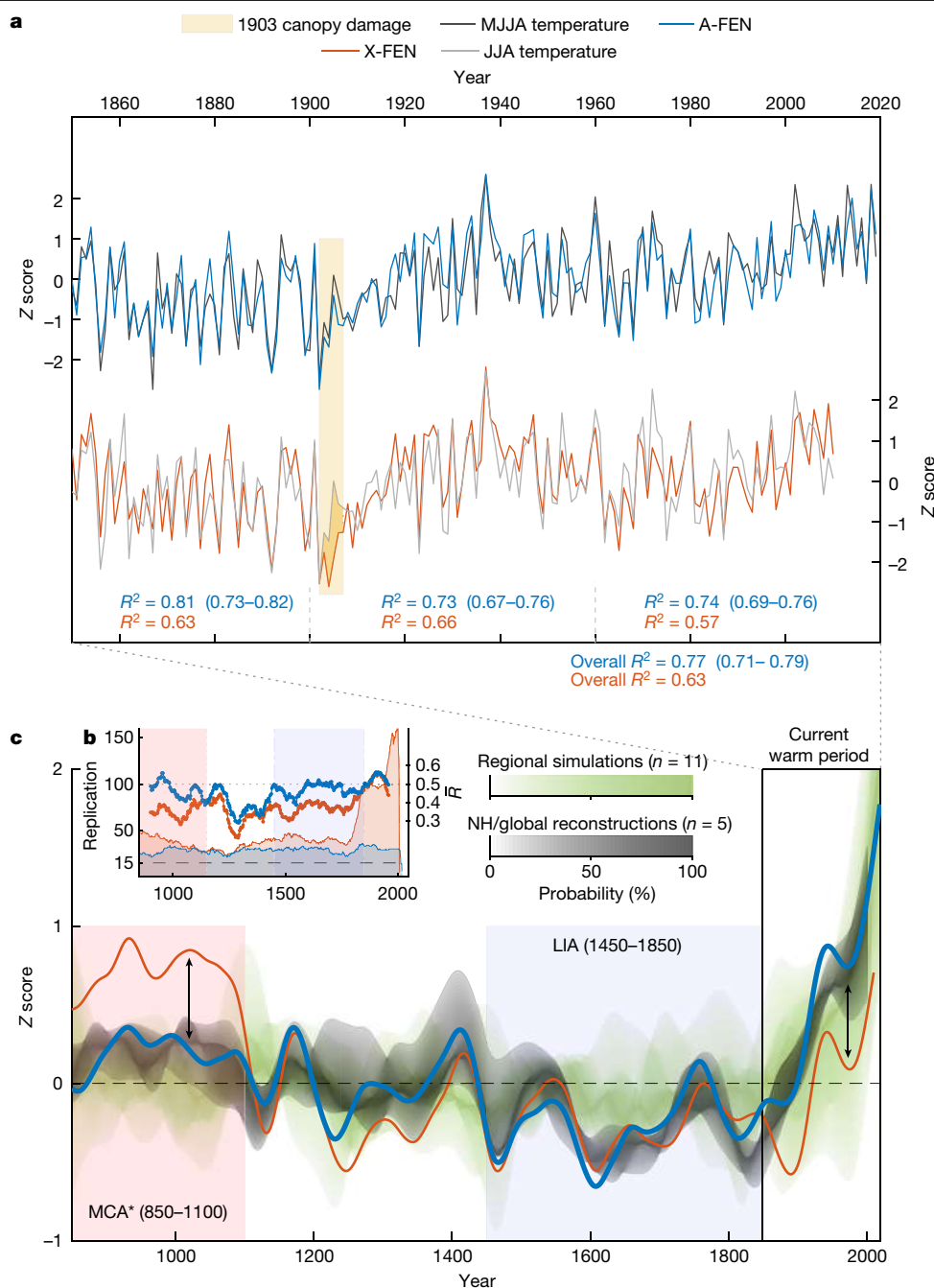
Fig. 2 | A-FEN has an excellent warm-season temperature reconstruction ability and reveals that the current climate is substantially warmer than that of the medieval period. a, A-FEN, produced in this study, calibrated using regional mean air MJJA temperatures[68] (R^2 ensemble range is shown in brackets, $\alpha = 0.05$) and corresponding results for the Fennoscandian X-ray-based dataset (X-FEN)[9] using JJA temperatures. The irregular winter and spring of 1902–1903 led to a massive dieback of yearly branch shoots in the region[55] (highlighted in yellow). For these years with extremely narrow rings, the X-ray technique struggles to measure high maximum-density values because of its comparatively lower effective measurement resolution[29] (Extended Data Fig. 4). R^2 values are

given overall and for the periods 1850–1899, 1900–1959, and 1960–2019 (blue) or 1960–2010 (red). **b**, Replication and pairwise inter-series correlation ($\overline{R}$) of A-FEN (blue) and X-FEN (red). **c**, Centennial-scale variations (Methods) compared between A-FEN, X-FEN, climate model simulations and Northern Hemisphere (NH) and global temperature reconstructions. The five large-scale reconstructions[1,9,10,38,39] and the 11 regionally extracted climate-model simulations[40–50] are represented by probabilistic percentile ranges (grey- and green-shaded areas, respectively). The vertical arrows highlight the overall discrepancies between the X-FEN dataset and the other data.

narrow rings on an average (Extended Data Figs. 4c and 5), suggesting latent measurement artefacts during the pre-instrumental era. In summary, A-FEN provides a more robust reconstruction of warm-season temperatures than X-FEN and the linear reconstruction model is well verified using three split-sample calibration and verification periods (Extended Data Table 1).

Revised long-term climate

A-FEN shows a slightly negative long-term trend between 850 and 1850, followed by a sharp temperature rise (Fig. 2). The peak medieval warm period in A-FEN and X-FEN occurs around 850–1100 (denoted as MCA*). The A-FEN dataset shows modest centennial-scale variability during
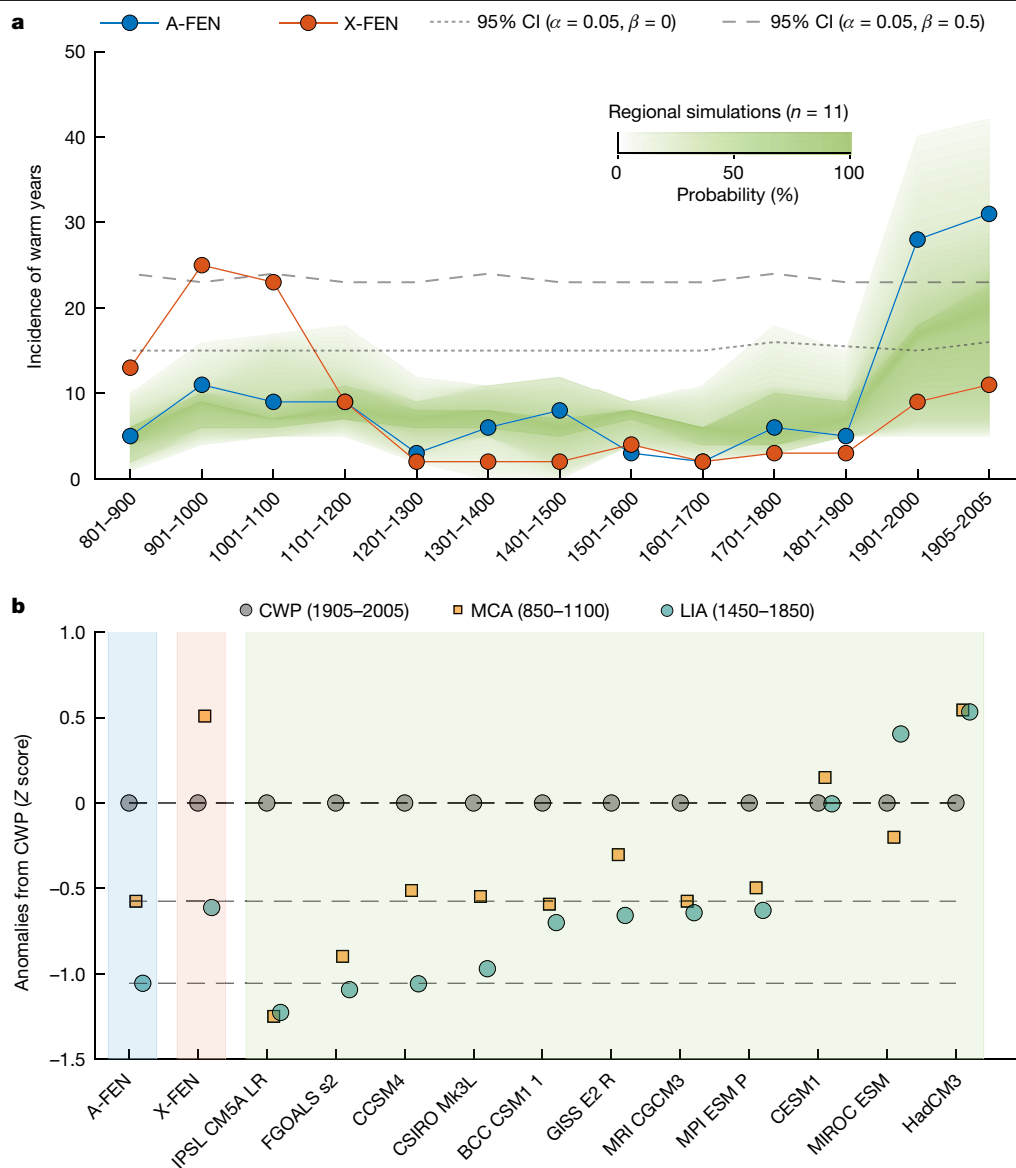
Fig. 3 | The climate models and A-FEN show unprecedented occurrences of warm extremes in the past century and comparable dynamics for medieval and current climates. a, A-FEN shows that warm years are significantly over-represented in the twentieth century. By contrast, X-FEN[9] shows significant over-representation of warm years during the MCA*. The simulations[40–50] (green shading) (probabilistic percentile range on the basis of respective incidences) largely agree with A-FEN suggesting a significant over-representation of warm summers in the most recent century. CI, confidence interval. **b**, Climate period mean values of A-FEN, X-FEN and the model ensemble during the medieval period (850–1100) and LIA (1450–1850), with the CWP adjusted to zero in all datasets, for ease of comparison.

MCA*, consistent with a lack of volcanic forcing and stable solar forcing[12]. The onset of the LIA is preceded by a relatively warm period around 1400, interrupted by rapid cooling[58], possibly driven by still-unidentified volcanism around 1453. Temperatures also briefly increase during the late eighteenth century. This temperature evolution is stable regardless of which 15 trees per year are resampled from A-FEN (Extended Data Fig. 6). It is remarkably similar to reconstructed Northern Hemisphere summer temperatures and even global annual temperatures (Fig. 2c). The multi-decadal fluctuations of A-FEN are similar to those of X-FEN, but X-FEN notably reconstructs MCA* as an exceptional warm period and CWP with only above-average temperatures (Fig. 2c). X-FEN is consistently included in hemispheric and global reconstructions (Fig. 2c); it will therefore also increase medieval and decrease current temperatures in these reconstructions. However, when several millennium-length Fennoscandian MXD datasets commonly used in larger-scale reconstructions[1,6,9,10,14,39,59–63] are analysed, the data show a

wide range of medieval temperature estimates (Extended Data Fig. 7a), thus underscoring the large uncertainty during this period. The inclusion of the new A-FEN reconstruction would probably reduce this uncertainty and provide a benchmark for a moderately warm medieval climate and a substantially warmer present.

Agreement between models and proxies

The spectral properties and first-order autocorrelation of A-FEN are similar to those of the model ensemble (Extended Data Fig. 8). A-FEN and most model ensemble members agree, particularly in the frequency of warm extremes with a significant over-representation of the 100 warmest summers in the previous century (Fig. 3a). The coldest years in the A-FEN and X-FEN datasets are often associated with volcanic activity—namely, the unidentified eruption in 1453, Eldgjá in 939, Huaynaputina in 1600 and Parker in 1640. However, two of the
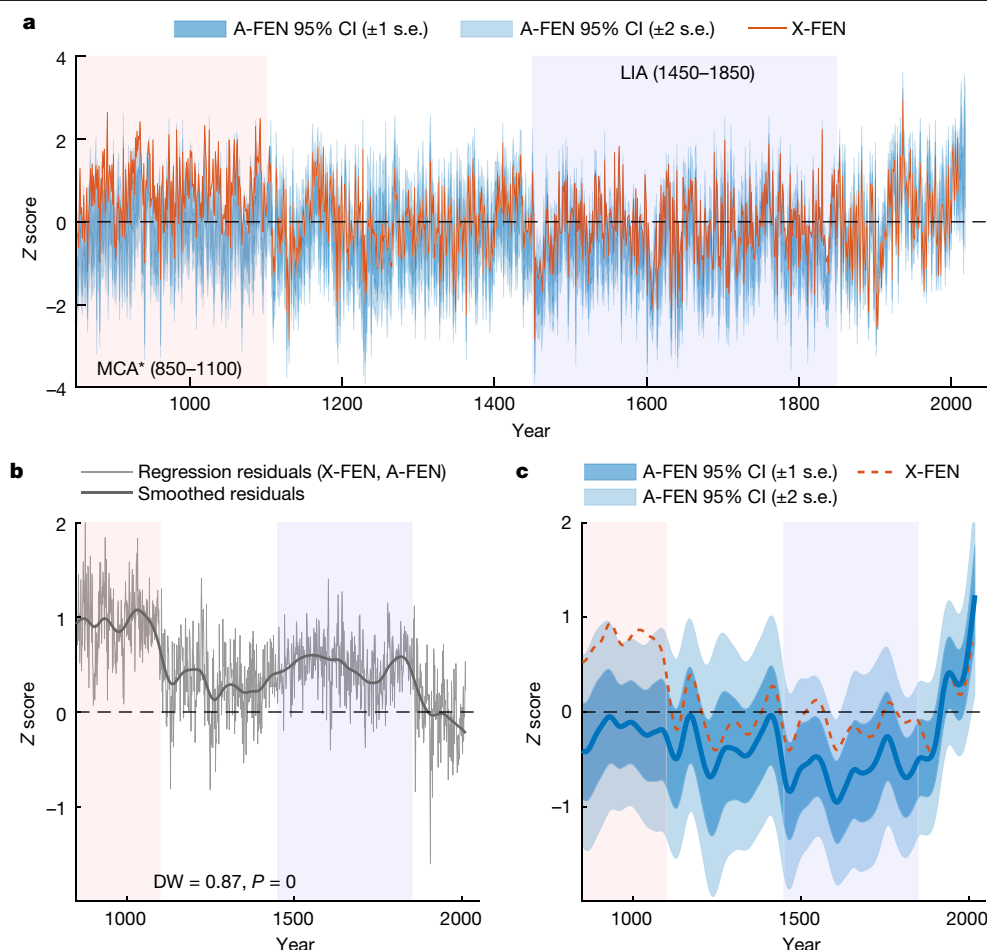
Fig. 4 | X-FEN shows the medieval climate as notably warmer than that estimated by A-FEN. a, Unfiltered reconstructions with 95% confidence intervals on the basis of ±1 and ±2 s.e. of the full set of trees (*n* = 188) included in the A-FEN dataset—that is, not based on the subsampled A-FEN ensemble members. The mean and variance of the reconstructions are adjusted to zero and unity over the 1850–2010 overlap. **b**, Regression residuals between X-FEN and A-FEN. The Durbin–Watson (DW) statistic show that the residuals are far from random, underscoring the systematic and notable differences between X-FEN and A-FEN. **c**, Same as in **a** but the data are low-pass filtered using 100-year splines.

largest volcanic eruptions of the past millennium—Samalas in 1257 and Tambora in 1815—do not seem to have caused detectable summer cooling in Fennoscandia (Extended Data Fig. 9). The climate models used in this study produce comparably distinct cooling patterns on an average, yet with important differences for specific events.

Model simulations of regional summer temperatures show overarching patterns that strongly resemble those of the A-FEN reconstruction, again highlighting an MCA* and CWP mismatch with X-FEN (Fig. 3b). At regional scales, such as in Fennoscandia, substantial internal climate variability—both implemented in models and imprinted in proxy observations—tends to diverge as these stochastic components are by definition not synchronized[64]. Therefore, it is particularly encouraging that the models and A-FEN agree, in terms of (1) statistical properties of the temperature history (for example, incidence of warm extremes, spectral characteristics), (2) responses to volcanic and anthropogenic forcing and (3) the amplitude and warming of the quietly forced climate variability around the medieval period. We also note that A-FEN is consistent with orbitally forced cooling over the past millennium[16], while characterizing the impact of anthropogenic forcing as much more influential.

Warm medieval period but hotter present

Unlike reconstructions relying on other tree-ring parameters, the A-FEN record realistically tracks the full breadth of recent climate extremes[21,23–25] and seems to provide a spectrally unbiased transformation from climate to proxy record[17–19,24,26] through tree biological processes. Although the debate over whether tree-ring proxies reflect climate change over millennial timescales[15,16,22] will continue, this study identifies a close agreement between model and proxy records over these timescales. The A-FEN reconstruction thus shows that the recent warming in Fennoscandia far exceeds the natural variability of the past millennium, in contrast to what was previously reported[3], and highlights the impact and occurrence of anthropogenic forcing on climate warming even at the regional scale. The difference compared with the X-FEN dataset is therefore highly consequential. The MCA* reconstructed using X-FEN is significantly warmer than the estimate of A-FEN for the same period (Fig. 4), which would de-emphasize the role of anthropogenic forcing on Fennoscandian climate variability. This large difference is difficult to explain with only technical improvements provided by the QWA approach and also referring to a slightly imperfect sample overlap (Methods) given the homogeneous sub-sampling results of A-FEN (Extended Data Fig. 6). However, the empirical evidence and theoretical advantages presented in this study strongly support the A-FEN representation of past temperature variability and are further strengthened by the compatibility with independent large-scale reconstructions and climate models.

Climate models integrate the current knowledge of system physics and are consistent with past climate dynamics at the event scale (for example, volcanic eruptions) as well as for broader trends (for example,

Article

an increase in current temperatures). The completely independent tree-ring anatomy observations instead draw strength from an exceptional fidelity to instrumental climate variations, are theoretically less susceptible to measurement artefacts, and exhibit a balanced dataset in terms of tree replication and correlations among trees outside the period of instrumental observations. Because the forced models, as well as A-FEN, are both unable to reproduce an exceptionally warm medieval period[3,4], their alignment thus provides compelling evidence for a more moderate MCA* than previously estimated[2,3,16,34,35,37,61]. The proxy–model alignment may have benefited further from the reduced biological noise in A-FEN, which more closely resembles the noise-free structure of model simulations[63].

Provided that revised temperature characterization and proxy capabilities using tree-ring anatomy become standard, the underestimates in temperatures observed in recent years, also known as divergence[65], seen in large-scale reconstructions[9,10,66] and perhaps other regional Fennoscandian tree-ring data (Extended Data Fig. 7), could possibly be mitigated. Furthermore, the uncertainties before 1400 could probably be reduced and provide a valuable new constraint on quietly forced natural climate variability. In conjunction with model simulations, reconstructions of current and medieval climates using tree-ring anatomy could therefore improve benchmarks of natural and anthropogenically forced climate variability and climate models and their projections.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-023-06176-4.

1. Neukom, R. et al. Consistent multidecadal variability in global temperature reconstructions and simulations over the Common Era. *Nat. Geosci.* **12**, 643–649 (2019).
2. Esper, J., Düthorn, E., Krusic, P. J., Timonen, M. & Büntgen, U. Northern European summer temperature variations over the Common Era from integrated tree-ring density records. *J. Quat. Sci.* **29**, 487–494 (2014).
3. Luterbacher, J. et al. European summer temperatures since Roman times. *Environ. Res. Lett.* **11**, 024001 (2016).
4. Fernández-Donado, L. et al. Large-scale temperature response to external forcing in simulations and reconstructions of the last millennium. *Clim. Past* **9**, 393–421 (2013).
5. Masson-Delmotte, V. et al. (eds) *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge Univ. Press, 2021).
6. Neukom, R., Steiger, N., Gómez-Navarro, J. J., Wang, J. & Werner, J. P. No evidence for globally coherent warm and cold periods over the preindustrial Common Era. *Nature* **571**, 550–554 (2019).
7. Frank, D., Esper, J., Zorita, E. & Wilson, R. A noodle, hockey stick, and spaghetti plate: a perspective on high-resolution paleoclimatology. *Wiley Interdiscip. Rev. Clim. Change* **1**, 507–516 (2010).
8. Anchukaitis, K. J. & Smerdon, J. E. Progress and uncertainties in global and hemispheric temperature reconstructions of the Common Era. *Quat. Sci. Rev.* **286**, 107537 (2022).
9. Wilson, R. et al. Last millennium northern hemisphere summer temperatures from tree rings: Part I: the long term context. *Quat. Sci. Rev.* **134**, 1–18 (2016).
10. Schneider, L. et al. Revising midlatitude summer temperatures back to A.D. 600 based on a wood density network. *Geophys. Res. Lett.* **42**, 4556–4562 (2015).
11. Zhao, B. et al. Prolonged drying trend coincident with the demise of Norse settlement in southern Greenland. *Sci. Adv.* **8**, eabm4346 (2022).
12. Bradley, R. S., Wanner, H. & Diaz, H. F. The Medieval Quiet Period. *Holocene* **26**, 990–993 (2016).
13. Knutti, R. & Hegerl, G. C. The equilibrium sensitivity of the Earth's temperature to radiation changes. *Nat. Geosci.* **1**, 735–743 (2008).
14. PAGES2kConsortium. A global multiproxy database for temperature reconstructions of the Common Era. *Sci. Data* **4**, 170088 (2017).
15. Cook, E. R., Briffa, K. R., Meko, D. M., Graybill, D. A. & Funkhouser, G. The 'segment length curse' in long tree-ring chronology development for palaeoclimatic studies. *Holocene* **5**, 229–237 (1995).
16. Esper, J. et al. Orbital forcing of tree-ring data. *Nat. Clim. Change* **2**, 862–866 (2012).
17. Franke, J., Frank, D., Raible, C. C., Esper, J. & Brönnimann, S. Spectral biases in tree-ring climate proxies. *Nat. Clim. Change* **3**, 360–364 (2013).
18. Esper, J., Schneider, L., Smerdon, J. E., Schöne, B. R. & Büntgen, U. Signals and memory in tree-ring width and density data. *Dendrochronologia* **35**, 62–70 (2015).
19. Zhang, H. et al. Modified climate with long term memory in tree ring proxies. *Environ. Res. Lett.* **10**, 084020 (2015).
20. Esper, J. et al. Ranking of tree-ring based temperature reconstructions of the past millennium. *Quat. Sci. Rev.* **145**, 134–151 (2016).
21. McCarroll, D., Young, G. H. & Loader, N. J. Measuring the skill of variance-scaled climate reconstructions and a test for the capture of extremes. *Holocene* **25**, 618–626 (2015).
22. Büntgen, U. Scrutinizing tree-ring parameters for Holocene climate reconstructions. *Wiley Interdiscip. Rev. Clim. Change* **13**, e778 (2022).
23. Battipaglia, G. et al. Five centuries of Central European temperature extremes reconstructed from tree-ring density and documentary evidence. *Glob. Planet. Change* **72**, 182–191 (2010).
24. Von Storch, H. et al. Reconstructing past climate from noisy data. *Science* **306**, 679–682 (2004).
25. Belmecheri, S., Babst, F., Wahl, E. R., Stahle, D. W. & Trouet, V. Multi-century evaluation of Sierra Nevada snowpack. *Nat. Clim. Change* **6**, 2–3 (2016).
26. Lücke, L. J., Hegerl, G. C., Schurer, A. P. & Wilson, R. Effects of memory biases on variability of temperature reconstructions. *J. Clim.* **32**, 8713–8731 (2019).
27. von Arx, G., Crivellaro, A., Prendin, A. L., Čufar, K. & Carrer, M. Quantitative wood anatomy—practical guidelines. *Front. Plant Sci.* **7**, 781 (2016).
28. Prendin, A. L. et al. New research perspectives from a novel approach to quantify tracheid wall thickness. *Tree Physiol.* **37**, 976–983 (2017).
29. Björklund, J. et al. Scientific merits and analytical challenges of tree-ring densitometry. *Rev. Geophys.* **57**, 1224–1264 (2019).
30. Björklund, J., Seftigen, K., Fonti, P., Nievergelt, D. & von Arx, G. Dendroclimatic potential of dendroanatomy in temperature-sensitive *Pinus sylvestris*. *Dendrochronologia* **60**, 125673 (2020).
31. Lopez-Saez, J. et al. Tree-ring anatomy of *Pinus cembra* trees opens new avenues for climate reconstructions in the European Alps. *Sci. Total Envir.* **855**, 158605 (2023).
32. Seftigen, K. et al. Prospects for dendroanatomy in paleoclimatology – a case study on *Picea engelmannii* from the Canadian Rockies. *Clim. Past* **18**, 1151–1168 (2022).
33. Allen, K. J., Nichols, S. C., Evans, R. & Baker, P. J. Characteristics of a multi-species conifer network of wood properties chronologies from Southern Australia. *Dendrochronologia* **76**, 125997 (2022).
34. Melvin, T. M., Grudd, H. & Briffa, K. R. Potential bias in 'updating' tree-ring chronologies using regional curve standardisation: re-processing 1500 years of Torneträsk density and ring-width data. *Holocene* **23**, 364–373 (2013).
35. Linderholm, H. W. & Gunnarson, B. E. Were medieval warm-season temperatures in Jämtland, central Scandinavian Mountains, lower than previously estimated? *Dendrochronologia* **57**, 125607 (2019).
36. Grudd, H. Torneträsk tree-ring width and density AD 500–2004: a test of climatic sensitivity and a new 1500-year reconstruction of north Fennoscandian summers. *Clim. Dyn.* **31**, 843–857 (2008).
37. Matskovsky, V. & Helama, S. Testing long-term summer temperature reconstruction based on maximum density chronologies obtained by reanalysis of tree-ring data sets from northernmost Sweden and Finland. *Clim. Past* **10**, 1473–1487 (2014).
38. Büntgen, U. et al. Prominent role of volcanism in Common Era climate variability and human history. *Dendrochronologia* **64**, 125757 (2020).
39. Guillet, S. et al. Climate response to the Samalas volcanic eruption in 1257 revealed by proxy records. *Nat. Geosci.* **10**, 123–128 (2017).
40. Wu, T. et al. An overview of BCC climate system model development and application for climate change studies. *J. Meteorol. Res.* **28**, 34–56 (2014).
41. Landrum, L. et al. Last millennium climate and its variability in CCSM4. *J. Clim.* **26**, 1085–1111 (2013).
42. Dufresne, J.-L. et al. Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. *Clim. Dyn.* **40**, 2123–2165 (2013).
43. Yukimoto, S. et al. A new global climate model of the Meteorological Research Institute: MRI-CGCM3 – model description and basic performance–. *J. Meteorol. Soc. Japan. Ser. II* **90A**, 23–64 (2012).
44. Bao, Q. et al. The flexible global ocean-atmosphere-land system model, spectral version 2: FGOALS-s2. *Adv. Atmos. Sci.* **30**, 561–576 (2013).
45. Phipps, S. et al. The CSIRO Mk3L climate system model version 1.0 – part 1: description and evaluation. *Geosci. Model Dev.* **4**, 483–509 (2011).
46. Miller, R. L. et al. CMIP5 historical simulations (1850–2012) with GISS ModelE2. *J. Adv. Model. Earth Syst.* **6**, 441–478 (2014).
47. Giorgetta, M. A. et al. Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5. *J. Adv. Model. Earth Syst.* **5**, 572–597 (2013).
48. Marsh, D. R. et al. Climate change from 1850 to 2005 simulated in CESM1(WACCM). *J. Clim.* **26**, 7372–7391 (2013).
49. Watanabe, S. et al. MIROC-ESM 2010: model description and basic results of CMIP5-20c3m experiments. *Geosci. Model Dev.* **4**, 845–872 (2011).
50. Johns, T. C. et al. Anthropogenic climate change for 1860 to 2100 simulated with the HadCM3 model under updated emissions scenarios. *Clim. Dyn.* **20**, 583–612 (2003).
51. Taylor, K. E., Stouffer, R. J. & Meehl, G. A. An overview of CMIP5 and the experiment design. *Bull. Am. Meteorol. Soc.* **93**, 485–498 (2012).
52. Braconnot, P. et al. Evaluation of climate models using palaeoclimatic data. *Nat. Clim. Change* **2**, 417–424 (2012).
53. Wigley, T. M. L., Briffa, K. R. & Jones, P. D. On the average value of correlated time series, with applications in dendroclimatology and hydrometeorology. *J. Appl. Meteorol. Climatol.* **23**, 201–213 (1984).
54. Helama, S., Melvin, T. M. & Briffa, K. R. Regional curve standardization: state of the art. *Holocene* **27**, 172–177 (2017).
55. Andersson, G. *Om talltorkan i öfra Sverige våren 1903* (Statens skogsförsöksanstalt, 1905).
56. Pallardy, S. G. *Physiology of Woody Plants* 3rd edn (Academic, 2008).
57. Vaganov, E. A., Hughes, M. K. & Shashkin, A. V. *Growth Dynamics of Conifer Tree Rings: Images of Past and Future Eenvironments* Vol. 183 (Springer Science & Business Media, 2006).
58. Abbott, P. M. et al. Cryptotephra from the Icelandic Veiðivötn 1477 CE eruption in a Greenland ice core: confirming the dating of volcanic events in the 1450s CE and assessing the eruption's climatic impact. *Clim. Past* **17**, 565–585 (2021).

59. Stoffel, M. et al. Estimates of volcanic-induced cooling in the Northern Hemisphere over the past 1,500 years. *Nat. Geosci.* **8**, 784–788 (2015).

60. Anchukaitis, K. J. et al. Last millennium Northern Hemisphere summer temperatures from tree rings: part II, spatially resolved reconstructions. *Quat. Sci. Rev.* **163**, 1–22 (2017).

61. McCarroll, D. et al. A 1200-year multiproxy record of tree growth and summer temperature at the northern pine forest limit of Europe. *Holocene* **23**, 471–484 (2013).

62. Neukom, R. et al. Inter-hemispheric temperature variability over the past millennium. *Nat. Clim. Change* **4**, 362–367 (2014).

63. PAGES2k-PMIP3 group. Continental-scale temperature variability in PMIP3 simulations and PAGES 2k regional temperature reconstructions over the past millennium. *Clim. Past* **11**, 1673–1699 (2015).

64. Lorenz, E. N. Deterministic nonperiodic flow. *J. Atmos. Sci.* **20**, 130–141 (1963).

65. D'Arrigo, R., Wilson, R., Liepert, B. & Cherubini, P. On the 'divergence problem' in northern forests: a review of the tree-ring evidence and possible causes. *Glob. Planet. Change* **60**, 289–305 (2008).

66. Büntgen, U. et al. The influence of decision-making in tree ring-based climate reconstructions. *Nat. Commun.* **12**, 3411 (2021).

67. Pawlowicz, R. M_Map: a mapping package for MATLAB, MATLAB package v.1.4m. https://www.eoas.ubc.ca/~rich/map.html (2020).

68. Morice, C. P. et al. An updated assessment of near-surface temperature change from 1850: The HadCRUT5 data set. *J. Geophys. Res. Atmos.* **126**, e2019JD032361 (2021).

Article

Methods

Tree-ring material

The trees used in this study were subsampled from material from two distinct but proximal temperature reconstructions on the basis of X-ray densitometry (MXD); they are known to palaeoclimatologists as the N-scan[16,20] (Finland) and the Torneträsk[34,36,69–71] (Sweden) MXD reconstructions. Data from these sites have been combined previously[2,37]. We used the arithmetic average of previously published data[2,37], developed and used in a previously published Northern Hemisphere temperature reconstruction[9] (X-FEN), for comparisons with the QWA data (A-FEN). The sample collection of N-scan was updated in 2011 and 2014. The living-tree material from 2014 has previously been used in publications exploring the performance of wood anatomical data compared with other microdensitometric techniques, including X-ray densitometry[29,30] (N-scan DC). The N-scan DC dataset, analysed using wood anatomical techniques in 2015–2017, consisted originally of 29 trees (two samples per tree). Here we randomly subsampled 10 trees (one sample per tree) to harmonize sample replication through time and to match replication with the Torneträsk dataset. The N-scan DC dataset was composited with data from six more living trees (sampling date 2011) as well as the original subfossil sample collection of N-scan (dead trees preserved in mountain lake sediments)[16]. The sample collection from Torneträsk was updated by sampling 14 living trees in May 2020; this material was composited not only with the dry deadwood material that was previously analysed using X-ray densitometry[69] but also with some recently collected dry deadwood samples previously analysed for only tree-ring width[34,36]. The new wood anatomical versions of N-scan and Torneträsk comprise measurements from 99 trees (671–2013 CE) and 89 trees (545–2019 CE), respectively, for a total of 188 trees. The combined dataset contains around 10 trees per year from each collection, amounting to a total sample replication of 20–30 trees per year for the period 850–2019 CE (Fig. 2). The latest versions of the Torneträsk and N-scan MXD datasets consist of 124 (441–2010 CE) and 294 trees (138 BCE to 2006 CE), respectively, for a total of 418 trees.

QWA data production

The wood anatomical analyses were made following state-of-the-art techniques[27], but with adjustments to protocols for more recently analysed subset members.

The thin sections from the N-scan DC samples were cut to a thickness of approximately 15 μm using a sledge microtome[72] equipped with Feather N35 disposable blades (Feather Safety Razor). Images of the sections were serially captured with a camera (Canon EOS 650D, Canon) mounted on a microscope (Olympus BX41, Olympus) at a resolution of 2.36 pixels per μm. Composite images of the sections were created using PTGui (New House Internet Services) stitching software. The second N-scan component was produced at the University of Padua (N-scan Padua) and constituted half of the full-length N-scan dataset (including six living trees). These samples were cut to thin sections of about 15-μm thickness with a Leica RM 2255 rotary microtome (Leica), digitized using a camera (Nikon Eclipse 80i, Nikon) connected to a light microscope and then stitched using PTGui (v.8.3.3, New House Internet Services). The remaining N-scan components, comprising half of the subfossil material, were developed using the latest protocol (as of 2022) at WSL (N-scan WSL). Each wood sample was cut into about 2.5-cm-long pieces and embedded in paraffin using an embedding station (Leica TP1020 combined with Leica HistoCore Arcadia, Leica). Thin sections of 12-μm thickness were produced using a Leica RM2245 rotary microtome (Leica) and images were captured with a Zeiss Axioscan.Z1 slide scanner (Carl Zeiss) at a resolution of 2.27 pixels per μm.

The entire Torneträsk dataset was produced using the latest WSL protocol. Regardless of cutting and image-capturing protocol, each thin section was double-stained with safranin and astra blue to increase contrast, then permanently fixed with Eukitt, Canada balsam or Euparal. The software ROXAS (v.3.1) was used to automatically detect anatomical structures of all tracheid cells[28,73] (75–100 radial files per ring) as well as annual-ring borders (Fig. 1). The latewood in each ring was defined using a Mork's index value greater than 1. That is, those tangential cell bands in which twice the double cell-wall thickness was larger than the lumen diameter across the radial direction on average were defined as latewood[74]. The total QWA dataset comprises about 50 million tracheid-cell measurements. A detailed protocol for anatomical data extraction has been published previously[30].

Chronology development

We used the delta radial-cell-wall thickness parameter as the single parameter for our A-FEN reconstruction. This parameter represents the difference between radial-cell-wall thickness in the latewood and earlywood, specifically maximum radial-cell-wall thickness − minimum radial-cell-wall thickness. This parameter was chosen among a range of parameters, such as maximum, latewood and delta[75] radial-cell-wall thickness, maximum and delta anatomical density[29] as well as tangential latewood lumen diameter and latewood width. These parameters all contain pronounced and almost identical temperature signals, in which the delta radial-cell-wall thickness has the best performance overall (Extended Data Fig. 2). All considered parameters were subjected to standardization and detrending using the widely used regional curve standardization (RCS) approach[71]. This technique is used to preserve as much long-term climate variability as the included samples permit[15]. The methodology has been fine-tuned over many years, but without changing considerably[54], at least when using multi-generational and replication-balanced datasets through time. Therefore, we used the conventional single-RCS approach (one regional curve for all parameter samples), following the principle of parsimony as follows. (1) Each collection was subsampled to a sample replication of fewer than 12 through time by randomly removing trees from potentially oversampled periods and then aligning all the chosen trees by the pith year (pith offsets were estimated and used). (2) One regional average curve was calculated as the arithmetic mean of all included trees per pith year. (3) Curves were then fitted with a series of customized functions to estimate smoothed growth curves, that is, regional curves. (3a) A cubic smoothing spline with a frequency cut-off of 50% at 50 years was fitted to the regional average. (3b) The cambial age at which the sample replication of the age-aligned samples was fewer than 10 trees was determined. (3c) From this age and 20 years before that, a mean value was calculated and appended to the 50-year spline from the replication threshold until the final year of the oldest tree. (3d) At around 50 years of cambial age, the trees have passed the juvenile phase and entered the mature phase, which generally displays a modest decline in growth values. At an age of 50 years and older, a cubic smoothing spline with a frequency cut-off of 50% at 500 years was fitted and appended to the 1- to 50-year segment of the regional curve function of step 3a. (3e) Finally, a cubic smoothing spline with a frequency cut-off of 50% at 50 years was fitted to the whole function of step 3d. Steps 1–3e were repeated 100 times and a mean of all 100 regional curves was calculated to provide the final and single regional curve per parameter chronology. This regional curve was subtracted from the complete set of measurements to create dimensionless tree-ring indices. Ratios were used for ring width and latewood widths.

We also developed datasets without idiosyncratic signal processing (that is, no RCS applied), to understand how much the RCS changed the obtained results. Although the detrended data show very little difference compared with non-detrended data, we use the detrended version because of obvious juvenile trends (from about 0 to 20 years) in the raw anatomical data parameters.

Because we used datasets developed under a stepwise progression of measurement protocols (relevant for only N-scan), we examined the different components for discrepancies in means and variances,

which are critical for the appropriate implementation of RCS. We found inconsistencies and accordingly harmonized the data on the basis of a concurrent overlap between each N-scan component of more than five samples. Therefore, we first calculated the mean and variance of the N-scan Padua component during the replication overlap of more than 5 samples with the N-scan DC component (only living trees). We then forced the N-scan DC component to take on the mean and variance of the N-scan Padua component. Thereafter, we repeated the operation with the N-scan WSL component, again using N-scan Padua as target. These operations were done before RCS because of the relatively low, and sometimes uneven, temporal distribution in replication of each component.

Torneträsk and N-scan were processed using separate regional curves and subsequently their means and variances were adjusted to zero and unity on the 850–2005 CE overlap and all indices were pooled into the A-FEN dataset. After this, the A-FEN dataset was subsampled into an ensemble of 100 datasets by means of random sampling (without replacement) aimed at a consistent replication of 15 trees per year for the full length of each dataset. This operation was used to show the stability of temperature signals and lower frequency variations through time as a function of changes in included trees. Moreover, creating distributions of datasets allowed us to assess whether the obtained results for A-FEN and X-FEN were significantly distinct from each other in any respect.

Climate-model ensemble

Simulations from 11 models (BCC-CSM1-1[40], CCSM4[41], IPSL-CM5A-LR[42], MRI-CGCM3[43], FGOALS-s2[44], CSIRO Mk3L[45], GISS E2 R[46], MPI ESM P[47], CESM1[48], MIROC ESM[49] and HadCM3[50]) contributing to the CMIP5–PMIP3 models[51,76] were used in the model–proxy comparison. The choice of models was based on the criteria of having essentially complete monthly surface-air temperature variables covering the past millennium (850–1849 CE) and the historical period (1850–2005 CE). The past1000 simulations were forced using reconstructed solar, volcanic, greenhouse-gas and aerosol forcings and partly using land-use changes, whereas the historical simulations included natural and anthropogenic forcings[51,76]. Monthly values of MJJA temperatures were extracted from the land area bound by latitudes 65–70° N and longitudes 15–30° E. A simple monthly arithmetic mean was calculated to directly correspond to the best-fit instrumental target season for the proxy data. Most analyses executed for the proxy data were also performed for the models after they were normalized to a mean of zero and a standard deviation of unity over the period 850–2005 CE.

Proxy calibration and temperature signals

The association with temperature was determined using Pearson correlations with monthly data from HadCRUT5 (ref. 68) (5° gridded monthly dataset, latitude 65–70° N, longitude 15–30° E). Coefficients of determination (R^2) were calculated for three non-overlapping periods (1850–1899, 1900–1959 and 1960 to the final year of each dataset) to check the stability in the best-fit target season MJJA for A-FEN and JJA for X-FEN. The A-FEN reconstruction was further validated using the reduction of error and coefficient of efficiency statistics[77]. The simple nonparametric test referred to as extreme value capture[21] was used to determine how well each datasets reflect both cold and warm extremes of the instrumental dataset. That is, the number of the correctly reconstructed extreme years captured beyond the thresholds defined by the upper and lower 10% of the measured temperature data (Extended Data Fig. 1). The calibration results were calculated for both RCS-processed analyses (Fig. 2 and Extended Data Table 1) and non-detrended datasets (Extended Data Fig. 3). High-pass-filtered datasets (using cubic smoothing splines[78] with a 50% frequency response cut-off at 40 years), untreated and low-pass-filtered datasets (using cubic smoothing splines[78] with a 50% frequency response cut-off at 5 and 10 years) were also used for the monthly temperature correlations.

All correlations were adjusted for loss of degrees of freedom depending on the autocorrelation in each series[79]. The reconstructions and models were compared by scaling each dataset with a mean of zero and standard deviation of unity over the period 850–2005 CE. The large-scale reconstructions[1,9,10,38,39] were scaled after smoothing to the mean and variance of the smoothed A-FEN to facilitate comparisons.

The pine-tree canopy damage

In the winter and spring of 1902–1903, a trauma outside of the growing season afflicted many pine trees in north and central Sweden, described by G. Andersson at the Governmental Forest Research Institution in 1905 (ref. 55). Several references in the study show that the unusual event was also noticed in Norway and as far east as the Kola Peninsula. In brief, the growing season of 1902 was cold and wet and the radial growth and cell dimensions underdeveloped. In mid-September 1902, heavy wet snow fell and accumulated onto the trees. A few days later, temperatures drastically dropped to nearly −10 °C, leading to damaged or broken tree branches. In the spring of 1903, unusual warmth early in the season was followed by icy cold and dry winds from the east, leaving pine shoots stunted. This arguably led to reduced canopy photosynthetic capacity[56], with persistent reductions in wood cell divisions, resulting in extremely narrow ring widths for up to 6 years. The importance of this episode relates to the temporary collapse of ring-width expansion controlled foremost by climate, whereas individual cell dimensions seem to be largely unaffected. If cell dimensions continue to form under dominant control of climate, this would theoretically imply that MXD, which is also predominantly a function of cell dimensions[29], would be indifferent to the canopy damage as well. Instead, MXD values seem suppressed for several years, implying that MXD measurements may be affected by the extremely narrow rings characterizing this episode. For interested readers not fluent in Swedish, we share some excerpts loosely translated to English from the original reference[55] (translation by J.B.).

Anyone travelling to the North in the summer of 1903 could not fail to observe, especially in younger pine stands, that individual shoots, or the entire upper shoot systems, had withered away. Here and there, especially in certain areas, the stands were rather badly ravaged and their appearance attracted the attention of both foresters and to a large extent laymen as well.

A condition that many forest district administrators paid attention to was the unusual development of the 1902 year's yearly shoot. On the one hand the shoots were abnormally long and on the other, poorly developed. Sometimes the length was more than double that of the shoots of the preceding summer of 1901, but the pine needles were short and lay snug onto the yearly shoots, different from what is usually the case in fully mature shoots late in the autumn.

Several opinions were raised that the direct cause of the pine shoot damage was the period of warm temperatures followed by icy cold that occurred in April 1903. April experienced summer warmth that was believed to have triggered the sap flow in the trees. Thereafter followed a grim late winter with large snow masses. Many acquired the opinion that the pine shoot damage was caused by the too early commencement of processes in the tissues in the spring of 1903, and that they later suffered dehydration from the physiologically cold and dry periods. These people also emphasize the unusually hardened ground frost and the long time it took to thaw, as an exacerbating circumstance for the trees to absorb water from the ground.

Microscopic investigations of 1902 year's tree rings revealed that they were narrower than corresponding rings of 1901 and that the cells in the wood were smaller and the cell walls were thinner, with a more pronounced resin content in the wood cells. For example, their mean tangential width was 0.0138 mm in 1902 and 0.0179 mm in 1901.

It is interesting that it took about 100 years after these observations to discover that the starting point of the pine-tree canopy damage was marked by so-called blue rings[80] in 1902. In double-stained

Article

microsections of wood, using safranin and astra blue, the cell walls appear blue because of markedly reduced lignification of the cell walls. A normally lignified cell wall should be coloured red because of the safranin. The previous observations[55] testify that the conditions for the trees in this period were unusually harsh and perhaps the combination of the underdeveloped wood formed during the very cold year of 1902 was more vulnerable to the late-season cold spell in September of that year, and therefore blue rings formed.

Full-length dataset assessments

The common signal among trees for each specific tree-ring parameter was established using the $\overline{R}$ statistic[53] (the mean pairwise correlation of all trees overlapping in time) using a moving window of 100 years with a lag of 10 years across the full period. A high $\overline{R}$ value indicates that the trees in a tree-ring chronology respond similarly to environmental forcing. Using the moving-window $\overline{R}$ and tree replication through time, the statistic expressed population signal[53] was calculated to gauge the robustness of each dataset (both A-FEN and X-FEN had expressed population signals above 0.85 for the full 850 to present periods). The $\overline{R}$ and replication for X-FEN were done using the series included in ref. 37. The uncertainty of A-FEN was calculated using ± 1 and ± 2 s.e. (standard deviation of the annual spread of values divided by the square root of the annual tree replication). The difference between X-FEN and A-FEN was also evaluated by examining their regression residuals using the Durbin–Watson statistic. We assessed the spectral characteristics of the proxy and model data using Thomson's multitaper method[81], applied with four windows, to transform the records from the time to the frequency domain within the MATLAB function pmtmPH.m (ref. 82). We used a superposed epoch analysis[83] (SEA) to evaluate the degree of cooling associated with volcanic aerosol forcing in the models and the corresponding cooling in the proxies. The used SEA variant was inspired by ref. 84 and the computing was done in MATLAB using the function developed in ref. 85. We compiled event lists from ref. 86, consisting of the 10 and the 30 largest (on the basis of sulfate aerosol injection) Northern Hemisphere events. We used only ref. 86 as the basis for the event lists because most models in our ensemble were forced according to ref. 86. We used a model ensemble mean in the SEA to explore the degree of volcanic cooling in the models. The incidence of warm extremes was quantified by first extracting the 100 warmest years from the common overlap 850–2005 CE of each simulation and reconstruction. The extracted extremes were then summed per century and compared with confidence intervals on the basis of the number of extremes happening by chance. Confidence intervals were calculated by generating 1,000 synthetic time series over the interval 850–2005 CE using random white noise processes and coloured noise processes with a β coefficient of 0.5 (β coefficients of white noise = 0 and pink noise = 1 for reference) and then taking the 97.5th percentile as the confidence limit.

Data availability

The A-FEN reconstruction is available at the National Centers for Environmental Information on the NOAA homepage (https://www.ncei.noaa.gov/access/paleo-search/?dataTypeId=3). The data used in the reconstruction are available at the NOAA International Tree Ring Data Bank. The data used to perform our analysis as well as our results are uploaded to Zenodo and are freely accessible using the following link: https://doi.org/10.5281/zenodo.7993298. Source data are provided with this paper.

Code availability

The code that supports the findings of this study is available alongside source data on the Zenodo repository and can be accessed using the following link: https://doi.org/10.5281/zenodo.7993298.

69. Schweingruber, F. H., Bartholin, T., Schär, E. & Briffa, K. R. Radiodensitometric-dendroclimatological conifer chronologies from Lapland (Scandinavia) and the Alps (Switzerland). *Boreas* **17**, 559–566 (1988).
70. Briffa, K. R. et al. A 1,400-year tree-ring record of summer temperatures in Fennoscandia. *Nature* **346**, 434–439 (1990).
71. Briffa, K. R. et al. Fennoscandian summers from AD 500: temperature changes on short and long timescales. *Clim. Dyn.* **7**, 111–119 (1992).
72. Gärtner, H., Lucchinetti, S. & Schweingruber, F. A new sledge microtome to combine wood anatomy and tree-ring ecology. *IAWA J.* **36**, 452–459 (2015).
73. von Arx, G. & Carrer, M. ROXAS – a new tool to build centuries-long tracheid-lumen chronologies in conifers. *Dendrochronologia* **32**, 290–293 (2014).
74. Denne, M. P. Definition of latewood according to Mork (1928). *IAWA J.* **10**, 59–62 (1989).
75. Björklund, J. A., Gunnarson, B. E., Seftigen, K., Esper, J. & Linderholm, H. W. Blue intensity and density from northern Fennoscandian tree rings, exploring the potential to improve summer temperature reconstructions with earlywood information. *Clim. Past* **10**, 877–885 (2014).
76. Schmidt, G. et al. Climate forcing reconstructions for use in PMIP simulations of the Last Millennium (v1.1). *Geosci. Model Dev.* **5**, 185–191 (2012).
77. National Research Council. *Surface Temperature Reconstructions for the Last 2,000 Years* (National Academies Press, 2007).
78. Cook, E. R. & Peters, K. The smoothing spline: a new approach to standardizing forest interior tree-ring width series for dendroclimatic studies. *Tree Ring Bull.* **41**, 45–53 (1981).
79. Matalas, N. C. Statistical properties of tree ring data. *Int. Assoc. Sci. Hydrol. Bull.* **7**, 39–47 (1962).
80. Piermattei, A., Crivellaro, A., Carrer, M. & Urbinati, C. The "blue ring": anatomy and formation hypothesis of a new tree-ring anomaly in conifers. *Trees* **29**, 613–620 (2015).
81. Percival, D. B. & Walden, A. T. *Spectral Analysis for Physical Applications* (Cambridge Univ. Press, 1993).
82. Huybers, P. pmtmPH.m v.1.0.0.0. https://www.mathworks.com/matlabcentral/fileexchange/2927-pmtmph-m (MATLAB Central File Exchange, 2022).
83. Haurwitz, M. W. & Brier, G. W. A critique of the superposed epoch analysis method: its application to solar–weather relations. *Mon. Weather Rev.* **109**, 2074–2079 (1981).
84. Brad Adams, J., Mann, M. E. & Ammann, C. M. Proxy evidence for an El Nino-like response to volcanic forcing. *Nature* **426**, 274–278 (2003).
85. Blarquez, O. & Carcaillet, C. Fire, fuel composition and resilience threshold in subalpine ecosystem. *PLoS ONE* **5**, e12480 (2010).
86. Gao, C., Robock, A. & Ammann, C. Volcanic forcing of climate over the past 1500 years: an improved ice core-based index for climate models. *J. Geophys. Res. Atmos.* **113**, D23111 (2008).
87. Toohey, M. & Sigl, M. Volcanic stratospheric sulfur injections and aerosol optical depth from 500 BCE to 1900 CE. *Earth Syst. Sci. Data* **9**, 809–831 (2017).

Author contributions J.B., G.v.A., K.S. and M.C. conceptualized the study. H. Grudd, B.E.G., J.E., M.C., E.P. and D.N. conducted fieldwork and provided physical samples for the wood anatomical analyses. G.v.A., M.C., M.V.F., S.K. and E.P. coordinated, processed and measured the wood anatomical data. J.B. and K.S. performed the analyses with input from H. Goosse, G.v.A., P.F., D.C.F. and M.S. J.B. wrote the paper and all authors have reviewed and helped to revise the paper.
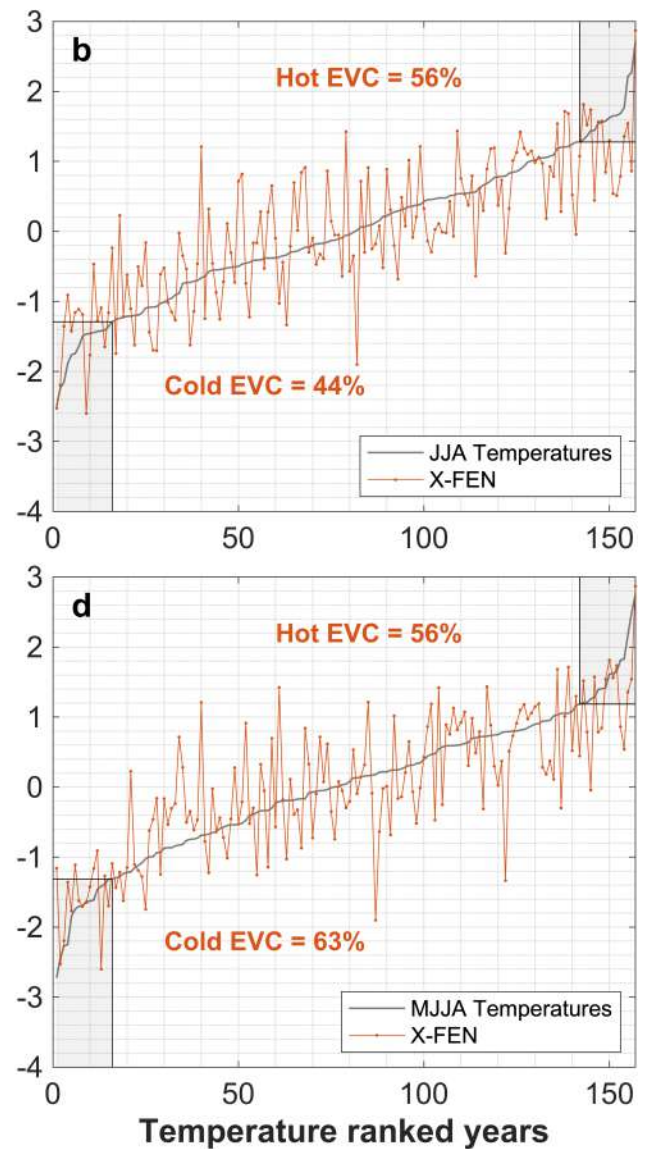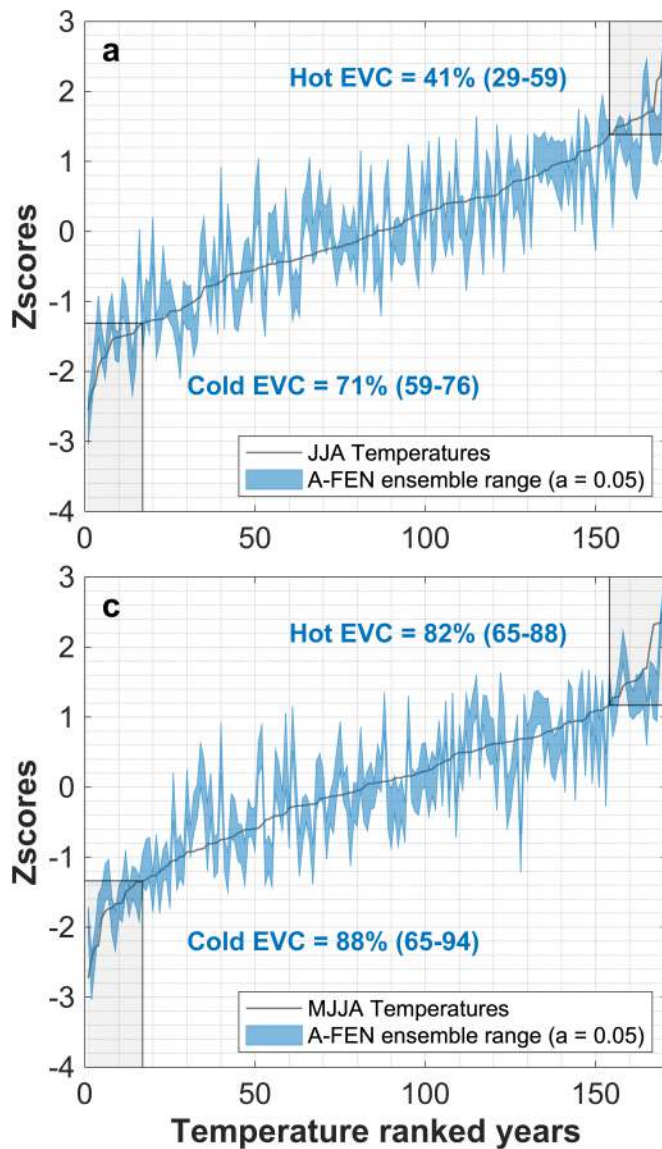
Extended Data Fig. 1 | Evaluation of the ability of the A-FEN and X-FEN reconstructions to capture extremes in climate targets. The instrumental temperature data was sorted from coldest to warmest and plotted together with the reconstruction values of the corresponding years. The grey boxes are bound by the 10% coldest and warmest years, and the 10th and 90th percentile of the zscore temperatures, respectively. If an extreme reconstruction value is found within the grey box, the extreme is defined as "captured". The sum of the captured values divided by the potential sum of values, is calculated and presented as a percentage of extreme value capturing (EVC). In McCarroll, et al.[21], a significance testing was implemented, and for 160–170 years of climate data, $p < 0.001$ is achieved if more than 40% of values are captured. **a)** A-FEN's ability to capture JJA temperature extremes. **b)** X-FEN's ability to capture JJA temperature extremes. **c)** and **d)** show same analysis as **a)** and **b)** but using the target MJJA. Both datasets thus display significant amounts of extremes captured, but the A-FEN captures significantly more than the X-FEN for the MJJA target season. The X-FEN captures a higher percentage of cold extremes if the MJJA target season is used but the same percentage of warm extremes regardless of target season. The rationale for using MJJA as the target season for X-FEN is thus less clearcut than for the A-FEN. JJA is the target season used in the publications originally presenting the MXD data[2,37] and is thus used in the main text for the other comparisons.
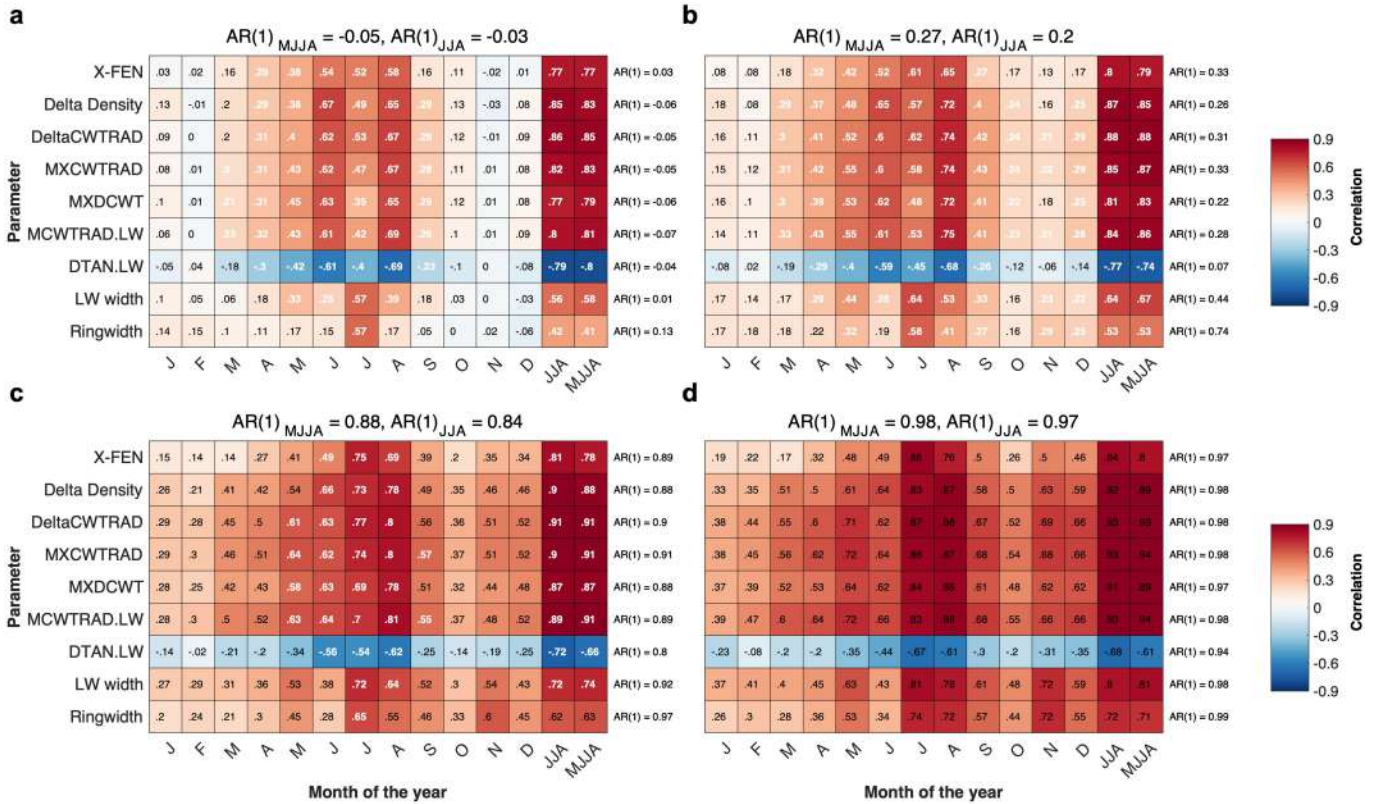
Extended Data Fig. 2 | Monthly climate correlations across a range of frequency domains. a) high-pass filtered data (cubic smoothing splines with 50 % frequency response cut-off at 40 years (HP40yrs)) correlated with identically treated temperature data. **b)** RCS detrended data correlated with untreated temperature data. **c)** low-pass filtered data (LP5yrs), and **d)** (LP10yrs), correlated with identically treated temperature data, respectively. The monthly temperature data were retrieved from HadCRUT5[68] (5° gridded monthly dataset, Lat. 65–70° N, Lon. 15–30° E). Correlation coefficients in white are significant at p < 0.01, and black coefficients are insignificant. When 10-year low-pass filtered data are used, the autocorrelation is so high that it is impossible to detect significance after adjusting for loss of degrees of freedom[79], why it is meaningless to continue the analysis over even lower frequencies. The parameters or reconstructions reside on the y-axis, and each monthly temperature or monthly target season on the x-axis. First order autocorrelation, AR(1), of the JJA and MJJA temperatures are given on top of each panel as a reference, and the tree-ring parameter AR(1) can be found in the right margin of each panel. The period of analysis covers the full length-overlap between all datasets (1850–2019 for anatomical parameters and 1850–2010 for the X-FEN). The results are very similar if the 1850–2010 period is used for the QWA data. The delta radial cell wall thickness (DeltaCWTRAD) parameter was established as predictor for the A-FEN reconstruction due to overall performance in the analysis.
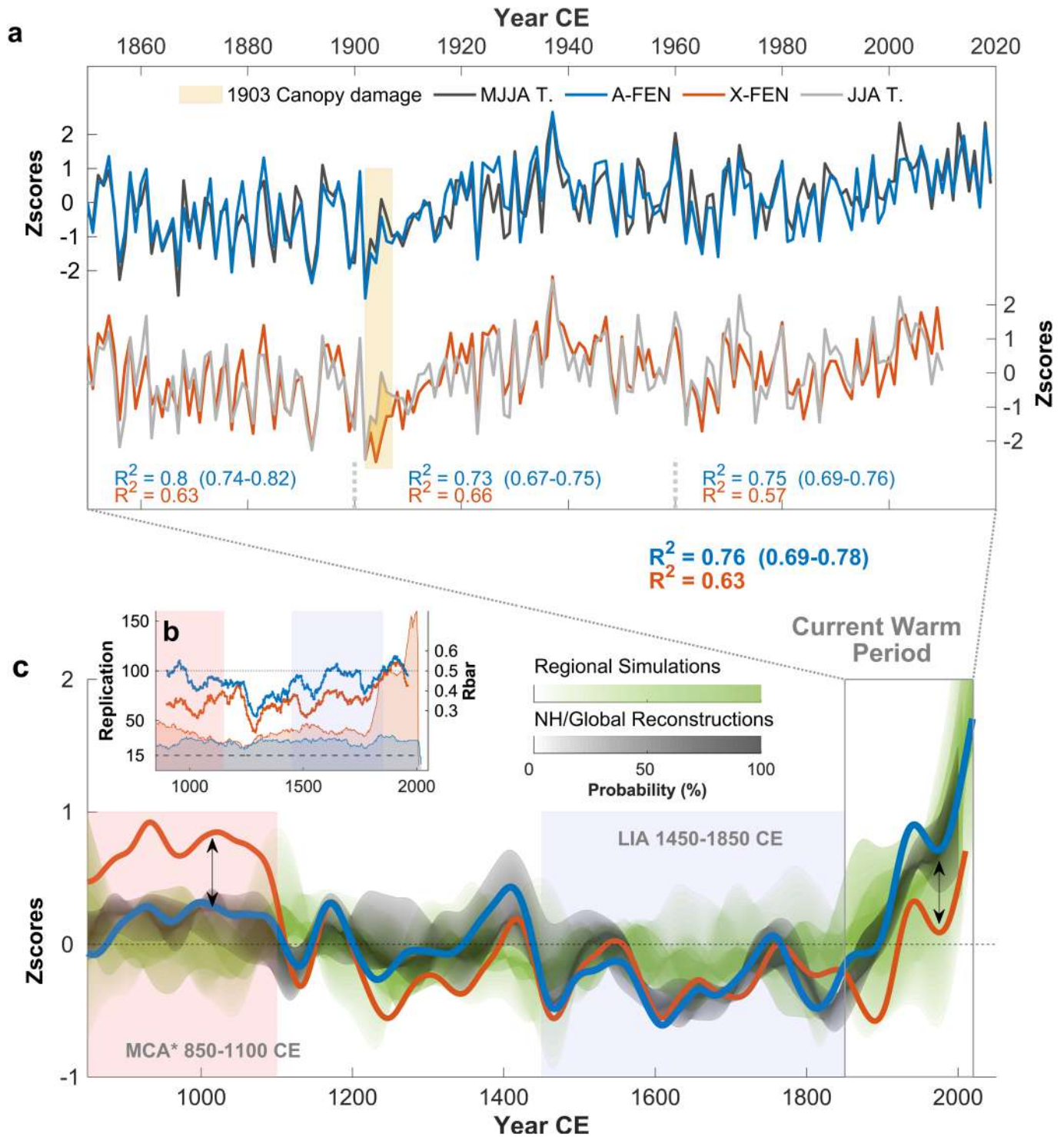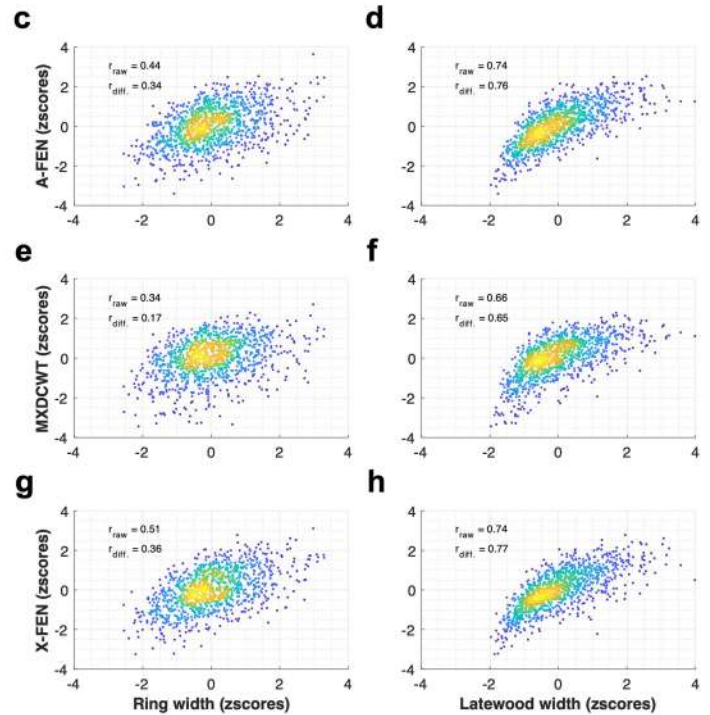
Extended Data Fig. 3 | Warm-season temperature reconstruction skill of A-FEN and X-FEN, as well as comparisons with existing large-scale reconstructions and regional climate simulations <u>but using QWA data that has not been detrended using the RCS approach.</u> The non-QWA datatypes are identical to Fig. 2 of the Main manuscript and the vertical arrows have the exact positions and dimensions as in Fig. 2 for reference. **a**) A-FEN (produced in this study) calibrated using regional mean air MJJA temperatures[68] (R^2 ensemble range within brackets (a = 0.05)), and results for the X-FEN (from Wilson, et al.[9]) using corresponding JJA temperatures. The irregular winter/spring of 1902/1903, led to a massive dieback of yearly branch-shoots in the region[55], highlighted by

the yellow area. In these years with extremely narrow rings, the X-ray technique struggles to measure high MXD values due to its comparatively lower effective measurement resolution[29] (see Extended Data Fig. 4). **b**) Replication and pairwise inter-series correlation ($\bar{R}$) of A-FEN in blue and the X-FEN in red. **c**) Centennial-scale variations (see Methods) compared between A-FEN, X-FEN, climate model simulations, and NH and global temperature reconstructions. The five large-scale reconstructions[1,9,10,38,39], as well as the eleven regionally extracted climate-model simulations[40–50] are represented by probabilistic percentile ranges. The vertical arrows highlight the overall discrepancies of the X-FEN compared to the other data.

Article
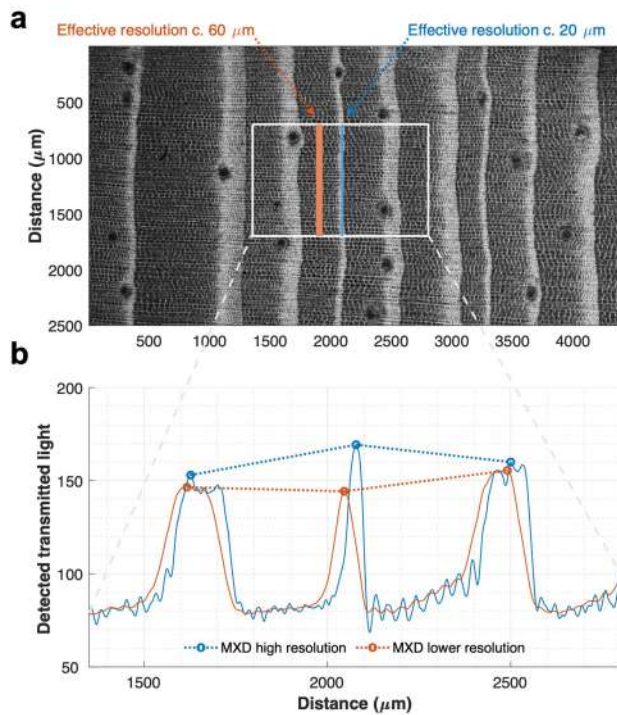
Extended Data Fig. 4 | Illustration of the issue with comparatively low measurement resolution for X-ray MXD. a) X-ray image with analysis track path indicated within the solid white rectangle, and examples of the effect of different effective measurement resolutions. **b)** The photosensors in a) build up measurement profiles, where the blue sensor builds the blue profile corresponding to a 20-micron effective measurement resolution, and the orange sensor builds up the orange profile corresponding to a 60 micron effective measurement resolution, approximating the effective measurement resolution of the X-ray methodology[29]. Note how the time series of MXD reflect inverse variations if developed using high-resolution or low-resolution equipment, i.e., the middle ring exhibits the lowest or highest value depending on resolution. The explanation for this is that very narrow latewood widths are associated with comparatively lower MXD values even though the "true" MXD value may be high. **c)** Relationships between TRW and A-FEN and **d)** LWW and A-FEN. **e)** Relationships between TRW and anatomical MXD (MXDCWT) and **f)** LWW and MXDCWT. **g)** Relationships between TRW and X-FEN and **h)** LWW and X-FEN. All datasets display correlations and using datapoints covering 850–2005 CE. Note how the X-FEN always is stronger correlated with TRW and LWW than the MXDCWT. A higher correlation is expected if TRW or LWW is affecting the measurement. Spearman rank correlation coefficients were used due to the possibly non-linear relationships between width and density. R_{raw} and $r_{diff.}$ refers to untreated and first differenced data prior to correlations, respectively.
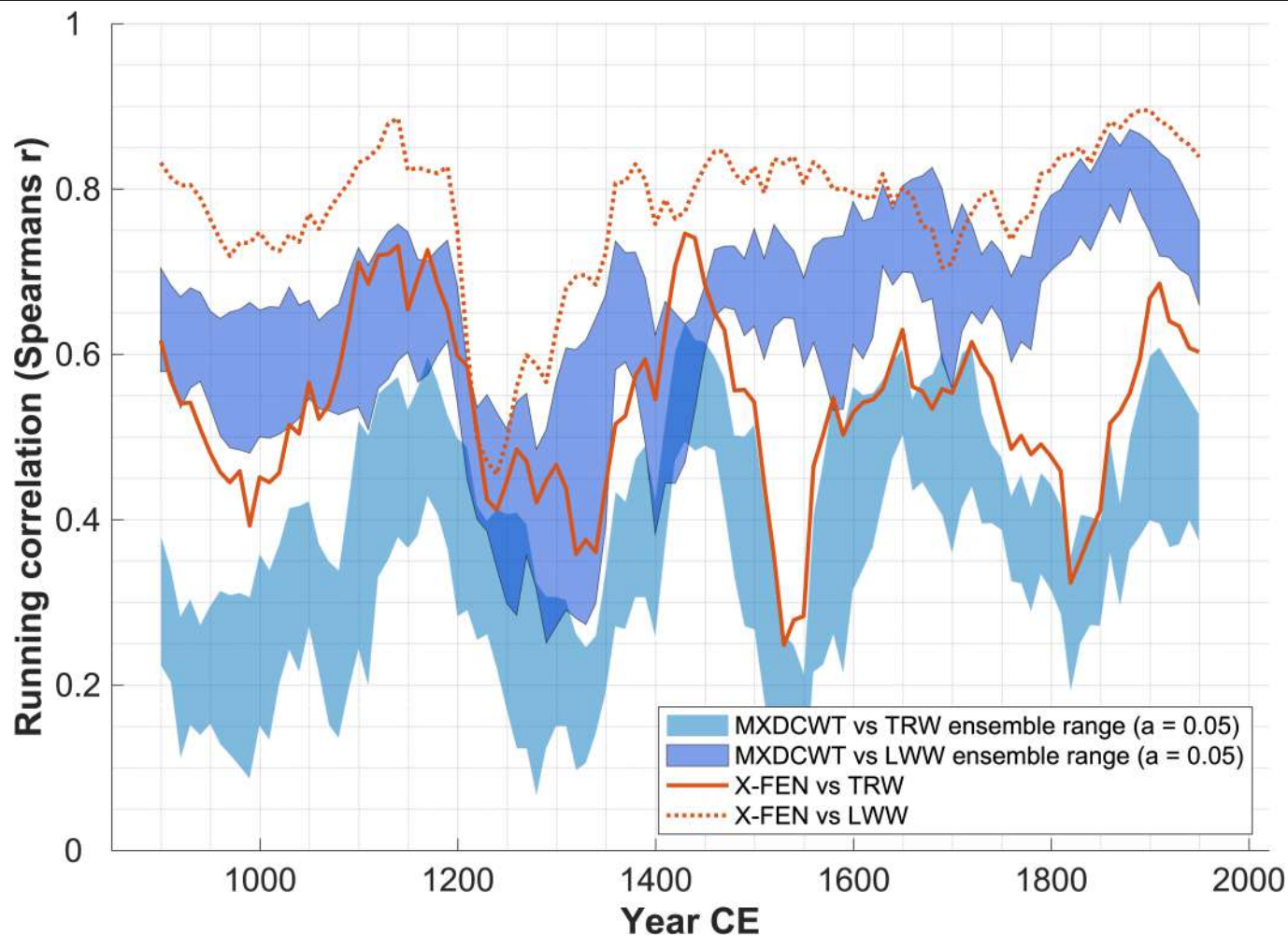
Extended Data Fig. 5 | Moving window correlation coefficients revealing that X-ray MXD exhibits a stronger relationship with ring width and latewood width than anatomical MXD. Ring width (TRW) versus anatomical MXD (MXDCWT) and X-FEN, as well as latewood width (LWW) versus MXDCWT and X-FEN. Spearman rank correlations were used on RCS-detrended chronologies with a 100-year base-lengths and 10-year overlaps. For the anatomical MXDCWT data, 100 sub-sampled chronologies with 15 trees/year were used to create ensemble ranges represented in blue shades. Deviations from these blue shaded areas represent significant differences (p < 0.05) from the TRW and LWW correlations with MXDCWT, respectively. The X-FEN correlations often reside outside the blue areas, and at higher correlations with TRW and LWW respectively, indicating occasionally stronger dependence of X-FEN on these parameters.
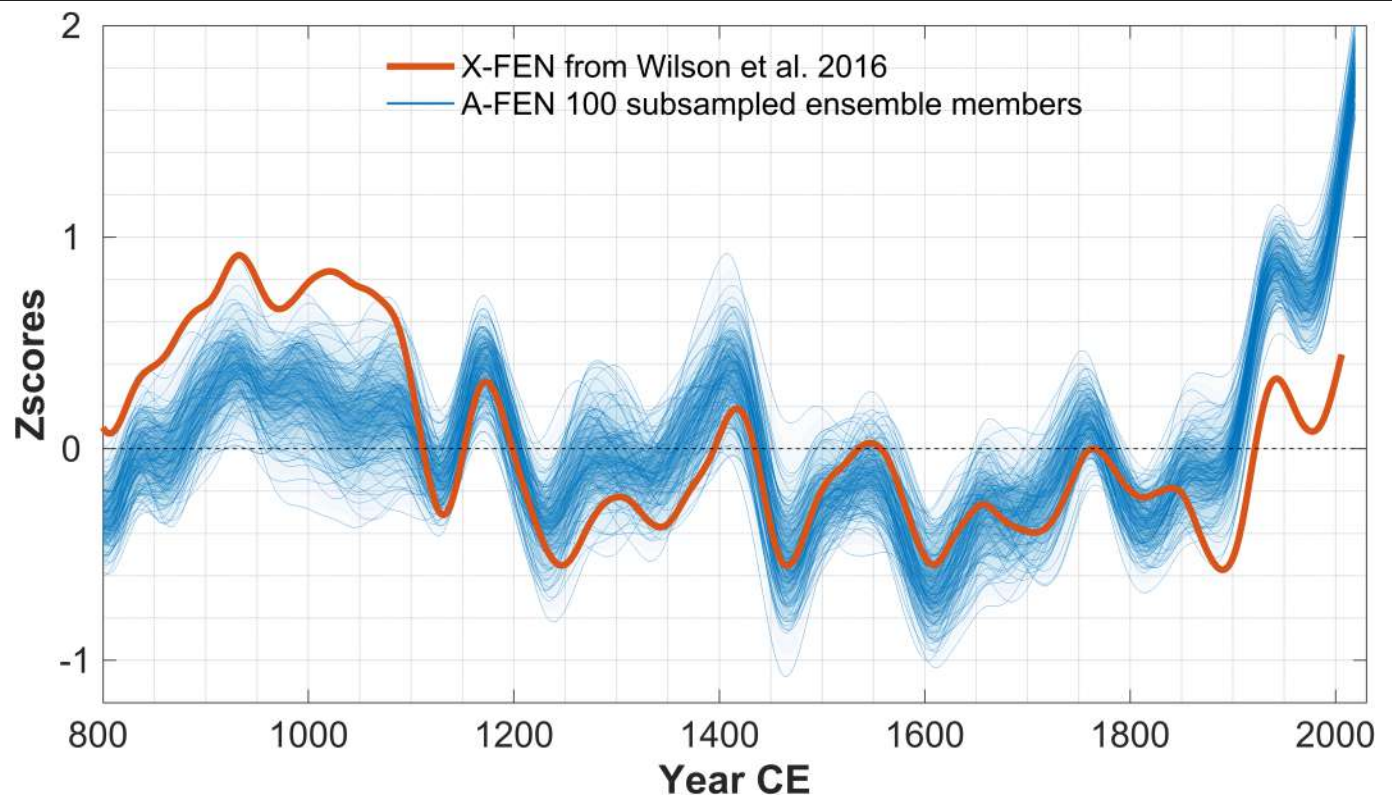
Extended Data Fig. 6 | Comparison of A-FENs complete set of ensemble members with X-FEN. RCS-detrended A-FEN (data from this study) versus the X-FEN (data from Wilson, et al.[9]), smoothed using cubic smoothing splines with 50% frequency response cut-off at 100 years. Note that no A-FEN ensemble member exhibit the protracted warmth during the MCA and the relatively low temperatures during the CWP, as does the X-FEN.
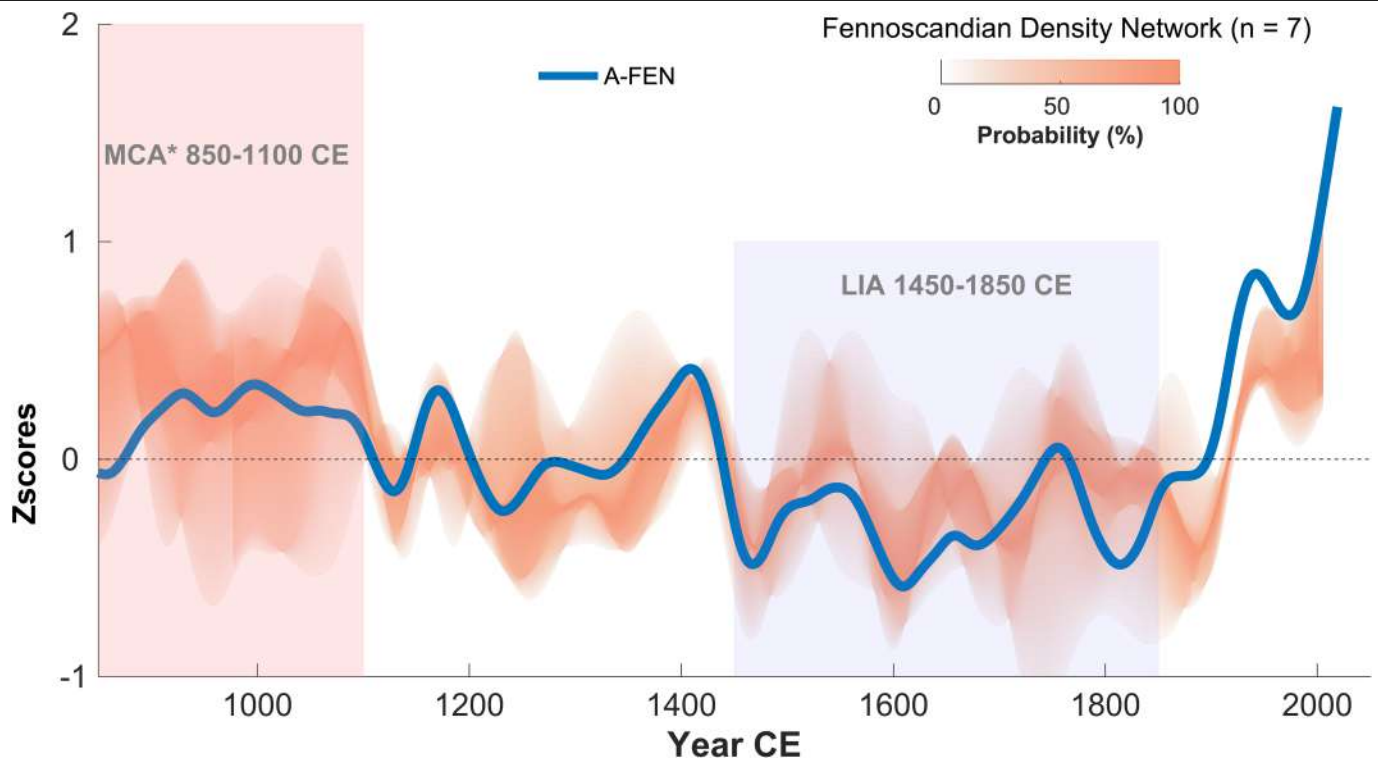
Extended Data Fig. 7 | Comparison of A-FEN versus a network of millennium-long Fennoscandian MXD datasets showing the wide range of medieval estimates and comparably modest modern warming. The MXD based temperature reconstructions from Fennoscandia are retrieved from Wilson, et al.[9], Schneider, et al.[10] and McCarroll, et al.[61] represented by a probabilistic percentile range.
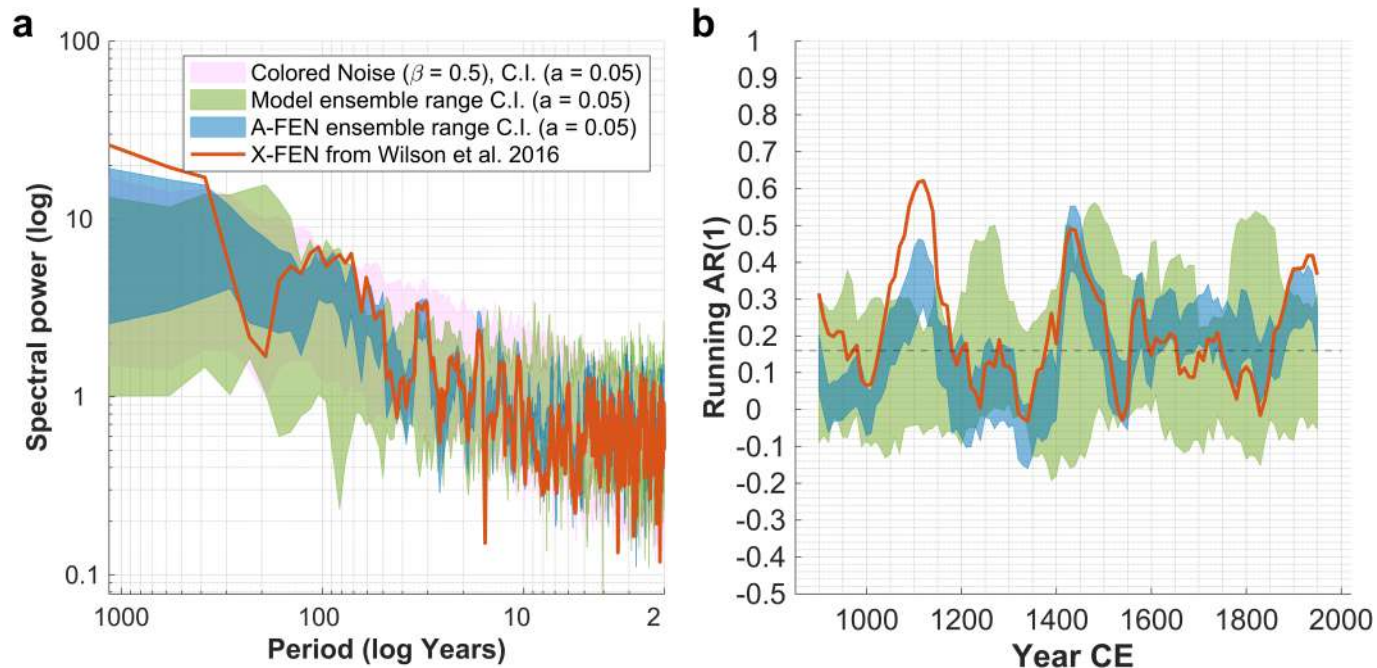
Extended Data Fig. 8 | Spectral properties and first order autocorrelation of the reconstructions and models compared in this study. a) Spectral properties of the A-FEN ensemble and X-FEN on the backdrop of the model ensemble range, as well as the range of a 1000 timeseries, of equal length to the A-FEN, of colored noise with a beta coefficient of 0.5. (Beta coefficient for White noise = 0, Pink noise = 1). **b)** Running autocorrelations AR(1) calculated for 100-year window lengths, shifted by 10-year lags.
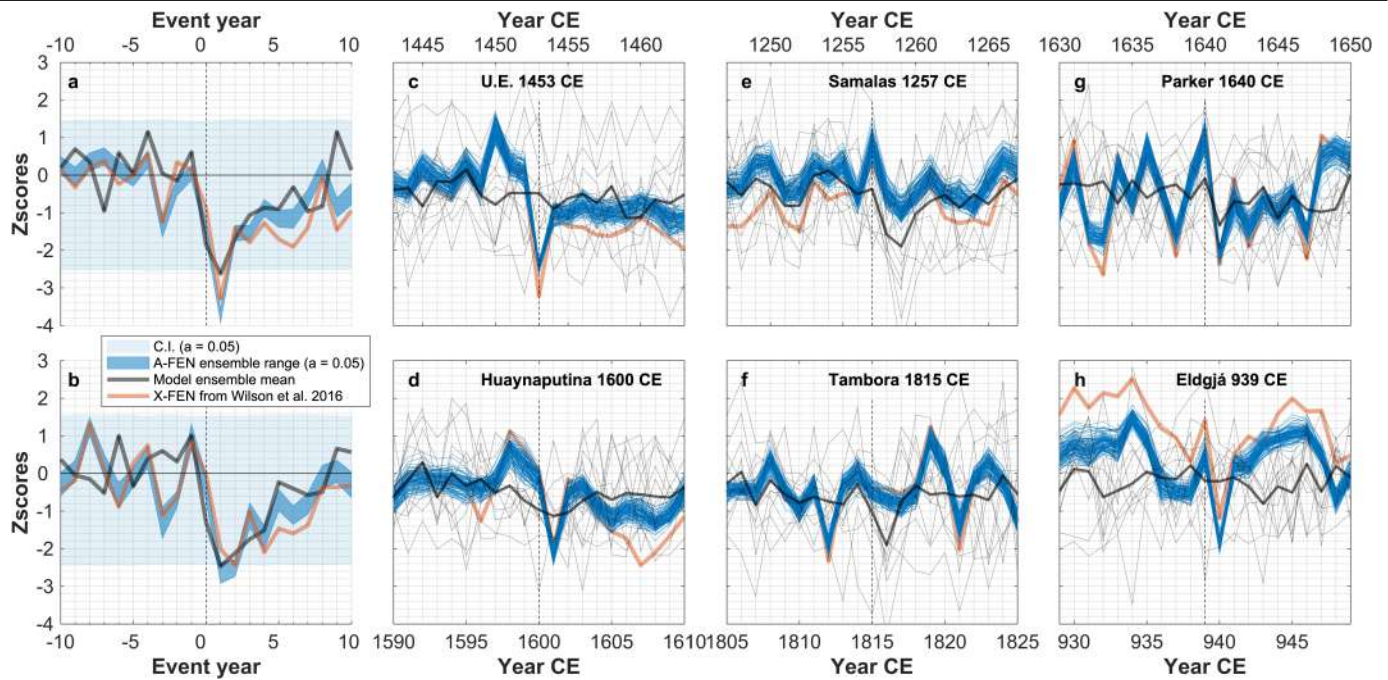
Extended Data Fig. 9 | Superposed epoch analysis and comparisons over individual eruption years for the A-FEN ensemble, X-FEN and the model ensemble. SEA's using Gao, et al.[86] event lists of the 10 **a**) and 30 **b**) of the largest (based on sulfate aerosol injection) northern Hemisphere events. The model simulations were all extracted from the corresponding grid cells Lat 65–70° N, Lon 15–30° E. We used only Gao et al as basis for the event lists because most models in our ensemble were forced with Gao et al, but note that this list may not be optimal for some models and the tree-ring data. We employed a model ensemble mean in the SEA, to explore the degree of volcanic cooling the models express. **c-h**) Proxy vs model response to some specific major volcanic events dated according to Toohey and Sigl[87]. The responses to U.E. 1453 CE, Huaynaputina and Eldgjá are pronounced in the proxy data but not in the models. The responses to Samalas and Tambora are pronounced in the models but not in the proxy data. The response to Parker is present in both models and proxy.

Article

Extended Data Table 1 | Reconstruction statistics for Fennoscandian tree-ring anatomy A-FEN

Calibration periods	1960–2019	1900–1959	1850-1899	1850-2019
Explained variance, R^2 adjusted	0.74	0.72	0.81	0.77
No. of observations, (d.f.)	60 (58)	60 (58)	50 (48)	170 (168)
First validation period	**1900–1959**	**1960–2019**	**1900–1959**	
Reduction of error (RE)	0.68	0.77	0.72	
Coefficient of efficiency (CE)	0.72	0.73	0.71	
Second validation period	**1850-1899**	**1850-1899**	**1960–2019**	
Reduction of error (RE)	0.78	0.77	0.69	
Coefficient of efficiency (CE)	0.81	0.77	0.64	

Reconstruction statistics were calculated using May-August from HadCRUT5[68] (5° gridded monthly dataset, Lat. 65–70° N, Lon. 15–30° E). The reconstruction statistics were calculated over three split calibration/validation periods.