

# SCIENTIFIC REPORTS



OPEN

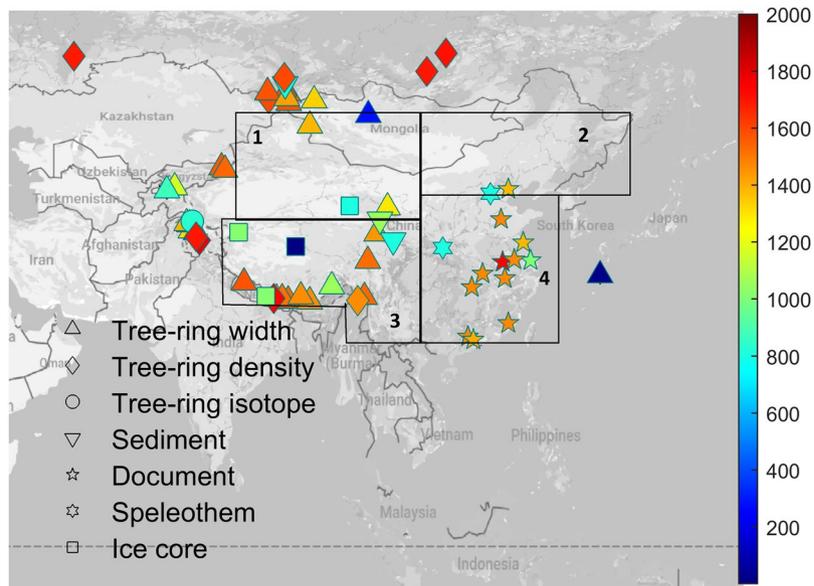
## East Asian warm season temperature variations over the past two millennia

Huan Zhang<sup>1</sup>, Johannes P. Werner<sup>2</sup>, Elena García-Bustamante<sup>3</sup>, Fidel González-Rouco<sup>4</sup>, Sebastian Wagner<sup>5</sup>, Eduardo Zorita<sup>6</sup>, Klaus Fraedrich<sup>6</sup>, Johann H. Jungclauss<sup>6</sup>, Fredrik Charpentier Ljungqvist<sup>7,8</sup>, Xihua Zhu<sup>9</sup>, Elena Xoplaki<sup>1</sup>, Fahu Chen<sup>10</sup>, Jianping Duan<sup>11</sup>, Quansheng Ge<sup>12</sup>, Zhixin Hao<sup>12</sup>, Martin Ivanov<sup>1</sup>, Lea Schneider<sup>1</sup>, Stefanie Talento<sup>1</sup>, Jianglin Wang<sup>13</sup>, Bao Yang<sup>13</sup> & Jürg Luterbacher<sup>1,14</sup>

East Asia has experienced strong warming since the 1960s accompanied by an increased frequency of heat waves and shrinking glaciers over the Tibetan Plateau and the Tien Shan. Here, we place the recent warmth in a long-term perspective by presenting a new spatially resolved warm-season (May–September) temperature reconstruction for the period 1–2000 CE using 59 multiproxy records from a wide range of East Asian regions. Our Bayesian Hierarchical Model (BHM) based reconstructions generally agree with earlier shorter regional temperature reconstructions but are more stable due to additional temperature sensitive proxies. We find a rather warm period during the first two centuries CE, followed by a multi-century long cooling period and again a warm interval covering the 900–1200 CE period (Medieval Climate Anomaly, MCA). The interval from 1450 to 1850 CE (Little Ice Age, LIA) was characterized by cooler conditions and the last 150 years are characterized by a continuous warming until recent times. Our results also suggest that the 1990s were likely the warmest decade in at least 1200 years. The comparison between an ensemble of climate model simulations and our summer reconstructions since 850 CE shows good agreement and an important role of internal variability and external forcing on multi-decadal time-scales.

Over the past two decades, many proxy-based temperature reconstructions at hemispheric and continental scale covering the past one to two millennia have been published<sup>1–4</sup>. Northern Hemisphere (NH) temperature reconstructions generally show a warm period ca. 1–300 CE, followed by a cooler period ca. 300–700 CE, warm conditions from ca. 800 to 1250 CE, and a cooler climate from about 1300 CE until the end of the 19th century<sup>1,2</sup>. The reconstructions agree with instrumental observations in showing a rapid temperature increase in the late 20th century with exceptionally high temperatures in the last few decades in the context of the past centuries<sup>1,2,5,6</sup>. Regional

<sup>1</sup>Department of Geography, Climatology, Climate Dynamics and Climate Change, Justus Liebig University Giessen, Giessen, Germany. <sup>2</sup>University of Bergen, Department of Earth Science and Bjerknes Centre for Climate Research, Allég. 41, NO-5020, Bergen, Norway. <sup>3</sup>Department of Energy, Renewable Energy Section, Research Center for Energy, Environment and Technology (CIEMAT), Madrid, Spain. <sup>4</sup>Department of Physics of the Earth and Astrophysics, IGEO (UCM-CSIC), Universidad Complutense de Madrid, Madrid, Spain. <sup>5</sup>Institute of Coastal Research, Helmholtz-Zentrum Geesthacht, D-21502, Geesthacht, Germany. <sup>6</sup>Max-Planck Institute for Meteorology, Bundesstrasse 53, D-20146, Hamburg, Germany. <sup>7</sup>Department of History, Stockholm University, Stockholm, Sweden. <sup>8</sup>Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden. <sup>9</sup>Center for Earth System Research and Sustainability, CliSAP, University of Hamburg, Hamburg, Germany. <sup>10</sup>MOE Key Laboratory of Western China's Environmental System, Lanzhou University, Lanzhou, 730000, China. <sup>11</sup>State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, 100093, Beijing, China. <sup>12</sup>Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, 11A, Datun Road, Chaoyang District, Beijing, 100101, China. <sup>13</sup>Key Laboratory of Desert and Desertification, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou, China. <sup>14</sup>Centre for International Development and Environmental Research, Justus Liebig University Giessen, Giessen, Germany. Correspondence and requests for materials should be addressed to H.Z. (email: [huan.zhang@geogr.uni-giessen.de](mailto:huan.zhang@geogr.uni-giessen.de))



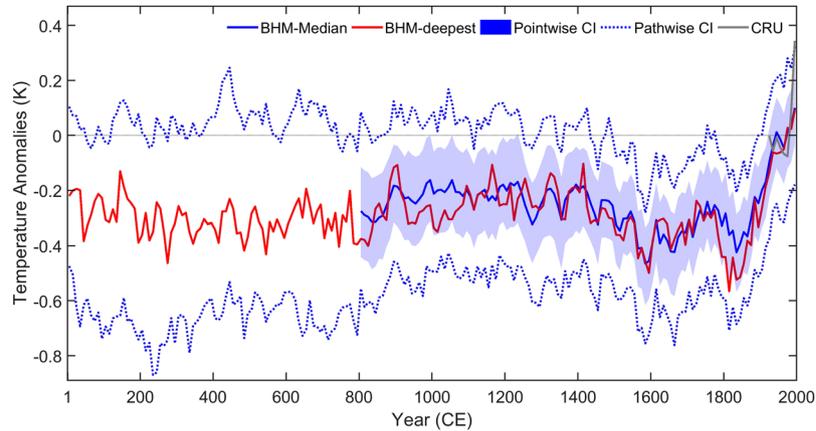
**Figure 1.** Locations of the proxies used in this study; the color indicates the starting year of the proxy record. The four selected sub-regions (with different climate) mainly include Northwest China-western Mongolia (marked as 1), Northeast China-eastern Mongolia (2), Tibetan Plateau (3) and Southeast China (4). The figure was generated using Matlab 2015b (<http://www.mathworks.com/>). The map in the figure was queried from Google Static Map APIs (<http://code.google.com/apis/maps/>).

temperature reconstructions from East Asia/China show a similar behavior to that of the NH, but they disagree on whether the warming over the last few decades has been unprecedented or not in the context of the past<sup>7,8</sup>.

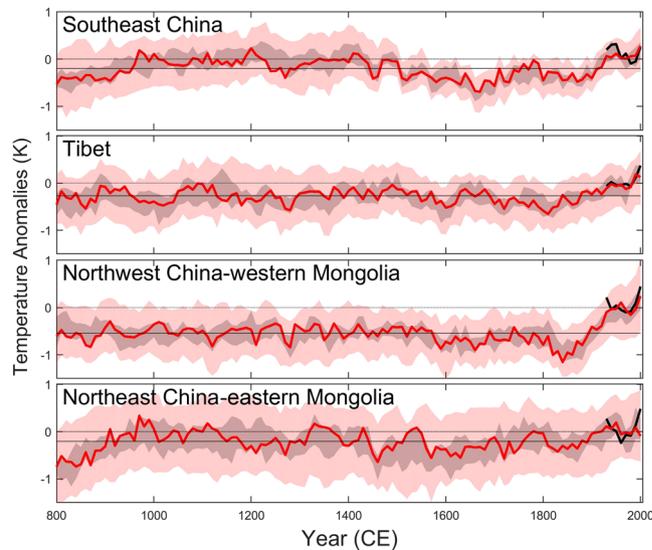
Climate field reconstructions (CFRs), complementing regional average reconstructions, offer temporal and also spatially resolved information over past temperature variability. So far, two field temperature reconstructions over East Asia have recently been published<sup>9,10</sup>. Cook *et al.*<sup>9</sup> reconstructed summer temperatures at annual resolution back to 800 CE using 229 tree-ring chronologies and applying a point-by-point regression method. Shi *et al.*<sup>10,11</sup> generated an annual-resolved temperature field reconstruction covering the last millennium based on 418 proxy records (392 being tree-ring chronologies), applying the Regularized Expectation Maximization method with Truncated Total Least Squares (RegEM-TTLS). The CFR techniques used in both of the two reconstructions are based on multivariate linear regression approaches, which are based on calibrating proxy records against instrumental data<sup>12–17</sup>. Apart from linear regression approaches, the Bayesian hierarchical model<sup>18</sup> (BHM), a probabilistic approach, is another CFR method that has been used to reconstruct the past regional temperature variability for Europe and the Arctic<sup>6,19,20</sup>. The advantages of the Bayesian method include: a more comprehensive assessment of uncertainties, no need for grid infilling (for instrumental and/or proxies), transparent assumptions about the spatial-temporal properties and no dependence on long-distance correlations that might not be stable in time. In addition, probabilistic ensembles of equally likely reconstructions provide unique means to evaluate e.g. trends or extreme periods<sup>20</sup>. BHM is suitable for estimating uncertainties of reconstructed temperatures<sup>21,22</sup>. In this study, we applied BHM to 59 selected temperature-sensitive proxy records (see Fig. 1 and the methods section) over East Asia (60°–160°E/10°–60°N) and generated a new warm-season (May–September) temperature reconstruction for the past 2000 years with decadal resolution and associated uncertainties. We compare our findings with earlier temperature reconstructions and with an ensemble of state-of-the-art climate model simulations of temperature over the period 850–2000 CE in order to assess the influence of changes in external forcing and the simulated internal climate variability over East Asia.

## Results and Discussion

Two regional average temperature time-series (anomalies w.r.t. 1961–1990 CE) together with their uncertainties are presented in Fig. 2. Point-wise median values covering the period 801–2000 CE, denoted as BHM-median; and path-wise median values covering the period 1–2000 CE, denoted as BHM-deepest. The regional average temperature reconstruction over the period 801–2000 CE is tested as robust against changes in the proxy-data network and methodology (the Pearson correlation coefficients are higher than 0.86 between the BHM and the composite reconstructions; see the Methods section for more details). BHM-median represents the best temperature estimates for each decade, but may suffer from loss of variance for early centuries and over some specific areas due to sparse proxy coverage. However, this potential issue has been solved for BHM-deepest. BHM-deepest is the median trajectory of reconstructed temperature variations if considering each ensemble member as a possible trajectory of past temperature (illustrated in the methods subsection “Deepest curve and path-wise confidence intervals”). Compared with BHM-median, BHM-deepest exhibits a higher temperature variability. These two BHM reconstructions (BHMs) share a large part of the inter-decadal temperature variability ( $r=0.83$ ,  $df=118$ ,  $p<0.01$ ) over their common period 801–2000 CE. The BHM reconstructions indicate that relatively warm East



**Figure 2.** Reconstructed area-weighted decadal temperature anomalies (w.r.t. 1961–1990 CE) using the full proxy network and Bayesian hierarchical model with instrumental data excluded. Solid blue line: the median values; Solid red line: the deepest curve; blue shades: 90% point-wise confidence intervals; dashed blue line: 90% path-wise confidence intervals; solid gray line: instrumental warm season temperature (Jones *et al.*<sup>39</sup>).



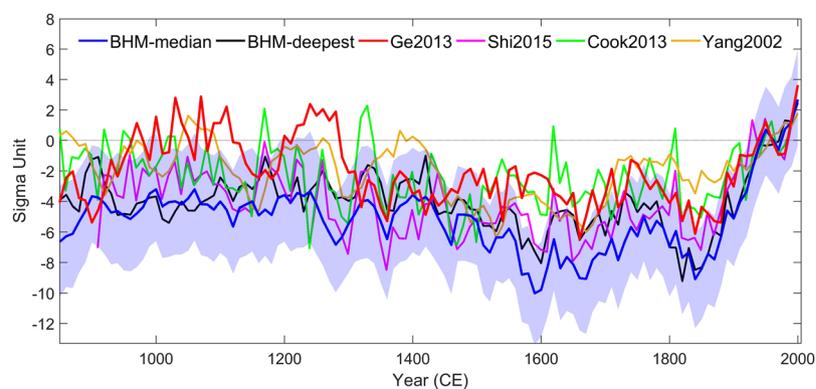
**Figure 3.** Reconstructed area-weighted temperature anomalies (with respect to 1961–1990 CE) using the 59 proxy record network and Bayesian hierarchical model with instrumental data excluded in the selected regions. Red curve: the deepest curve; thick black curve: decadal mean instrumental temperature; light black line: regional average temperature with respect to the whole period (801–2000 CE); gray shades: the variance range of the ten deepest curves; red shades: 90% path-wise confidence intervals.

Asia warm-season conditions prevailed from the beginning of the Common Era until the 3<sup>rd</sup> century CE, and were followed by generally cooler conditions in the subsequent centuries (Fig. 2). Rather high temperatures are found ca. 900–1200 CE, in the early 14<sup>th</sup> century and at the turn of the 15<sup>th</sup> century. The relatively warm conditions were followed by overall cooler conditions from 1450–1850 CE. Since the 1850s, the warm-season temperature has been increasing, a trend that holds until the present with a warming pause between ca. 1940 and 1970 CE. This is consistent with instrumental observations. Joint information from our regional average temperature reconstruction and the instrumental data indicates that the 1990s were likely the warmest decade in the last 2000 years, as it exceeds the upper boundary of 90% path-wise confidence intervals (Fig. 2).

The BHM spatial temperature reconstructions over four sub-regions indicate that the warmer conditions were spatially heterogeneous from the 9<sup>th</sup> century to the end of the 13<sup>th</sup> century (Fig. 3). For instance, warmer conditions can be found over south-east China, but anomalous lower temperatures were reconstructed over the Tibetan Plateau (from mid-10<sup>th</sup> century to end of the 11<sup>th</sup> century and again from mid of the 12<sup>th</sup> century to end of the 13<sup>th</sup> century). In contrast, the cold conditions from the 16<sup>th</sup> century to the 19<sup>th</sup> century covered almost the whole of East Asia. Particularly cold periods are prominent in the mid-17<sup>th</sup> century over south-east China and in the 19<sup>th</sup> century over northwest China and western Mongolia. The reconstructions indicate that, the temperature in the 1990s is likely to be the highest in the last millennium over the Tibetan Plateau<sup>23,24</sup>, northwest China and

(a)						
Reconstruction name	Number of proxy data	Method	Season	Region	Climate field reconstruction?	Publication
Shi2015	418 multi-proxy (including 392 tree-ring chronologies)	Regularized expectation maximization algorithm	Summer	Asia	Yes	<sup>10</sup>
Cook2013	229 tree-ring proxies	point-by-point regression	Summer	East Asia	Yes	<sup>5,9</sup>
Ge2013	22 multi-proxy	Composite method	Annual	China	No	<sup>26</sup>
Yang2002-weighted	9 multi-proxy	Composite method	Annual	China	No	<sup>27</sup>
(b)						
Cor.	BHM-median	BHM-deepest	Shi2015	Cook2013	Ge2013	
BHM61-deepest	<b>0.83</b>					
Shi2015	<b>0.68</b>	<b>0.52</b>				
Cook2013	<b>0.46</b>	<b>0.40</b>	<b>0.79</b>	—		
Ge2013	<b>0.55</b>	<b>0.44</b>	<b>0.47</b>	0.24	—	
Yang2002	<b>0.62</b>	<b>0.45</b>	<b>0.45</b>	<b>0.28</b>	<b>0.51</b>	

**Table 1.** (a) Selected published temperature reconstructions over China or Asia. (b) Pearson correlation analysis among different reconstructions of decadal temperature for the period 901–1990 CE,  $df = 107$  ( $p < 0.01$  are in bold). The critical value for significant correlation at the level  $p = 0.01$  is 0.25. It should be kept in mind that the real  $p$ -value is probably much higher than 0.01 due to strong autocorrelation in the reconstructions.



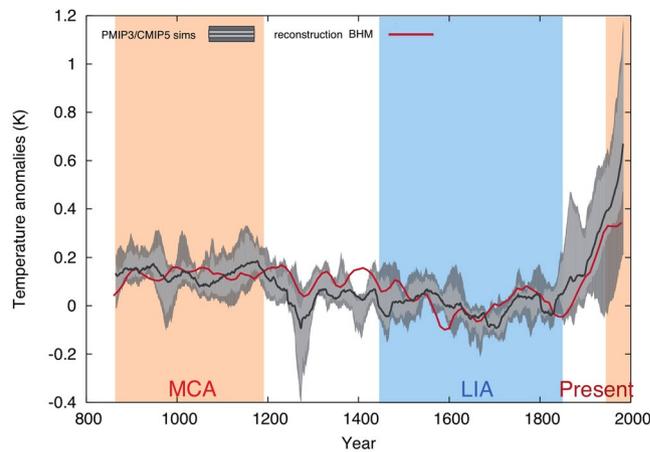
**Figure 4.** BHM reconstructions from this study in comparison with published evidence for East Asia (all reconstructions are presented at decadal resolution and standardized with respect to 1921–1990 CE). Dashed black lines are 90% point-wise confidence intervals of BHMs.

western Mongolia<sup>25</sup>. The warm conditions in the 1990s are not exceptional over south-east China. Equally warm conditions prevailed during medieval times<sup>26</sup> and during the period 1921–1940 CE.

**Comparison with China/ East Asia temperature reconstructions.** We compared the BHMs reconstructions with four other regional average temperature reconstructions over China / East Asia (Yang<sup>27</sup>, Cook<sup>9</sup>, Ge<sup>26</sup>, and Shi<sup>10</sup>, see Table 1 for a description of reconstructions and cross-correlation analysis) for the last 1150 years. The reconstructions differ regarding their amplitudes of variation for the instrumental period (Fig. S6). For a fair comparison, all reconstructions were standardized with respect to the instrumental period 1921–1990 CE (Fig. 4).

The reconstructions roughly agree on the timing of three characteristic periods: a multi-centennial warm period ca. 900–1200 CE, the cold “Little Ice Age” ca. 1450–1850s CE and a continuous warming after the 1850s. They differ in terms of the shape and strength of temperature variations and the timing of cold and warm periods. Moreover, Shi2015 and our new BHM reconstruction suggests, that China/East Asia experienced likely the warmest decade during the 1990s for more than 1100 years.

Ge2013 and our BHMs share a large portion of proxies and show a similar temperature variability for the last half millennium (with the exception that the BHMs reveal a stronger centennial scale variability). They differ in both the shape and strength of the low-frequency variability before the 13th century. The BHMs suggest a modest warming and relatively weak temperature fluctuation during the medieval time while Ge2013, Cook2013 and Yang2002 suggest that a few decades during the MCA were as warm as the 1990s. Generally, the proxy data networks become sparser back in time and the reconstructed variance is increasingly sensitive to the influence of individual proxy records. The strong warming and temperature fluctuation of Ge2013 before the 15th century seem to be mainly influenced by two low-resolution proxies ( $>10a$ , “Hong2000” and “Ge2003”) over eastern China<sup>26,28</sup>. It is difficult to estimate the biases in the two proxies but they might be over-weighted when we



**Figure 5.** Simulated and reconstructed East Asia warm-season land temperature anomalies (with respect to 1500–1850 CE) for the last 1200 yr (850–1980; The decadal-resolute reconstructions are 3 decades moving average, and the annual model data are 31-yr filtered). BHM59 reconstructed temperature are shown dark green while BHM13 appear in light green over the spread of model run. The ensemble mean (heavy black line) and the band accounting for 50% and 80% (shading) of the spread are shown for the model ensemble.

consider the spatial decorrelation length of temperature. The better spatial coverage in the proxy network applied for the BHM59 may provide a more robust temperature estimation in this early period compared to Ge2013.

The discrepancies between the reconstructions may be from different (but partly overlapping) proxy networks, different data selection strategies, and different reconstruction methods. But most important is the data selection step. And from this perspective, Ge2015 and our BHM59 are probably more reliable than the others. Looking into the sub-regional temperatures reconstructed in Ge *et al.*<sup>26,28</sup>, we find agreement with our study: western China experienced warmer conditions in the 1990s compared to any decade of the MCA. Further, eastern China may have experienced similar warm conditions climate during some decades of the MCA.

Shi2015 and Cook2013 are based on climate field reconstructions. Cook *et al.*<sup>9</sup> used only tree-ring chronologies mainly from high latitudes or western mountainous areas. Shi *et al.*<sup>10</sup> employed tree-ring-dominant (>93%) proxy records (all 341 tree-ring chronologies provided by Cook *et al.*<sup>9</sup> and another 51 tree-ring chronologies obtained through private communication). As tree-ring growth is influenced by several environmental factors during the growing season, it is necessary to check their temperature sensitivity. Note that, most of the tree ring chronologies were also used to reconstruct the PDSI (Palmer Drought Severity Index) over monsoon Asia<sup>29</sup>. Cook2013 was pointed out as having “substantial uncertainties in low frequencies” by the PAGES 2K Consortium<sup>5</sup>. This shortcoming is addressed by removing tree-ring chronologies with low temperature sensitivity and by adding data from other proxy archives. Yang2002 and Ge2013 are regional average temperature reconstructions based on the composite methods using both a smaller number of proxies.

**Comparison with GCM simulations.** Figure 5 shows the BHM59 reconstructions and the median of the ensemble of simulations along the 10, 25, 75 and 90 percentile uncertainties based on the ensemble spread<sup>30</sup>. This spread reflects differences in the models and the individual forcing implementation and also internal variability in decadal to multi-decadal temperature variations. The latter can be seen in multiple realizations of the same model and it has been shown that individual ensemble members can deviate substantially from the ensemble mean even under strong external forcing<sup>31</sup>. Reconstructions and simulations are filtered using a 31-yr moving average filter to highlight the low-frequency variability and to compare temperature variations influenced by external forcing.

The reconstruction for the East Asian temperatures correlates well with the median of the ensemble of simulations ( $r = 0.77$ ,  $df = 113$ ,  $p < 0.01$ ). The external forcing used in the model experiment explains a significant fraction of East Asia warm season temperature variations during 850–2000 CE. The imprint of large volcanic eruptions (e.g. Samalas 1257 CE, Kuwae 1453 CE) is clearly visible, where simulations tend to show stronger short-term responses than reconstructions and instrumental observations<sup>32</sup>. Changes in solar activity can directly influence earth climate via thermal processes, changes in tropical convective activity or indirectly by changes in ozone photochemistry and related changes in stratospheric circulation eventually propagating into the troposphere<sup>33,34</sup>. Due to the comparatively small changes in solar activity during the last millennium the impact on climate might be seen as being supportive in amplifying prolonged cool phases in conjunction with increased volcanic activity. Solar forcing<sup>33,34</sup> likely contributed to episodes of cooling during periods of low solar irradiance (Fig. S8): the Oort Minimum (ca. 1040–1080 CE), the Wolf Minimum (ca. 1280–1350 CE), the Spörer Minimum (ca. 1460–1550 CE), the Maunder Minimum (ca. 1645–1715 CE) and the Dalton Minimum (ca. 1790–1820 CE). However, as some solar minima coincide with strong volcanic events, it is difficult to distinguish their influence on multi-decadal temperature variability from the effects of solar forcing.

The simulations and the reconstruction agree on the timing and duration of the Medieval Climate Anomaly (MCA, ca. 900–1200 CE) and the Little Ice Age (LIA, ca. 1450–1850 CE) as well as the amplitude of the low-frequency variability. A few cooling intervals, possibly being due to low-solar-activity, in the simulations cannot be found in our East Asia regional average reconstruction. For instance, the simulated temperature minimum

over the late Maunder Minimum (late 17<sup>th</sup> century/early 18<sup>th</sup> century) is not recorded in proxy records over East Asia; the simulated cold period during the MCA, attributed to the Oort Minimum, appeared over the Tibetan Plateau but not in the other regions of East Asia (see Fig. 3). Assuming the BHM reconstruction is the best currently available reconstruction, it suggests that the influence of solar forcing on regional decadal temperature variation may be much more heterogeneous than ensemble means of the GCMs. The more homogeneous temperature patterns in response to forcing in the model simulations have also been demonstrated in the study from PAGES 2k-PMIP3 group<sup>35</sup>.

The simulations and the reconstruction suggest a similar warming rate over the period 1850–1950 CE (0.04 °C/decade) and the temperature over the last thirty years of the 20<sup>th</sup> century is likely the warmest 30-years in the last 1150 years. However, a few differences are noticeable: Firstly, the reconstruction suggests that the rapid warming started around mid of 19<sup>th</sup> century<sup>36</sup>, that is two to three decades later than in the simulations (in which the warming started in the 1820s, the timing for the peak of Dalton minimum and Tambora eruption<sup>33,34,37</sup>). Secondly, there was no obvious multi-decadal time warming pause at the end of 19<sup>th</sup> century in the reconstruction. Thirdly, the simulations do not show a warming pause during 1940–1970 CE, which can be seen in both the reconstruction and the instrumental temperature (Fig. 2). The simulated rapid warming is mainly attributed to the greenhouse gases (see Model settings in Table S2). However, anthropogenic aerosols might have also played an important role in the temperature variability in the second half of the 20<sup>th</sup> century<sup>38</sup>. We note that, for most of the pre-industrial last millennium (850–1850 CE), the BHM reconstruction lies within the multi-model ensemble spread (Fig. 5). Two exceptions are seen: The cooling at the end of the 16<sup>th</sup> century is slightly larger in the reconstruction (possibly due to a cluster of volcanic eruptions). The reconstruction also suggests a warmer interval at the turn of the 15<sup>th</sup> century while the simulations show a cooler climate driven by low-solar forcing.

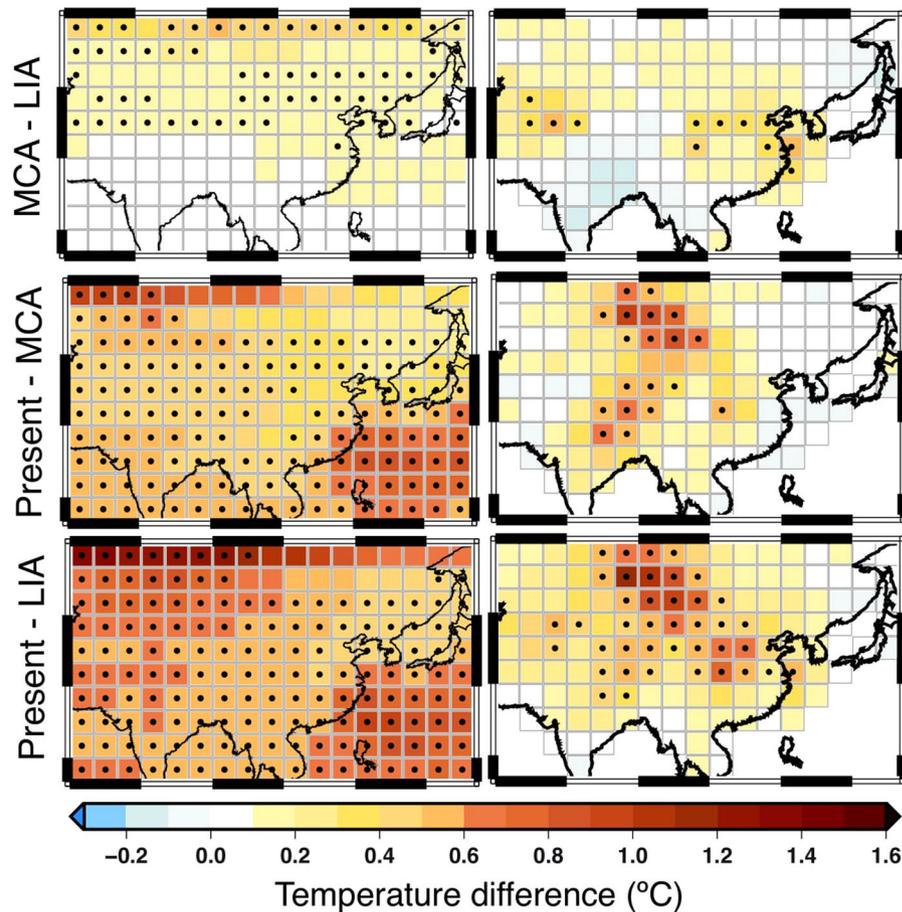
Similar to<sup>6</sup>, we compared the simulations with the reconstruction on the spatial temperature differences for the transitions: MCA (900–1200 CE) minus the LIA (1450–1850 CE), present-day (1950–2000 CE) minus the MCA and present-day minus the LIA (Fig. 6). The simulated differences are statistically significant at the 5% level at most of the grid-points, indicating a high agreement among the model ensemble members and suggesting that changes in external forcing have had a prominent influence on past East Asia warm-season temperature variations. Both the simulation ensemble mean and the reconstruction suggest that southeast China and the region north of 30° N experienced warmer climate during the MCA than during the LIA (top panels in Fig. 6). Moreover, the reconstruction and the simulation agree on the warmer conditions during the present time compared to the LIA (mid panels in Fig. 6). Disagreements between simulations and reconstructions can be found for the temperature difference between present-day and MCA (bottom panels in Fig. 6). The simulations show warmer conditions over the whole domain, similar to the temperature differences between the present-time and the LIA but with smaller anomalies. Our reconstruction indicates that the warming may be more spatially heterogeneous, i.e. warmer conditions over the middle part of the domain and cooler conditions over the lower reaches of the Yangtze River. Regional-scale differences are often observed comparing proxy-based climate reconstructions and model simulations<sup>35</sup>. The reconstructions are based on “real world” proxies, which allow an assessment of the true climate trajectory (taking into account the associated uncertainties in the reconstruction). The regional temperature series derived from climate simulations represent single trajectories that are consistent with changes in external forcings (orbital, volcanic, solar, greenhouse gases and land use) applied to the models as boundary forcing. The presence of internal climate variability, uncertainties in the external forcings as well as model errors limit temporal consistency between individual model trajectories and empirical reconstructions. For continental scales, the impact of changes in external forcings should be large enough to allow for a first order comparison. Still some characteristics (monsoonal changes in tropics, changes in extratropical circulation, changes in sea surface temperatures in adjacent continental areas etc.) can modify the direct impact of external forcings on East Asian temperatures in both, empirical reconstructions and model simulations.

We present a novel climate field reconstruction for East Asia using selected multi-proxy evidence. We extend available reconstructions to 2000 years and also assess uncertainties by applying a stochastic model for a multi-proxy network over a larger area. It generally compares well with the ensemble model simulation for the last 1150 years, which not only assigns the climate models confidence but also offers physical explanation for the past temperature evolution in the reconstruction. This gridded product will be used to analyze spatial modes of long-term temperature variability within the East Asia area and further compare with the simulations under various individual forcing, with the aim to assess for example the sub-regional effects of volcanic eruption, solar- and land-use change etc.

## Data and Methods

**Instrumental data.** The instrumental temperature field used for the calibration of the Bayesian hierarchical model was derived from the global CRUTEM4v dataset, hereafter referred to as CRU<sup>39</sup>. This dataset has been used to calibrate proxy data in previous works (e.g. a European summer temperature reconstruction<sup>6</sup>).

This dataset comprises monthly mean temperature anomalies (with respect to 1961–1990 CE) with 5.0° × 5.0° resolution spanning the period 1850–2010 CE, but only data from the year 1921 onward have been used in this study as most meteorological stations are established after 1921 CE over East Asia. The region (60° – 160° E/10° – 60° N) comprises in total 132 grid points, excluding grid cells over the ocean or islands. As in<sup>6</sup>, gaps with missing observations were filled using the Regularized Expectation Maximization algorithm with ridge regression<sup>40</sup> (RegEM). Then the warm season (May–September) mean temperature at each grid point and each decade was calculated using the infilled dataset for the comparison with the reconstructions. Non-infilled CRU was used as input data for the BHM reconstructions. Regional anomalies (w.r.t. 1961–1990 CE) of warm season temperature over East Asia using non-infilled CRU and infilled CRU are represented in Fig. S1 including the decadal-mean time series derived from the infilled CRU data.



**Figure 6.** Simulated (left) and reconstructed (right) warm season (May–September) temperature differences for three periods: MCA (900–1200) minus LIA (1450–1850); present-day (1950–2000) minus MCA; and present-day minus LIA. Model temperature differences indicate average temperature changes in the ensemble of available model simulations. Reconstructed temperature differences with the BHM methods with full proxy network are shown in the right column. Simulations have been weighted by the number of experiments considered from each model (see Table S2 in the supplementary material). Dots indicate significant ( $p < 0.1$ ) changes in the reconstructions; in the simulation ensemble a dot indicates at least 80% of agreement in depicting significant ( $p < 0.05$ ) changes of the same sign. Figure was plotted using Generic Mapping Tools (<http://gmt.soest.hawaii.edu/>).

Meteorological station data were used to select temperature-sensitive tree-ring proxy records. The station dataset includes monthly temperature data from 710 stations covering the period 1961–2000 CE, and the data were collected and processed by the National Meteorological Information Center of the China Meteorological Administration.

**Proxy data.** Proxy data with at least decadal resolution were used, including tree-ring widths (TRW), tree-ring maximum latewood density (MXD), tree-ring isotopes, historical documentary data, ice cores, speleothems, lake sediments, and marine data. TRW and MXD proxy data were mainly taken from PAGES2k<sup>5,9</sup> and a few MXD records were from the International Tree Ring Data Bank (ITRDB, <http://www.ncdc.noaa.gov/paleo/treering.html>).

TRW growth, and thus the TRW, is influenced by several environmental factors during the growing season, such as temperature and/or precipitation in water- and/or energy limited regions<sup>41</sup>. In contrast, MXD is generated at the end of the growing season and thus should not reflect a precipitation signal<sup>42</sup>. However, they still need to be checked for temperature sensitivity. To make sure the MXD and TRW data used for the reconstruction are temperature sensitive, the following steps were carried out:

- (1) The tree-ring records located over China or at the Chinese border were compared with the closest meteorological station data within 500 km at an annual scale covering the period 1961–1990 CE (This period is chosen while most meteorological stations over Tibetan Plateau are established after the late 1950s, and lots of TR proxy records do not cover the 1990s). Meteorological station data rather than CRU were used for the proxy-data screening because we found the tree-ring proxy records have a higher cross-correlation with near-by station data. A proxy record is retained when the Pearson correlation coefficient with the station data is higher than 0.25 (the critical value at one-tail 10% significance level for 30 years).

- (2) The tree-ring records without a nearby meteorological station were compared with CRU gridded data for the periods 1921–2000 CE and 1961–1990 CE. The proxy record is retained when one of the correlation coefficients is significant (higher than the critical value at a one-tail 10% significance level).
- (3) In data-rich regions, i.e. regions where more than one proxy record is available within 50 km, only a single record is kept. We chose the record that has higher correlation with nearby records within 500 km over the last 500 years and a longer record length. Thus, we avoid involving too much local climate information which would decrease the spatial correlation length as well as overweighting regions with many proxies compared to regions that contain only sparse proxy information.

After the outlined procedure, 29 TRW and eight MXD chronologies are kept. Non-tree-ring proxy records were selected based on published articles, which have been used by Ge *et al.*<sup>26</sup> or Ljungqvist *et al.*<sup>43</sup> (see Table S1). Together with one tree-ring isotope record, twelve historical documentary series, four ice cores, three speleothems, three lake sediments, and one marine record, we have in total sixty-one warm-season temperature-sensitive proxy data from different regions over East Asia (Table S1). Their locations are shown in Fig. 1. Here, we acknowledge a potential uncertainty generated from different optimal seasons some proxy records have.

After the selection, all of the proxy records at annual resolution were averaged to decadal scale. In this study, we reconstructed decadal-resolution temperature rather than annual-resolution temperature as our proxy records with annual resolution are largely restricted to the western mountainous area or high-latitude areas.

**Global climate model (GCM) simulation data.** We adopted the eleven millennium-length simulations performed with nine state-of-the-art GCM following the PMIP3/CMIP5 standards<sup>34,44</sup> (see Table S2). Their regional temperature simulations over different continents have been compared with reconstructions<sup>35</sup>. The geographical window used in this work to represent the area of East Asia is 60°–160° E/10°–60° N. The variable considered is the 2 m air temperature. The simulated monthly series have been prepared to mimic the resolution and seasonality of the reconstruction. Also, all simulations have been remapped to a common 5° × 5° grid for the spatial comparison.

**Bayesian hierarchical model (BHM).** In the frame of non-regression approaches, BHM interprets surface temperature anomalies as a first-order auto-regression (AR1) process in time with exponentially decaying spatial covariance. In this study, instrumental data were modeled as temperature anomalies with white noise; proxy records like documentary data were modeled as a linear function of temperature anomalies plus a white noise term; and proxy records like TRW which exhibit strong persistence<sup>45</sup> were modeled in three terms: a linear function of temperature anomalies, a AR1 process term (describing that the tree ring growth in the present-year is associated with the one in the previous-year) and a white noise term. To test the robustness of BHM field reconstructions and the corresponding regional average temperature time series, we generated two ensembles of CFRs using the full proxy network (59 proxies) and a sub dataset (denoted here Frozen1000) including only 13 proxy records which cover the whole last millennium (see Table S1). Each ensemble contains 1000 members and each member is one CFR going back to the beginning of the Common Era. The CFR members were taken from the predictive run (without instrumental data as input), which enables the comparison between the reconstruction and the instrumental data. The regional average temperature reconstructions generated from the two ensembles generally agree with each other (see “Bayesian hierarchical model” section in the supplementary file for more about the method, its implementation and related results). The spatial mean time series for each ensemble member was computed by averaging area-weighted (weighted by multiplying the cosine of latitude) values in relevant grid boxes. The temperature anomaly estimation for the ensemble median was calculated in two ways: 1) the point-wise median ensemble of reconstructions and the 90 percentiles confidence level, which takes the 50<sup>th</sup> percentile at each decade as the best estimate of temperature anomaly (the median) and the 5<sup>th</sup> and 95<sup>th</sup> percentiles produce 90% point-wise uncertainty envelopes; and 2) the path-wise median ensemble of reconstructions and the 90 percentiles confidence intervals, which are addressed in “*Deepest curve and path-wise confidence intervals*”.

**Composite method.** The proxy data were first standardized individually, and then the average was calculated to present regional temperature variations. This method could sufficiently reproduce large-scale temperature variations out of small and heterogeneous datasets<sup>27,46,47</sup>.

**Deepest curve and path-wise confidence intervals.** As proxy records and reconstructions normally have high autocorrelation<sup>45,48</sup>, we considered each ensemble member as a function generated by continuous smooth dynamics. We drew information from a set of functional data by applying function data analysis. As an important concept to sort functional data, data depth introduces a measure to rank a sample of function data from the center outwards. In general, each ensemble member was drawn as a continuous curve. Then we computed the depth values of all the curves and ranked the curves by decreasing depth values following a modified half region depth approach<sup>49</sup>, which defines band depth by taking a graph-based approach.

Each ensemble member was considered as a function  $y_i(t)$ ,  $i = 1, \dots, 1000$ , generated by the stochastic process  $Y$ . The depth of  $y$  was measured as

$$H(y) = \min\{SL(y), IL(y)\}. \quad (1)$$

$SL$ , the superior lengths, indicates the “proportion of time” that the stochastic process  $Y$  is greater than  $y$ ; and  $IL$ , the inferior length, indicates the “proportion of time” that the stochastic process  $Y$  is smaller than  $y$ . They can be calculated in the following manners:

$$SL(y) = \frac{1}{n\lambda(I)} \sum_{i=1}^n \lambda\{t \in I: y(t) \leq y_i(t)\},$$

$$IL(y) = \frac{1}{n\lambda(I)} \sum_{i=1}^n \lambda\{t \in I: y(t) \geq y_i(t)\} \quad (2)$$

where  $\lambda$  stands for the Lebesgue measure on  $\mathbf{R}$ .

Then  $y_i(t)$ ,  $i = 1, \dots, 1000$  can be ranked as  $y_{[1]}(t), \dots, y_{[1000]}(t)$  according to decreasing values of  $H(y_i)$ , where  $y_{[1]}(t)$  denotes the deepest curve or simply the median curve and  $y_{[1000]}(t)$  the most outlying curve. The deepest curve has the shortest distance to the mean. We used this concept to find the most central ensemble member  $y_{[1]}(t)$  and let it represent the median trajectory of reconstructed temperature variations. The upper and lower boundaries of the band that is bounded by the curves  $y_{[1]}(t), \dots, y_{[900]}(t)$  are used as path-wise 90% confidence intervals. Estimating uncertainties in a path-wise manner enabled us to judge whether one event is an outlier from the perspective of the whole time period rather than at one specific time interval. The concept of path-wise uncertainties was adopted in former paleo-climate reconstructions<sup>19,50</sup>.

**Robustness test of the regional average temperature reconstruction.** We applied two strategies to generate different regional average temperature reconstructions. The first one is to use the same data network (the full data network) but two different methods (BHM and composite method, the particularities of both methodologies has been briefly addressed above and more details can be found in the supplementary file). The second approach makes use of the same method, BHM, but two different data networks (the full data network versus “Frozen1000”). The four regional average temperature reconstructions generated as explained before are shown in Figs S2, S3. Correlation analysis among the four reconstructions shows that they are highly correlated with each other ( $r > 0.86$ ,  $p < 0.01$ ). They all indicate a multi-century warm period around 900–1200 CE followed by a multi-century long cool period around 1450–1850 CE and a continuous warming tendency after the 1850s decade with the modern warming during the second half of 20<sup>th</sup> century. The temperature in the 1990s is seen noticeably higher than any previous warm conditions.

**Gridded temperature verification.** The null hypothesis,  $H_0$ , states that the observed temperature anomaly is well “predicted” by the reconstructions against the alternative hypothesis,  $H_A$ , that the observed temperature anomaly is far different from the reconstructions. The null hypothesis is tested for each decade at each grid point for the period 1921–2000 CE. Let  $T$  denotes the reconstructed temperature and  $T_{CRU}$  the observed temperature, assuming that the observed record only falls at the tail of the empirical distribution of the 1000 ensemble members. The corresponding  $P$  value is calculated as following<sup>51</sup>, which indicates the proportion of reconstructed temperature anomalies which are more extreme than the observed temperature anomalies.

$$p = 2 \min \left( \frac{1}{1000} \sum_{i=1}^{1000} I(T_i < T_{CRU}), \frac{1}{1000} \sum_{i=1}^{1000} I(T_i > T_{CRU}) \right) \quad (3)$$

When this  $P$  value is sufficiently small, we reject  $H_0$ . Then we follow the false discovery rate procedure<sup>52,53</sup> to reject  $H_0$  at all locations  $i$  for which  $p_i \leq p_k$ , where

$$k = \max_{i=1, \dots, n} \left\{ i: p_{[i]} \leq q \frac{i}{n} \right\} \quad (4)$$

with  $p_{[i]}$ ,  $i = 1, \dots, n$  are ranked  $p$ -values in an ascending order.

The field reconstruction has been verified against the instrumental record at each grid point and for each decade during 1921–2000 CE. Only for a few grid cells, the reconstruction is statistically significantly different from the instrumental temperature target (Fig. S4).

## References

1. Masson-Delmotte, V. *et al.* Information from Paleoclimate Archives In: *Climate Change 2013: The Physical Science Basis Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge University Press, 2013).
2. Christiansen, B. & Ljungqvist, F. C. Challenges and perspectives for large-scale temperature reconstructions of the past two millennia. *Reviews of Geophysics*, 40–96 (2017).
3. Anchukaitis, K. J. *et al.* Last millennium Northern Hemisphere summer temperatures from tree rings: Part II, spatially resolved reconstructions. *Quaternary Science Reviews* **163**, 1–22 (2017).
4. Wilson, R. *et al.* Last millennium northern hemisphere summer temperatures from tree rings: Part I: The long term context. *Quaternary Science Reviews* **134**, 1–18 (2016).
5. PAGES 2k Consortium. Continental-scale temperature variability during the last two millennia. *Nat. Geosci.* **6**, 339–346 (2013).
6. Luterbacher, J. *et al.* European summer temperatures since Roman times. *Environmental Research Letters* **11**, 024001 (2016).
7. Ge, Q. S. *et al.* Coherence of climatic reconstruction from historical documents in China by different studies. *Int. J. Climatol.* **28**, 1007–1024 (2008).
8. Ge, Q., Liu, H., Ma, X., Zheng, J. & Hao, Z. Characteristics of temperature change in China over the last 2000 years and spatial patterns of dryness/wetness during cold and warm periods. *Adv. Atmos. Sci.* **34**(8), 941–951 (2017).
9. Cook, E. *et al.* Tree-ring reconstructed summer temperature anomalies for temperate East Asia since 800 C.E. *Clim. Dynam.* **41**, 11–12 (2013).
10. Shi, F. *et al.* A multi-proxy reconstruction of spatial and temporal variations in Asian summer temperatures over the last millennium. *Clim. Change* **131**, 663–676 (2015).

11. Smerdon, J. E., Kaplan, A. & Chang, D. On the origin of the standardization sensitivity in RegEM climate field reconstructions. *Journal of Climate* **21**, 6710–6723 (2008).
12. Luterbacher, J., Dietrich, D., Xoplaki, E., Grosjean, M. & Wanner, H. European seasonal and annual temperature variability, trends, and extremes since 1500. *Science* **303**, 1499–1503 (2004).
13. Xoplaki, E. *et al.* European spring and autumn temperature variability and change of extremes over the last half millennium. *Geophys. Res. Lett.* **32**, L15713 (2005).
14. Riedwyl, N., Küttel, M., Luterbacher, J. & Wanner, H. Comparison of climate field reconstruction techniques: Application to Europe. *Climate Dyn.* **32**, 381–395 (2009).
15. Mann, M. E. *et al.* Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia. *P. Natl. Acad. Sci.* **105**, 13252–13257 (2008).
16. Smerdon, J. E., Kaplan, A., Chang, D. & Evans, M. N. A Pseudoproxy Evaluation of the CCA and RegEM Methods for Reconstructing Climate Fields of the Last Millennium. *J. Climate* **23**, 4856–4880 (2010).
17. Guillot, D., Rajaratnam, B. & Emile-Geay, J. Statistical Paleoclimate Reconstructions via Markov Random Fields. *The Annals of Applied Statistics* **9**(1), 324–352 (2015).
18. Tingley, M. P. & Huybers, P. A Bayesian algorithm for reconstructing climate anomalies in space and time. Part 1. *Development and applications to paleoclimate reconstructions problems. J. Climate* **23**, 2759–2781 (2010).
19. Tingley, M. P. & Huybers, P. Recent temperature extremes at high northern latitudes unprecedented in the past 600 years. *Nature* **496**, 201–205 (2013).
20. Werner, J. P., Divine, D. V., Ljungqvist, F. C., Nilsen, T. & Francus, P. Spatio-temporal variability of Arctic summer temperatures over the past 2 millennia. *Clim. Past* **14**, 527–557, <https://doi.org/10.5194/cp-14-527-2018> (2018).
21. Tingley, M. P. & Huybers, P. A Bayesian algorithm for reconstructing climate anomalies in space and time. Part 2. Comparison with the Regularized Expectation-Maximization Algorithm. *J. Climate* **23**, 2782–2800 (2010).
22. Werner, J. P., Luterbacher, J. & Smerdon, J. E. A Pseudoproxy Evaluation of Bayesian Hierarchical Modelling and Canonical Correlation Analysis for Climate Field Reconstructions over Europe. *J. Climate* **26**, 851–867 (2013).
23. Wang, J., Yang, B. & Ljungqvist, F. C. A millennial summer temperature reconstruction for the Eastern Tibetan Plateau from tree-ring width. *Journal of Climate* **28**, 5289–5304 (2015).
24. Cai, D., You, Q., Fraedrich, K. & Guan, Y. Spatiotemporal temperature variability over the Tibetan Plateau: Altitudinal dependence associated with the global warming hiatus. *J. Climate* **30**, 969–983 (2017).
25. Davi, N. K. *et al.* A long-term context (931–2005 CE) for rapid warming over central Asia. *Quaternary Science Reviews* **121**, 89–97 (2015).
26. Ge, Q., Hao, Z., Zheng, J. & Shao, X. Temperature changes over the past 2000 yr in China and comparison with the Northern Hemisphere. *Climate of the Past* **9**, 1153–1160 (2013).
27. Yang, B., Braeuning, A., Johnson, K. R. & Shi, Y. F. General characteristics of temperature variation in China during the last two millennia. *Geophys. Res. Lett.* **29**, 1324 (2002).
28. Ge, Q. S. *et al.* Temperature variation through 2000 years in China: An uncertainty analysis of reconstruction and regional difference. *Geophys. Res. Lett.* **37**, L03703, <https://doi.org/10.1029/2009GL041281> (2010).
29. Cook, E. R. *et al.* Asian monsoon failure and megadrought during the last millennium. *Science* **328**(5977), 486–489 (2010).
30. Jansen, E. *et al.* “Palaeoclimate.” Climate Change 2007: *The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge University Press, 2007).
31. Otto-Bliesner, B. *et al.* Climate Variability and Change since 850 C.E.: An Ensemble Approach with the Community Earth System Model (CESM). *Bull. Amer. Meteor. Soc.* **97**, 735–754 (2016).
32. Brohan, P. *et al.* Constraining the temperature history of the past millennium using early instrumental observations. *Clim. Past* **8**, 1551–1563 (2012).
33. Schmidt, G. A. *et al.* Climate forcing reconstructions for use in PMIP simulations of the last millennium (v1.0). *Geoscientific Model Development* **4**(1), 33–45 (2011).
34. Schmidt, G. A. *et al.* Climate forcing reconstructions for use in PMIP simulations of the last millennium (v1.1). *Geoscientific Model Development* **5**, 185–191 (2012).
35. 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).
36. Abram, N. J. *et al.* & PAGES 2k Consortium. Early onset of industrial-era warming across the oceans and continents. *Nature* **536**, 411–418 (2016).
37. Sigl, M. *et al.* Timing and climate forcing of volcanic eruptions for the past 2,500 years. *Nature* **523**(7562), 543–549 (2015).
38. Folini, D. & Wild, M. The effect of aerosols and sea surface temperature on China's climate in the late twentieth century from ensembles of global climate simulations. *J. Geophys. Res. Atmos.* **120**, 2261–2279 (2015).
39. Jones, P. D. *et al.* Hemispheric and large-scale land-surface air temperature variations: An extensive revision and an update to 2010. *J. Geophys. Res.: Atmospheres* **117**, D5 (2012).
40. Schneider, T. Analysis of incomplete climate data: estimation of mean values and covariance matrices and imputation of missing values. *J. Clim.* **14**, 853–871 (2001).
41. Briffa, K. R. *et al.* Tree-ring width and density data around the Northern Hemisphere: Part 1, local and regional climate signals. *Holocene* **12**, 737–757 (2002).
42. Bradley, R. S. *Paleoclimatology: reconstructing climates of the Quaternary 400–403* (Academic Press, 1999).
43. Ljungqvist, F. C. *et al.* Northern Hemisphere hydroclimate variability over the past twelve centuries. *Nature* **532**(7597), 94–98 (2016).
44. Taylor, K. E., Stouffer, R. J. & Meehl, G. A. An Overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.* **93**, 485–498 (2012).
45. Zhang, H. *et al.* Modified climate with long term memory in tree ring proxies. *Environmental Research Letters* **10**(8), 084020 (2015).
46. Crowley, T. J. & Lowery, T. S. How warm was the Medieval Warm Period? A comment on ‘Man-made versus natural climate change?’. *Ambio* **39**, 51–54 (2000).
47. Jones, P. D., Osborn, T. J. & Briffa, K. R. Estimating sampling errors in larger-scale temperature averages. *J. Climatol.* **10**, 2548–2568 (1999).
48. Franke, J., Frank, D., Raible, C. C., Esper, J. & Brönnimann, S. Spectral biases in tree-ring climate proxies. *Nature Climate Change* **3**(4), 360–364 (2013).
49. López-Pintado, S. & Romo, J. On the concept of depth for functional data. *Journal of the American Statistical Association* **104**, 718–734 (2009).
50. McShane, B. B. & Wyner, A. J. A statistical analysis of multiple temperature proxies: are reconstructions of surface temperatures over the last 1000 years reliable? *Ann. Appl. Statist.* **5**, 5–44 (2011).
51. Davidson, R. & MacKinnon, J. G. The power of bootstrap and asymptotic tests. *Journal of Econometrics* **133**, 421–441 (2006).
52. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Royal Stat. Soc. B*, 289–300 (1995).
53. Ivanov, M., Warrach-Sagi, K. & Wulfmeyer, V. Field significance of performance measures in the context of regional climate model evaluation. Part 1: temperature. *Theor Appl Climatol.* **1**–19 (2017).

## Acknowledgements

The authors thank the reviewers for their constructive criticism and important suggestions. The authors also thank very much the proxy and instrumental data provider. Huan Zhang, Jürg Luterbacher, Johann H. Jungclaus, Sebastian Wagner and Eduardo Zorita acknowledge support from the German Science Foundation project “Attribution of forced and internal Chinese climate variability in the common eras”. Jürg Luterbacher, Lea Schneider, Stefanie Talento, Jianglin Wang and Bao Yang acknowledge JPI-Climate/Belmont Forum collaborative Research Action “INTEGRATE, An integrated data-model study of interactions between tropical monsoons and extratropical climate variability and extremes”. Jürg Luterbacher acknowledges the Climate Science for Service Partnership China project (CSSP): Digitisation and quality control of subdaily meteorological data from Asian stations in the late 19th and early 20th century. Fidel J. González-Rouco acknowledges support from ILMoelS CGL2014–59644-R. Elena García-Bustamante thanks the European Framework programm FP7 ERA-NET project “NEWA: New European Wind Atlas”, funded by the European Commission. The reconstructions can be downloaded from the NOAA paleoclimate homepage: <https://www.ncdc.noaa.gov/paleo/study/23491>.

## Author Contributions

J.L., J.W. and E.G. planed the project. H.Z. collected the proxy records and performed the analyses and generated the reconstruction with support from J.W. and J.L. E.G. and F.G. performed the analyses with model simulations. H.Z. wrote the initial version with help from J.W., J.L., E.G., E.X., F.G., S.W., E.Z., K.F., J.H.J., F.C.L. and X.Z. All authors discussed the results and commented on the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-26038-8>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018