CrossMark

ORIGINAL PAPER

Multiscale combination of climate model simulations and proxy records over the last millennium

Xin Chen[1,2] · Pei Xing[3] · Yong Luo[1,2] · Suping Nie[4] · Zongci Zhao[1,2] · Jianbin Huang[1,2] · Qinhua Tian[4]

Abstract To highlight the compatibility of climate model simulation and proxy reconstruction at different timescales, a timescale separation merging method combining proxy records and climate model simulations is presented. Annual mean surface temperature anomalies for the last millennium (851–2005 AD) at various scales over the land of the Northern Hemisphere were reconstructed with $2° × 2°$ spatial resolution, using an optimal interpolation (OI) algorithm. All target series were decomposed using an ensemble empirical mode decomposition method followed by power spectral analysis. Four typical components were obtained at inter-annual, decadal, multidecadal, and centennial timescales. A total of 323 temperature-sensitive proxy chronologies were incorporated after screening for each component. By scaling the proxy components using variance matching and applying a localized OI algorithm to all four components point by point, we obtained merged surface temperatures. Independent validation indicates that the most significant improvement was for components at the inter-annual scale, but this became less evident with increasing timescales. In mid-latitude land areas, 10–30% of grids were significantly corrected at the inter-annual scale. By assimilating the proxy records, the merged results reduced the gap in response to volcanic forcing between a pure reconstruction and simulation. Difficulty remained in verifying the centennial information and quantifying corresponding uncertainties, so additional effort should be devoted to this aspect in future research.

✉ Yong Luo
Yongluo@mail.tsinghua.edu.cn

1 Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing, China

2 Joint Center for Global Change Studies, Beijing, China

3 Beijing Municipal Climate Center, Beijing, China

4 National Climate Center, China Meteorological Administration, Beijing, China

1 Introduction

Understanding spatiotemporal climate variation over the past centuries is extremely important for exploring the variability of the climate system at different timescales and corresponding responses to external forcings. Climate system model simulation and proxy records-based reconstruction both provide insights into the variation of spatial patterns and hemispheric and continental averages of climatic variables in past times (Jones and Mann 2004; Jones et al. 1998; Jones et al. 2009). The Coupled Model Intercomparison Project Phase 5 (CMIP5) and Paleoclimate Modeling Intercomparison Project Phase 3 (PMIP3) have gathered a large set of last millennium results from various climate system models and spatial resolutions. There has been substantial research into gridded surface temperature reconstruction over the last millennium based on different types of proxy records (Cook et al. 2013; Mann et al. 2009; Shi et al. 2012; Shi et al. 2015), using a class of methods referred to as climate field reconstruction (CFR). Both of these methodologies have laid a solid foundation for further exploration of past climate changes.

The methodology of CFR aims to reconstruct the spatial pattern and temporal evolution of climatic variables (e.g., temperature and precipitation) in the past, based on the screened proxy records (Cook et al. 1999, 2010; Tingley & Huybers, 2010a, b; Jones et al. 2009; Mann et al. 1998, 2005, 2008; Zhang et al. 2004). The problems with these methods have

been widely discussed (Jones et al. 1998; Jones et al. 2009), including limited spatial coverage of proxy data and decreasing temporal availability in earlier times. Other difficulties lie in the uncertainties associated with the calibration method, especially regarding tree ring-based reconstruction of low-frequency variations (Briffa et al. 1996; Christiansen 2011; Christiansen et al. 2009; Esper et al. 2002). State-of-the-art climate system models can not only provide regular gridded results for climatology but also physically coordinated simulated results. In addition, these well-developed, coupled climate system models can faithfully reproduce climate responses to external forcings, especially at greater spatiotemporal scales (Fernández-Donado et al. 2013). However, for smaller scales, there remain uncertainties, some of which originate from the impact of simulated internal variability (Deser et al. 2012; Deser et al. 2014). The internal variability in climate simulation can be comparable to climate variability at continental spatial scales and multidecadal centennial timescales (Goosse et al. 2005, 2012). However, this internal variability makes it more difficult to analyze actual climatic responses to external forcings at shorter timescales (Huber and Knutti 2014) and to identify significant response signals at global or regional scales (Jones and Mann 2004).

Given both of these approaches are associated with advantages and uncertainties, in recent years, more studies have emerged which apply data assimilation techniques to combine the paleo-reconstruction and physical models. Von Storch (2000) first suggested using data assimilation (DA) in paleoclimatology studies, and a method of Data Assimilation Through Upscaling and Nudging (DATUN) was outlined aiming to nudge the climate model closer to the reconstructions. Traditional DA approaches were applied to weather prediction, but since then, several methods have been proposed to assimilate the proxy records in either an online or offline way. Barkmeijer (2003) proposed a method forcing singular vectors that use the reconstructed atmospheric pattern index (e.g., north annular mode (NAM)) and simulated large-scale information to force the model forecasting. This was applied to examine the NAM variation from 1790 to 1820 (Widmann et al. 2010). Goosse et al. (2006) first applied the particle filter-based ensemble-selection approach to combine the selected simulation samples by weights related to the distance values of the proxy-based reconstructed data. However, simulation ensembles were constrained by the reconstructions. Some newly published research has made great incremental improvements on (1) adopting real reconstruction results, such as PAGES-2k reconstructed temperature series (Goosse 2016; Matsikaris et al. 2015); (2) applying forward modeling and sequential ensemble Kalman Filter techniques (Acevedo et al. 2016); and (3) considering time-averaged data assimilation algorithms, which were performed in an offline way (Steiger and Hakim 2016; Steiger et al. 2014). These studies treated the existing proxy reconstructions as direct

targets or constraints, so as to either nudge the model dynamically or recombine the simulated members. From the perspective of data assimilation techniques, the uncertainties of reconstruction were not considered specifically (or were considered in a static and simple way). The reported results of most recent studies were mainly based on the pseudo-proxy frame. However, it is important to consider the uncertainties of real proxy records. Although the newly developed merging scheme based on the optimal interpolation (OI) algorithm of Chen et al. (2015) considered uncertainties from the scaling process of tree ring chronologies with the variance matching (VM) method, there were two weaknesses of the combined approach: (1) merging was accomplished based on Community Climate System Model version 4 (CCSM4) simulation, and this method uses statistics of simulation results within a fixed time window to estimate the background error covariance matrix. In this way, climate variability within that window was included in the calculated background error covariance matrix (Evensen 2003); and (2) with the application of the OI algorithm in local form, the radius for searching nearby tree ring sites was also fixed which may blur the variability at different timescales. Therefore, these variabilities should be treated individually in an adaptive way, and the search radius should depend on the timescale in some way. Some new proxy-based temperature reconstruction studies with a scale decomposition method have provided greater insight into climate change at various timescales (Shi et al. 2011; Xing et al. 2016; Zheng et al. 2015).

In the present study, using proxy records, the merging method developed by Chen et al. (2015) is further improved in three regards: (1) the model-simulated temperature (background) field is replaced by fully forced members of a Last Millennium Ensemble (LME) simulation of the National Center for Atmospheric Research Community Earth System Model (CESM version 1.1) (Otto-Bliesner et al. 2015); (2) the proxy records must also be screened at various levels. An ensemble-based merging is done for the components of different timescales (inter-annual, decadal, multidecadal). Eventually, the corresponding variability for temperature evolution is obtained; and (3) the radius of local search, which determines the proxy records used for a certain grid, is quantified by multiple merging trails ahead of time. Verification tests were also conducted to show the potential application of this method.

This paper is organized as follows. Section 2 introduces (1) three datasets, instrument data, proxy records and climate model (CESM) simulation ensemble samples, together with some corresponding processes (e.g., using the RegEM algorithm to fill the proxy record within the same period as the instrument data span, and removing bias for the simulation ensemble members), and (2) a merging method based on the simulated ensemble members, including timescale decomposition and the OI algorithm. Since the latter algorithm is applied within a localized area, the scheme for quantifying the

appropriate search radius is also introduced in this section. Section 3 presents the validation results with independent testing carried out by separating the data into calibration and verification sets. Climate variability for different timescales of merged results was compared to analyze improvement relative to the original model simulation. Surface temperature response to external volcanic forcing is evaluated and discussed. In the final section, principal conclusions and prospects are presented.

2 Data and methods

### 2.1 Instrument data and proxy records

The Climatic Research Unit time-series (CRU TS v2.1), global gridded monthly $0.5° × 0.5°$ land surface temperature dataset (Mitchell and Jones 2005), from 1901 to 2000 was used to (1) calibrate the proxy records and (2) validate the merged results and quantify parameters within the merger frame. The annual mean surface temperature anomalies were extracted and interpolated onto $2° × 2°$ longitude-latitude target grids (blue dots in Fig. 1) with a bilinear scheme. The Berkeley Earth Surface Temperature dataset (BEST) (Rohde et al. 2013) was used to evaluate improvement in merged results and was interpolated by the same interpolation scheme onto $2° × 2°$ regular grids. Before using the interpolated dataset, anomalies were readjusted with respect to the reference period 1951–1980.

Three types of proxy records were collected, i.e., tree ring chronologies, lake sediment records, and ice core data. The candidates mainly came from the PAGES-2k network (Ahmed et al. 2013), among which only records reported as positively correlated with temperature were incorporated in this analysis. There were also other chronologies, gathered from the China Meteorological Data Sharing System Service and the literature. All these collected records were initially decomposed into typical components on inter-annual, decadal, multidecadal, and centennial timescales. Then, the correlation coefficients were calculated between these proxy components and the corresponding instrument components (i.e., CRU components). Only the components with significant correlations were adopted (see Sect. 2.3). A total of 323 proxy records were ultimately selected for merging (see Tables s1, s2, and s3 in the Supplementary Material), including 317 tree-ring (including types of total ring width and maximum latewood density), 6 ice core, and 2 lake sediment chronologies. Before merging, all these proxy records were normalized. The ice core and sediment records with lower temporal resolution were interpolated into annual resolution using spline functions. For convenient computing and comparison, we used the RegEM algorithm (Mann et al. 2008; Mann et al. 2009; Schneider 2001) to fill in missing values in proxy chronologies during the overlap period (1901–2000).

2.2 Model simulation dataset

We used ten fully forced runs from ensemble simulation results of the CESM (version 1.1) LME simulation (Otto-Bliesner et al. 2015) as the background field. In the LME experimental design, there are 30 members including ten runs with full forcing and other 20 runs driven by a single forcing
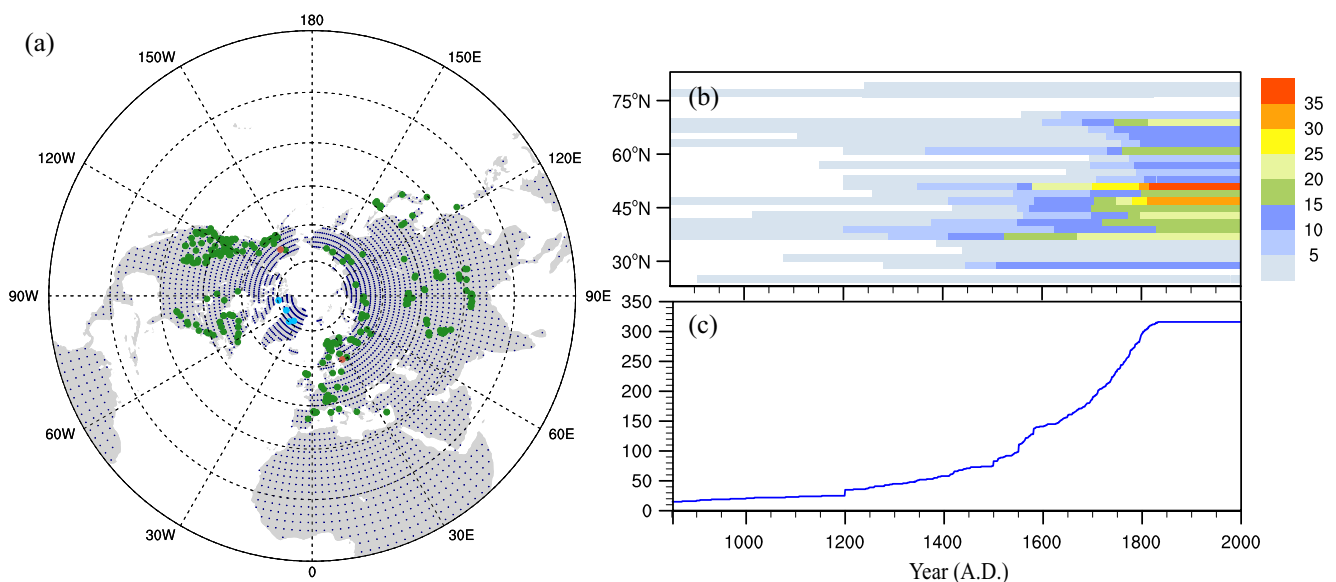
Fig. 1 a The target $2° × 2°$ degree grid over the land of the Northern Hemisphere (*dark blue dots*), tree rings (*green pies*), sediment records (*orange pies*), and ice core series (*light blue pies*). **b** The absolute number of proxy data for each $2°$ latitude belt of the Northern Hemisphere. **c** Availability of proxy records over the last millennium

(total solar irradiation, volcanic eruptions, greenhouse gases, land use and land cover, orbital variations and anthropogenic sulfate aerosols). The simulation period was 850 to 2005, and resolution was ~2° for atmosphere and land process components and ~1° for ocean and sea ice components. The difference from experiments with full forcing was in small, random perturbations imposed on an initial air temperature field. Similarly, annual mean surface temperature was extracted from the simulated monthly data, and then interpolated into a regularized 2° × 2° target latitude-longitude grid box (Fig. 1) to maintain consistency with the interpolated CRU dataset.

2.3 Multiscale approach and proxy record screening

For more effective use of the proxy records and extracting climate variability at different timescales, the ensemble empirical mode decomposition (EEMD) method was used in a hierarchical way (Huang and Wu 2008; Wang et al. 2010; Wu and Huang 2009). The EEMD method was developed from the empirical mode decomposition (EMD) method aiming to obtain the independent components at different typical periods. EEMD/EMD has been widely used in geo-scientific, climatic, and related studies, typically as a novel method to explore climate change at various timescales and extract intrinsic nonlinear trends (Franzke 2012; Franzke 2014; Ji et al. 2014). In this study, there were three main steps to decompose the target series into given timescales, and those series mainly refer to the model-simulated temperature series, instrument series on a specific grid, or a normalized proxy chronology series (Fig. 2a). In the first step, the target series was decomposed into several intrinsic mode functions (IMFs) and residuals. According to the EMD algorithms, for a particular time series, the number of IMFs (including the residuals) is certain, and this is determined by the series length. In the second step, power spectral analysis was applied to diagnose the main period for each IMF. For the final step, according to the diagnosed period, the obtained IMFs were reduced to four

components with typical timescale, inter-annual (<10 years), decadal (10–30 years), multidecadal (30–90 years), and centennial (>90 years). One target time series can be written as the sum of components (Eq. 1).

$$Ser_i(t) = \sum_{h=1}^{N_i} imf_h(t) + \text{resi} = \sum_{j=1}^{4} \text{comp}_j(t) \qquad (1)$$

where N_i is the total number of IMFs for the ith target time series [$Ser_i(t)$]. Subscript j denotes the timescale of each component [$comp_i(t)$]: $j = 1$ for annual scale, $j = 2$ for decadal scale, $j = 3$ for multidecadal scale, and $j = 4$ for centennial scale. These components of model simulation and instrument data were expected to contain more physical meaning and reveal climate change at multiscales. Therefore, the components at a certain timescale were used as merging candidates.

To construct the proxy networks for the four typical timescales, correlations between the proxy components (inter-annual, decadal, multidecadal, and centennial) and the counterparts of instrument data were estimated and corresponding significance was tested. Allowing for the completeness of instrument data and availability of proxy records over 1901–2000, we chose a 60-year period (1921–1980) to calculate correlation coefficients to screen the proxy components for all four timescales, using only those that passed significance tests. Given the reduced degrees of freedom for each series during screening, Monte Carlo red noise simulation (Mann et al. 2007; Mann et al. 2008) was used to quantify the threshold of correlation (see Appendix 1), and only the significant series were screened out. This approach can not only identify climate response recorded in proxy chronologies at varying temporal and spatial scales but also make better use of the proxy records for each component. The conventional scaling method of VM (Jones et al. 1998; Juckes et al. 2007; Lee et al. 2008; Shi et al. 2011) was used to scale the proxy components in order to transform the dimensionless indices to surface temperature components based on the interpolated CRU series at

Fig. 2 **a** Process of obtaining components of typical timescales by applying EEMD and power spectral analysis. **b** The main merging process to produce the EMM results. Together with the uncertainties estimated through a "leave-one-out" process, the reconstructed typical temperature components were used as inputs for the OI algorithm
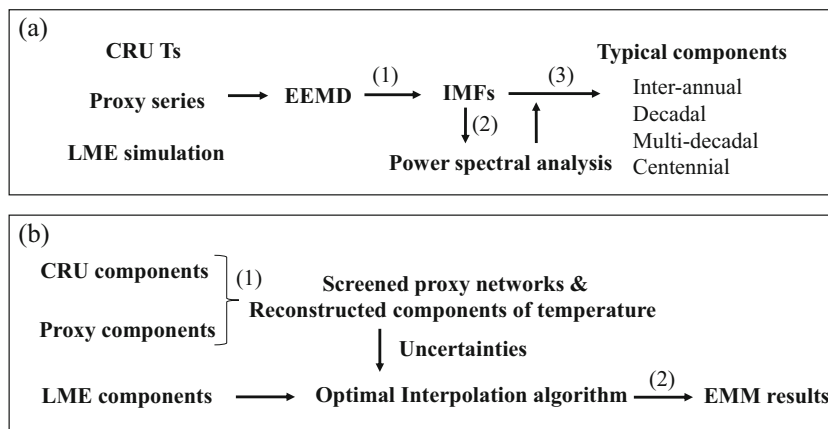
the nearest grid, following the EEMD process. The four typical proxy networks and scaled temperature components were used as "observation" candidates for the merging process. For each component, uncertainty was estimated with scaled verification results of the "leave-one-out" process (Shi et al. 2011). This type of uncertainty was further used as values of observation error variance matrix in the merging process.

2.4 OI algorithm and validation

The OI algorithm-based merging method (MM) for combining the climate model simulation and tree-ring chronologies was applied by Chen et al. (2015), where only one set of model simulation (CCSM4) (Landrum et al. 2013) was used to reconstruct surface temperatures over North America. Technically, the improvement of the OI algorithm applied in this study mainly lies in two aspects: (1) the estimation of the background error covariance matrix and (2) the searching radius for local proxy data. A background error covariance matrix was coarsely estimated in Chen et al. (2015) that simulated temperature within the fixed 30-year window chosen as an ensemble. Our current research gathered the ensemble of last millennium simulation (i.e., LME) and applied an ensemble-based merging method (EMM, Fig. 2b), for which the background error covariance matrix was calculated from the simulated ensemble members. For the components at various timescales, the matrices were computed independently and the structure was updated dynamically with time (see Appendix 2). By updating the optimal weights simultaneously, the OI algorithm (Gandin & Hardin 1965) was applied to merge the simulated ensemble mean and reconstructed components of four typical timescales individually. The purpose of EMM is to merge the climate model simulations with proxy data at multiscales, so as to construct a new platform to merge different types of proxy data. In addition, the improved OI algorithm was applied locally, i.e., within the area defined by a diagnostic "optimal" radius.

To seek an objective radius, we designed several groups of merging trials with linearly increasing distances in advance. These were set for different components to exhaustively discover the optimal search radius. The difference in root mean square error (RMSE) relative to CRU between the trial results and LME simulation was used as the metric. The rule was that, for each target grid, the distance corresponding to the minimum RMSE was set as the optimal search radius, and the RMSE for the merged results with trial radius had to be smaller than the RMSE of the original LME simulation. These trials were performed for the period 1921–1980, when the proxy records were scaled. The linear increment of radius candidates for attempts were set as follows: for the inter-annual component, 500–2000 km with 100-km interval; for the decadal component, 500–3000 km with 200-km interval; for multidecadal and centennial components, 500–4500 and 500–5000 km, both with 500-km interval.

By applying the diagnostic radius, independent validation was used to validate the inter-annual, decadal, and multidecadal components. The screened proxy components were first scaled for 1951–2000, and the temperature anomalies were merged with the background field, i.e., LME simulation for 1911–1950. Therefore, the instrument data from this period became independent of the data used to scale the proxy components, and the error statistic and RMSE were derived.

For the centennial component, the period for validation was not sufficiently long to validate it directly. Considering that (1) IMFs decomposed by EEMD are basically zero-mean, and the system bias should only exist in the residual term (i.e., "resi" term in Eq. 1 in Sect. 2.3) contained in the centennial component, the purpose of validation is to test the capability of the merging method to reduce uncertainty rather than the bias for monotonic time series; and (2) merger of the longest component was achieved with the full-period (1911–2000), calibrated candidates of both the proxy records and background field, the bias against the CRU component was removed. Thus, in this study, we mainly focus on the validations for inter-annual, decadal, and multidecadal components.

3 Results and discussion

3.1 Calculated optimal radius

Optimal radii for all the grids were derived from the preliminary experiments for four typical components as stated in the last section. The zonal mean curves of calculated radius (Fig. 3) indicated larger distances at both low and high latitudes than at mid-latitudes, especially for 30°–50° N. This pattern is very similar to the previous result of a regression-based method (Jones et al. 1998) that defined a correlation decay distance varying from 1000 to 3000 km. The optimal radius for the inter-annual and decadal components mainly fell within this range and longer-timescale components (multidecadal and centennial) were larger. Physically, the barotropic atmospheric state over lower latitude areas is more
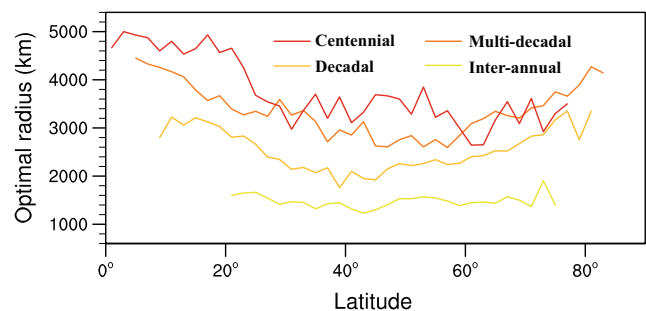
Fig. 3 Zonal mean optimal search radius (km) calculated for inter-annual, decadal, multidecadal, and centennial components over the lands of the Northern Hemisphere

likely to show homogeneous temperature patterns at longer timescales, during which the impact of ENSO oscillations at inter-annual timescale would be canceled out. For the higher latitude area, the ocean state and related phenomena (e.g., Atlantic Meridional Overturning Circulation and interactions between the ocean circulation and sea ice changes) tend to exhibit variabilities at longer timescales (e.g., multidecadal and centennial). Therefore, the surface temperature should also display a more homogeneous state. Additionally, the polar area became more sensitive to global warming, and thus, longer-term warming should be more common in higher latitude areas. Thus, the difference in the radius (high values over high latitudes and lower values over mid-latitudes) for the lower-frequency components (decadal, multidecadal, and centennial modes) would be more obvious than inter-annual modes.

The search radius in this study determines the selection of nearby proxy data directly, and the optimal values should represent a compromise between two aspects. The first is the complex topography. Because the underlying topography and ground properties directly impact local temperature covariance with places far away, the spatial correlation should be constrained within the local area (Hasler et al. 2011;

Lookingbill et al. 2003). The second aspect lies in the actual availability of distributed proxy records for the target grid, which requires that the scanning process be somewhat longer. The calculated optimal radius for different components can also be treated as compatible with the temporal and spatial variabilities of surface temperature.

3.2 Validation of merged results

As described in Sect. 2.4, the difference of RMSE relative to instrument data between LME and validated EMM results and correlation for the three typical components (inter-annual, decadal, multidecadal) over 1911–1950 is shown in Figs. 4 and 5. It is evident that the inter-annual component displays greater improvement than the other components.

The climate model could not reproduce the high-frequency variability of phase transition within strict calendar years, which is believed to be impacted by internal variability. By merging the proxy records, the inter-annual component can alleviate this shortcoming in EMM. The improvement of EMM results (increased correlation with independent instrument data and negative RMSE differences) was mainly across
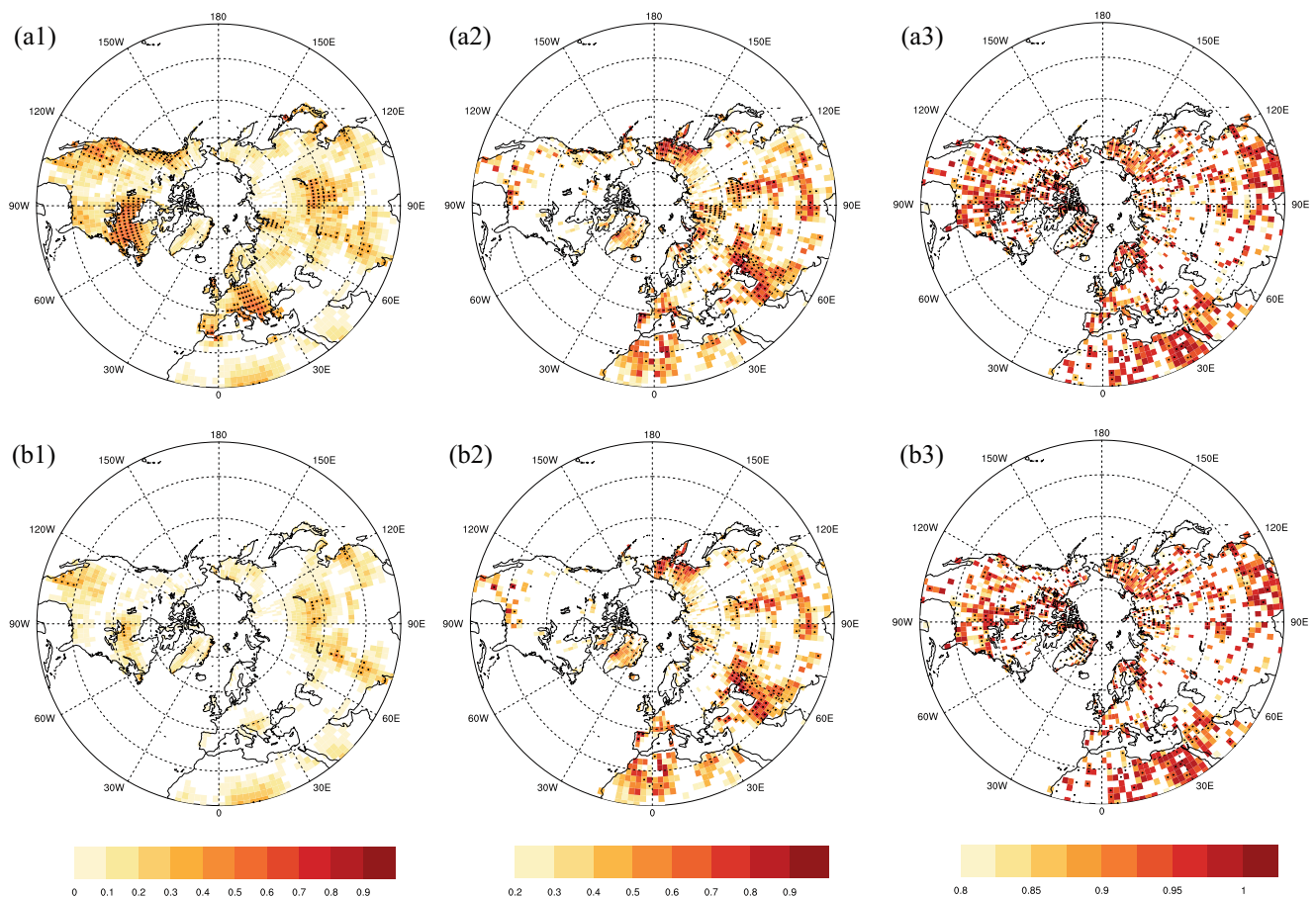
Fig. 4 *a1–a3* Correlation coefficients between the merged results and instrument CRU data for inter-annual (*left*), decadal (*middle*), and multidecadal (*right*) components. *b1–b3* Correlation between LME and CRU data at corresponding components (inter-annual, decadal, multidecadal) during 1911–1950 AD. The stippled grids indicate the 95% significant level given by the Monte Carlo simulation
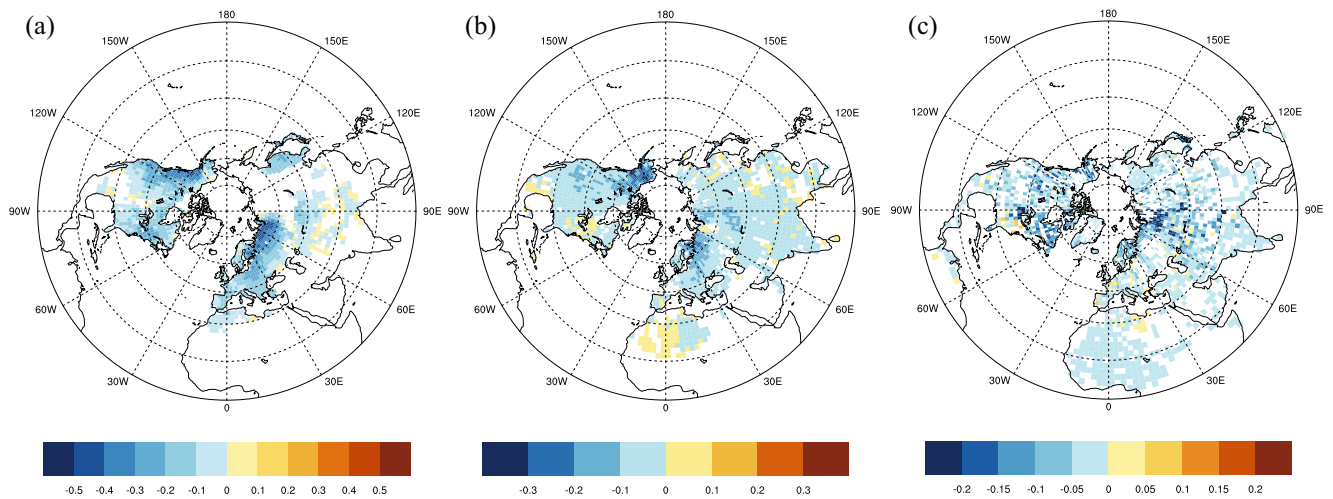
Fig. 5 Difference of root mean square error (RMSE) between components of EMM and LME simulation for **a** inter-annual, **b** decadal, and **c** multidecadal components during the validation period (1911–1950). Zero values have been masked out

the western part of Europe, the west coast and eastern portion of North America, and the polar region of Siberia (Figs. 4 (a1, b1) and 5a). The area with positive significant correlation coefficients at mid-latitudes increased by 15–20% (Fig. 6a). The standard deviation of the errors with respect to CRU at proxy sites over the entire Northern Hemisphere (NH) declined by ~0.15 °C. A significance test for correlation was conducted using Monte Carlo red noise simulation. There was less improvement for longer timescales (decadal and multidecadal), with increases of significant correlation area of 5–15% (decadal, Fig. 6b) and 5–10% (multidecadal, Fig. 6c). In addition, the zonal average RMSE reduction

(Fig. 6d, blue and green curves) for decadal and multidecadal periods became less significant as well. It is reasonable to infer that the merging procedure is more likely to highlight information at shorter timescales. First, fewer available proxy records could be merged at longer timescales. For example, for the decadal component, 110 records were used for merging, about half those used for the inter-annual timescale (209). Second, the inherent long-term climate variabilities were much weaker than those in the short term. Finally, although studies have extensively addressed the capability of proxy data recording of long-term climate change signals, especially regarding tree-ring chronologies (e.g., tree ring density and
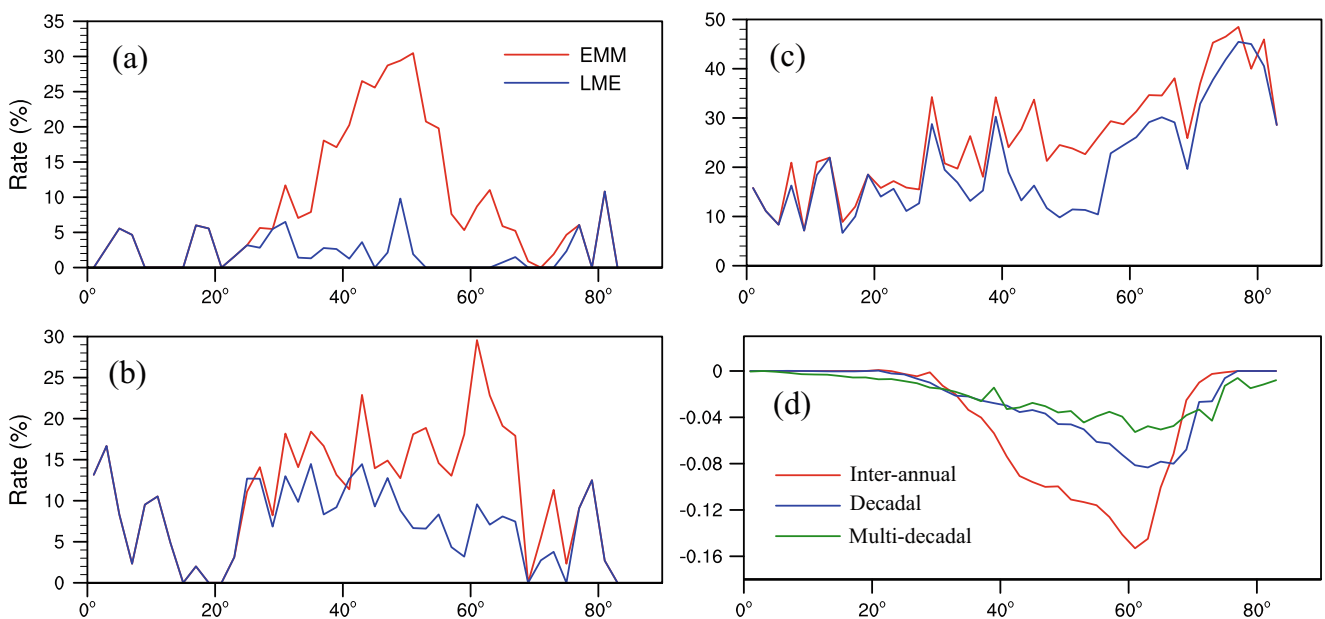
Fig. 6 Proportion of area along zonal belts with significant positive correlation of both the EMM components (*red*) and LME simulation (*blue*) relative to instrument data for **a** inter-annual, **b** decadal, and **c** multidecadal components. **d** Zonal averaged RMSE difference for inter-annual (*red*), decadal (*blue*), and multidecadal components (*green*)
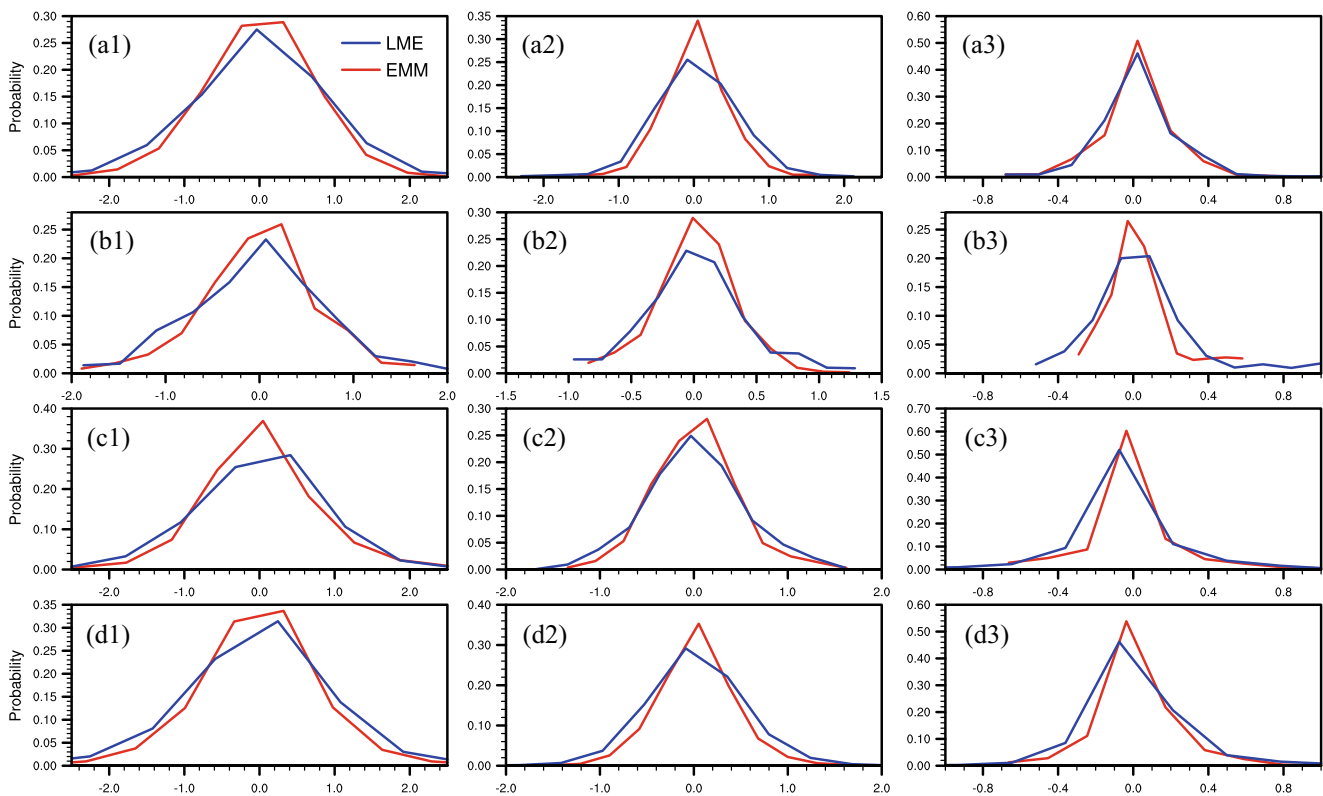
Fig. 7 *Row plots* present the statistical error distribution of both LME (blue curve) and EMM (red curve) on the proxy sites over North America (*a1–a3*), East Asia (*b1–b3*), Europe (*c1–c3*), and the Northern Hemisphere (*d1–d3*) for the components of inter-annual (*left column*), decadal (*middle column*), and multidecadal (*right column*)

width), there remain substantial uncertainties in analyzing relatively long-term variability (Christiansen 2011; Christiansen and Ljungqvist 2011; Esper et al. 2002). We extracted the validation results over three regional domains, Europe (35° to 70° N–10° to 40° E), North America (22.5° to 77.5° N–

Table 1 Regional (NA, EA, and EU) and Northern Hemispheric statistical standard deviation (°C) of errors for each component of LME and EMM validation results

Region	Results	IA	D	MD
North America	LME	0.88	0.55	0.21
	EMM	0.72	0.43	0.20
	Reduction	0.16	0.12	0.01
Eastern Asia	LME	0.73	0.41	0.27
	EMM	0.62	0.32	0.17
	Reduction	0.11	0.09	0.1
Europe	LME	0.87	0.52	0.33
	EMM	0.76	0.45	0.28
	Reduction	0.11	0.07	0.05
Northern Hemisphere	LME	0.94	0.51	0.27
	EMM	0.79	0.40	0.22
	Reduction	0.15	0.11	0.05

Reduction is referred to that of LME values minus EMM values in this table

42.5° to 67.5° W), and eastern Asia (23° to 54° N–60° to 160° E). It is clear that, for all three domains, the error range of EMM results (red lines in Fig. 7) became smaller than those of LME simulation (blue lines). Several factors led directly to error reduction: (1) the good quality of proxy records used in this study, which have been reported or proved to have correlations with temperature; (2) the climate response at different timescales was highlighted by the process of mode decomposition; and (3) the effectiveness of proxy components screening. As Table 1 has shown, for the results over North America, the standard deviation of errors for the inter-annual component (0.16 °C) is much larger than that of the multidecadal component (0.01 °C). Such a contrast became smaller in the

Table 2 Scheme for combining the merging results

Combination schemes	Inter-annual	Decadal	Multidecadal	Centennial
LME	LME_IA	LME_D	LME_MD	LME_C
Mixed-IA	EMM_IA	LME_D	LME_MD	LME_C
Mixed-IAD	EMM_IA	EMM_D	LME_MD	LME_C
Mixed-IADMD	EMM_IA	EMM_D	EMM_MD	LME_C
Full-EMM	EMM_IA	EMM_D	EMM_MD	EMM_C

LME_*: components of the ensemble mean at inter-annual (IA), decadal (D), and multidecadal (MD); EMM_*: the merged components

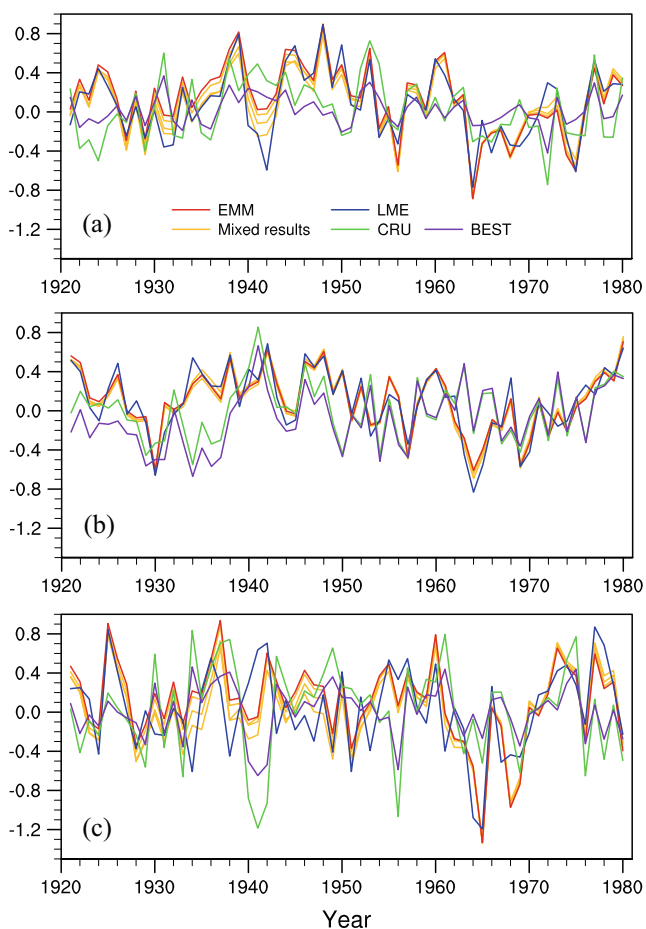**Table 3** Correlation coefficient between regionally averaged time series of merged results and CRUTS (1921–1980)

Zonal Belt	LME	Mixed-IA	Mixed-IAD	Mixed-IADMD	Full-EMM
Northern Hemisphere (1° N–87° N)	0.688	0.722	0.737	0.730	0.717
Extra-tropical area (31° N–87° N)	0.671	0.720	0.734	0.729	0.719
Arctic region (61° N–87° N)	0.459	0.525	0.541	0.542	0.540

results for Europe and eastern Asia. Two main causes might be inferred. One is that a larger proportion (near half) of the maximum latewood density (MXD) chronologies were used in North America, and these MXD chronologies tend to show a better response to the relatively high-frequency (inter-annual) temperature variability, especially for temperature variations over the growing season (April–September). However, climatic changes on long timescales recorded in density chronologies (not regional curve standardization chronologies) are likely to be underestimated, as discussed in previous studies (Briffa et al. 2004; Briffa et al. 1996). This chronology utility for North America might somewhat explain the greater improvements for the inter-annual component. The other causes concerning the availability of instrument data should also be

Fig. 8 Regional series of temperature anomalies (°C) relative to 1951–1980, the area-weighted series of **a** North America, **b** Eastern Asia, and **c** Europe. Mixed results are referred to the "Mixed-IA," "Mixed-IAD," and "Mixed-IADMD"

considered. One is that the number of observation stations in eastern Asia was smaller than in North America and Europe (Mitchell and Jones 2005), especially for years prior to 1951. This difference might have caused uncertainties in validation of the inter-annual component over eastern Asia. The other is the representativeness of the instrument dataset, as in Michells et al.'s (2005) study, where the regions with no station observations would be filled using stations within the correlation decaying distance; thus, the longer-term variabilities should be more realistic.

3.3 Surface temperature improvement in merged results

Merged surface temperature variability over the past centuries for each timescale was obtained using the full-period (1901–2000) components calibrated by instrument data. Since the EMM could be treated as a mixture of simulation and proxy records, we compared the datasets combining different components from the ensemble mean of LME and EMM, so as to further evaluate the effect of the assimilation. Thus, four sets of merged results were generated. As Table 2 shows, there were mixed results consisting of the same four typical components as the ensemble mean of LME with components replaced by those of EMM only at certain scales. By calculating the correlation coefficients of the combined annual resolved series with CRU for the entire land area of the NH (1° N–87° N), NH extratropics (31° N–87° N), and the polar areas (61° N–87° N) (Table 3), the improvements showed some spatial preference in extra-tropical and Arctic areas. Correlations with CRU for both the merged results (this term refers to the mixed datasets in Table 2) and LME across the Arctic were lower than those for the entire NH or extra-tropical NH. This is reasonable because some modeling studies have shown that it remains a major challenge to simulate climate change in the Arctic region using coupled climate models (Holland and Bitz 2003; Vavrus et al. 2012), especially for calculating poleward heat transfer by ocean and atmosphere and related sea ice change in a long-duration simulation (Jungclaus et al. 2014; Zhang 2015). This directly affects the surface energy balance in the Arctic and generates further uncertainty in simulated surface temperatures (Pithan and Mauritsen 2014). The validation results in Figs. 4 and 5 reveal that by merging the proxy records in the Arctic region, the merged surface temperature clearly shows RMSE reduction and enhanced correlation (e.g., north of Europe, Siberia, and Alaska). It should be stressed that the correlation of the average temperature time

Table 4 Correlation coefficients between series of results of EMM and LME relative to the time series of instrumental data (1921–1980)

| Region	Instrumental datasets	LME	Mixed-IA	Mixed-IAD	Mixed-IADMD	Full-EMM
North America	CRUTs	0.151	0.249	0.297[a]	0.328[a]	0.337[a]
	BEST	0.089	0.210	0.274[a]	0.292[a]	0.299[a]
Eastern Asia	CRUTs	0.282[a]	0.366[a]	0.399[a]	0.402[a]	0.392[a]
	BEST	0.178	0.248	0.277[a]	0.291[a]	0.266[a]
Europe	CRUTs	−0.082	0.274[a]	0.372[a]	0.378[a]	0.375[a]
	BEST	−0.090	0.256[a]	0.354[a]	0.358[a]	0.355[a]

[a] Ninety-five percent confidence level

series of merged data increased greatly with respect to the LME (Table 3), although only the inter-annual component was replaced in the original LME in the combination scheme Mixed-IA. It is reasonable that the combination results of Mixed-IADMD always displayed a higher correlation than that of Mixed-IA (Table 3) over NH, extra-tropical and the polar areas. In the validation results, the inter-annual components displayed a more significant improvement over mid-latitudes (Fig. 6), while the other longer-term components (decadal and multidecadal) show a higher proportion of

improved area over higher latitudes (north of 60° N) that would be merged into the Mixed-IADMD and Mixed-IAD. In addition, the correlation of Mixed-IADMD was slightly higher than that of Full-EMM, which should be also related to the processing of centennial components as discussed in Sect. 2.4 (Fig. 8).

To evaluate the merged combination results more objectively, another set of instrument data, the BEST, was adopted for comparison. Although there remains some uncertainty between different instrument datasets (Harris et al. 2014; Rohde

Fig. 9 Comparison of the same-sign rate (SSR) of EMM_IA (**a**, **b**) and LME_IA (**c**, **d**) inter-annual component over the land 20° N north with BEST (*left column*) and CRU (*right column*)
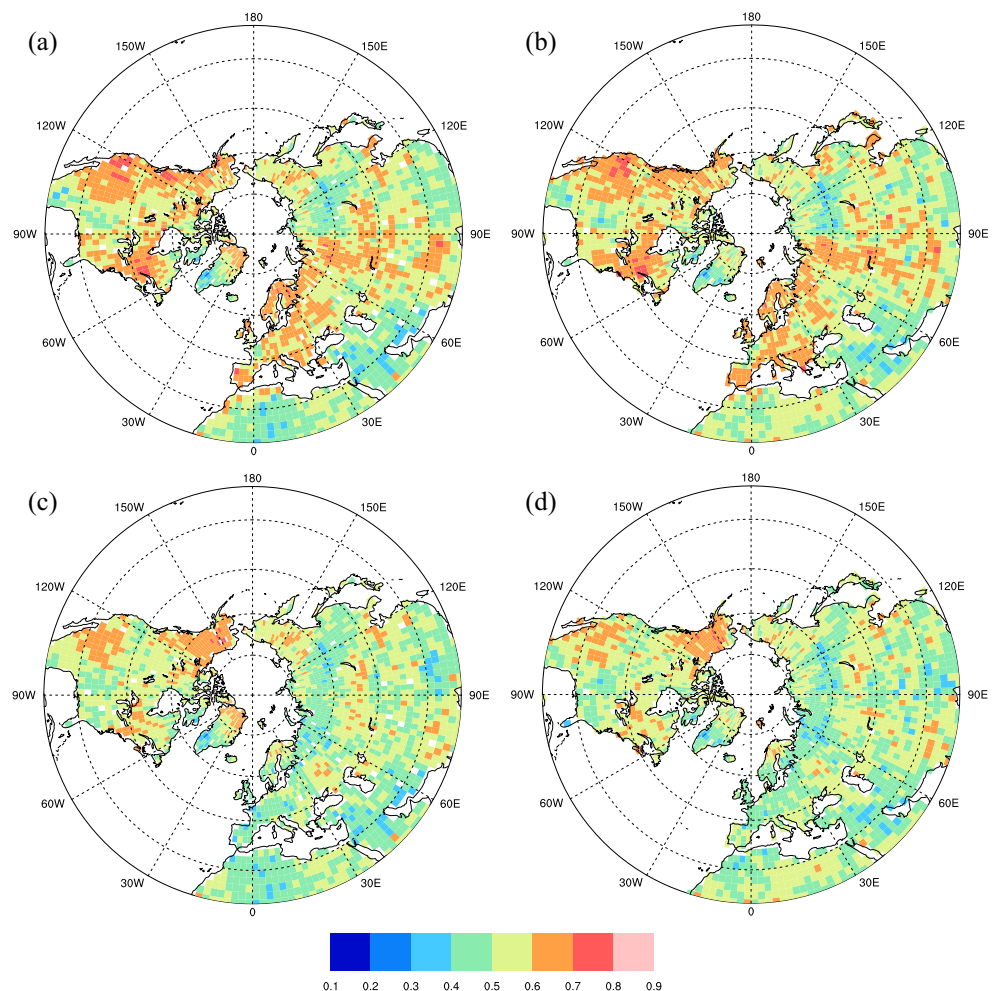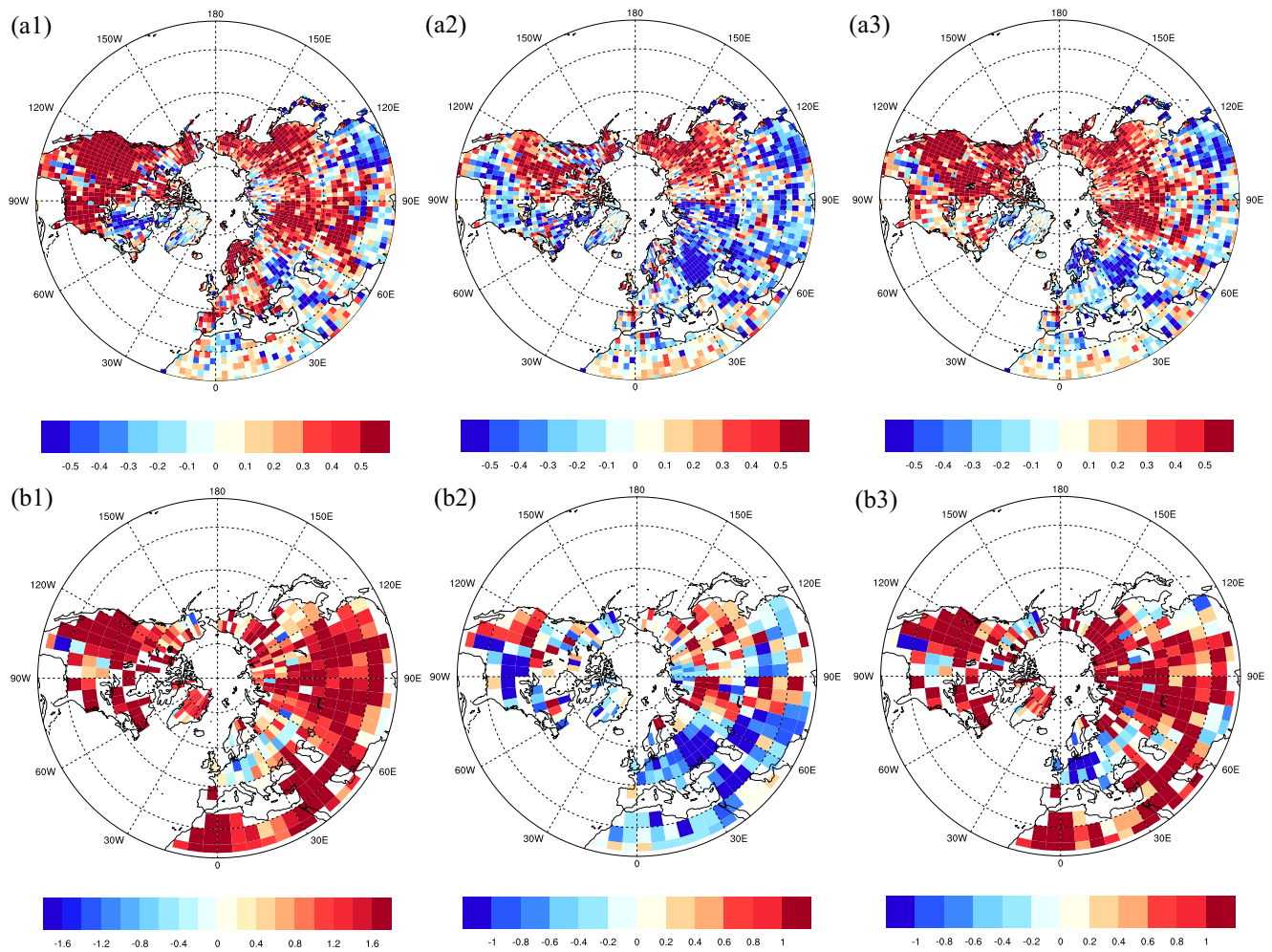
Fig. 10 The spatial pattern of surface temperature anomaly response differences to three typical volcanic eruptions 1815 (*left*), 1783 (*middle*), and 1641 (*right*). *a1–a3* Between EMM results and LME simulations. *b1–b3* Between Mann09 reconstruction and LME simulation. The reference period is 1951–1980

2013), it is important to evaluate the merged results with an independent benchmark. In Table 4, the mixed results and Full-EMM, with different combination schemes, also exhibit higher correlations with BEST and the increment of correlations are significant as well. Because the climate model-simulated results tend to be influenced by the phase of internal variability, the improvements from merging proxy records should be measured by how often the phase transition was actually corrected. Although we have quantified the error ranges, the incorporated series were error contaminated. We have to acknowledge that there remained errors in proxy-scaled temperature components, since the scaled VM could not get rid of these, which produce a compatible variance. It would be clearer to stress the correction by merging with the phase transition metric, not just the reduced values of error. In this sense, we examined the same-sign rate of the inter-annual variability (EMM_IA) by comparing the normalized series on each grid with instrument data (CRU and BEST) for the period 1921–1980 (Fig. 9). Two target series were normalized,

and the same-sign rate was calculated by counting the same-sign pairs. It was more convenient to examine the same-sign-rate, with zero-mean inter-annual component; otherwise, the trend term (longer-term variability) might blur the phase transition. The same-sign rate spatial pattern is very similar to the correlation and RMSE patterns for the inter-annual component of validation results. The greatest improvement was over Europe, where the average LME-simulated temperature series was poorly correlated with the instrument measurements (Table 4). The improvement over western and southern Europe was greater, by 20–30% in Spain and areas north of the Mediterranean Sea.

3.4 Volcanic response

A great number of studies have addressed the capability of model simulation and proxy-based reconstruction in estimating the response to volcanic eruptions (Anchukaitis et al. 2012; D'Arrigo et al. 2013; Mann et al. 2012; Stine and

Huybers 2014; Tingley et al. 2014). The issue of volcanic response at short timescales remains open for both climate model simulation and proxy-based reconstruction (Driscoll et al. 2012; Solomon 2011). We compared the EMM results (combination of all merged components) and spatiotemporal reconstruction of Mann et al. (2009) (Mann09 hereafter) for three typical post-sixteenth century volcanic eruptions, i.e., 1815 (Tambora), 1783 (Laki), and 1641 (Parker). This was because prior to 1600, only a third of proxy records are available. According to the volcanic reconstruction of Gao et al. (2008), Tambora was the strongest among the three eruptions. Three-year average surface temperature anomalies (the eruption year and subsequent 2 years) were calculated for the LME simulation, EMM results, and Mann09 reconstruction. Differences between EMM and Mann09 relative to LME are shown in Fig. 10. Spatial patterns of the differences of EMM were similar to those of Mann09 for the eruptions in 1783 (Fig. 10, middle) and Parker (Fig. 10, right). Moreover, the magnitude of EMM became smaller than that of Mann09. This means that the innovation was not as great as the discrepancy between pure simulation and reconstruction (Mann09), especially for Tambora and Parker, both of which were in tropical regions. The average NH surface temperature (north of 20° N) response of Mann09, EMM, and LME to the Tambora eruption cooling was −0.73, −1.78, and −2.01 °C, respectively. For the Parker eruption, these were −0.51, −1.35, and −1.53 °C. It can be inferred that the EMM results can reduce the discrepancy of the response between the pure simulation and empirical method-based reconstruction. Estimation of volcanic response by the EMM provides new insights by combining proxy reconstruction and model simulation in a statistical way. This is because the merging is achieved for different scales, and so the innovation is reflected as a combination of those scales. As shown in Fig. 10 (a1–a3), the innovation for all land grids contain combinations of all the components. Inter-annual and decadal scale innovation tended to be in certain local or regional areas and other areas retained LME-simulated values. In addition, to avoid the use of false information, the local OI only used proxy-scaled observation points within the range of optimal radius. Therefore, in areas with favorable improvement of short-term variability, we strongly believe that a more objective response has been achieved. Nevertheless, it remained difficult to provide a fully optimized picture of the volcanic response.

4 Conclusions and prospects

This work presents a new merging method that combines climate model (CESM)-simulated surface temperature and proxy records, using data assimilation techniques from a multiscale perspective. Because of the strengths and weaknesses of the two components of the combination, well-dated tree-ring and other types of proxy records could be used to correct inter-annual simulation, which might be obscured by internal variabilities. Moreover, the climate model-simulated temperature field at large spatiotemporal scales revealed good performance for lower-frequency (decadal and multidecadal) variation. Therefore, the scale-separation method provides a new approach to the aforementioned combination, which is optimal for reducing the weaknesses of the two components alone. The local OI method was applied to merging typical components with optimal radius, which constrains climate change signals to appropriate spatiotemporal scales for the background field. The independent validation results indicate that EMM had a smaller error range than LME for the inter-annual, decadal, and multidecadal components. RMSE and correlation in various latitudinal belts and regions both demonstrated that merging improvement for the inter-annual component was more substantial than the other components, especially for vast areas of Europe and the western and eastern coastal areas of North America. The decadal component for the area of Siberia near 60° N and several scattered grids over North America was improved also significantly. By calculating the SSR for the merged inter-annual component especially, it was inferred that the phase transition of surface temperature at high frequency was able to be corrected, and the spatial pattern was very similar to that of the validation results.

The response to volcanic forcing of the merged results in three typical eruption years (1815, 1783, and 1641) was inspected. Although the spatial difference of merged results and LME was similar to that of proxy-based reconstruction, the magnitude of response was moderate, which indicates that the discrepancy between the climate model simulation and pure reconstruction was somewhat reduced. The issue of how to estimate the response to volcanic forcing either by improving climate model simulation or further exploring and inspecting the proxy records remains open.

Finally, for the uncertainties in the merged results, three key aspects might affect these. The first is the representativeness of the proxy records. In this study, those records with significant correlation (tested by the Monte Carlo method in Appendix 1) with annual mean surface temperature were used. However, the records might indicate the temperature changes in some specific months (e.g., growing season) rather than the annual mean values in reality. We can anticipate that such a kind of uncertainty would get weaker in longer-timescale components (multidecadal and centennial scale). The second is the quality of the background field, which might be impacted by the factors of simulated internal variabilities, climate forcings, and the model structures. Third, as a data assimilation methodology (the OI algorithm), the background covariance matrix also played a critical role to determine how much the innovation (the difference between the reconstruction and background field) would be spread to different grids.

This proposed EMM considers the compatibility of climate change signals at different timescales and treats them individually. This provides a platform to combine the proxy records with diverse temporal resolutions and climate model-simulated outcomes under different physical frameworks. In the near future, there should be additional effort in the following three areas: (1) before the merging process, the VM scaling method was performed (although it has the potential to objectively preserve the variance of instrument datasets, the transfer functions were not constrained to give the best fit for the CRU in the calibration period); (2) the scale separation-based merging highlighted the advantage of short-term variability, but the issue of how to validate the centennial variability in real cases remains open; and (3) different types of proxy records have their own unique advantages in recording climate change over long periods, which were not assessed specifically in this work. In the long run, it is critical to explore the compatibility of long-term variability, using different types of proxy records and climate model-simulated results over the last millennium. Additionally, discerning the type of uncertainties in proxy scaling and simulation results would be helpful in assimilating diverse proxy records in a more meaningful way.

As climate models continue to develop and proxy records become more numerous over land areas in the near future, more effective combinations of the model simulation and proxy reconstruction will become increasingly important to understanding actual climate changes in the past.

Appendix 1

A Monte Carlo simulation test process was applied to quantify the significant levels of correlation coefficient for the time series with reduced freedom (Mann et al. 2007; Mann et al. 2009; Xing et al. 2016). The Monte Carlo simulated series was generated with the same auto-correlations as the original series by a first-order auto-regression (AR-1) model plus red noise. There were 3000 simulated series produced for target

component series, i.e., the four typical components of instrument series (CRU series). This Monte Carlo test was applied to (1) screen the proxy records candidates and (2) test the significance of correlation between the segments (the components at different timescale) of merged results and instrument series.

Appendix 2

The localized OI algorithm used in this paper at four typical timescales used a dynamically updated B matrix within the same frame as follows in Eq. 2.

$$T_i^a = T_i^b + \sum_k^M G_{ik} \left[T_k^o - H(T^b) \right] \tag{2}$$

where T_i^a is the analyzed (merged) component at a certain timescale. T_i^a is the ensemble mean of background. M is the number of available proxy records within the range with diagnostic radius. H is the observation operator, in this study functions to transfer the nearest value on regular grid to the proxy sites. G is the gain matrix value for the ith grid and kth proxy component. The G matrix is the solution of the set of linear equations (3):

$$BH^T = G(R + HBH^T) \tag{3}$$

where the R matrix is "observation error covariance matrix." In this study, we use the diagonal form of R, which is estimated by the covariance of scaled proxy components at certain timescales. B is the background error covariance matrix. B was not calculated directly and it was stored as BH^T (or HB) implicitly.

When deriving G, as the H operator transferred the nearest grid value to the scattered proxy site, it is likely that BH^T might become ill-conditioned; in that case, the linear equations would be solved by the technique of singular value decomposition (SVD), and the truncated results would be given.

References

Acevedo W, Reich S, Cubasch U (2016) Towards the assimilation of tree-ring-width records using ensemble Kalman filtering techniques. Clim Dyn 46:1909–1920

Ahmed M, Anchukaitis KJ, Asrat A et al (2013) Continental-scale temperature variability during the past two millennia. Nat Geosci 6:339–346

Anchukaitis KJ, Breitenmoser P, Briffa KR et al (2012) Tree rings and volcanic cooling. Nat Geosci 5:836–837

Barkmeijer J, Iversen T, Palmer T (2003) Forcing singular vectors and other sensitive model structures. Q J Roy Meteor Soc 129:2401–2423

Briffa KR, Jones PD, Schweingruber FH et al (1996) Tree-ring variables as proxy-climate indicators: problems with low-frequency signals.

Climatic Variation and Forcing Mechanisms of the Last 2000 Years. Springer, 9–41

Briffa KR, Osborn T, Schweingruber FH (2004) Large-scale temperature inferences from tree rings: a review. Glob Planet Chang 40:11–26

Chen X, Xing P, Luo Y et al (2015) Surface temperature dataset for North America obtained by application of optimal interpolation algorithm merging tree-ring chronologies and climate model output. Theor Appl Climatol. doi:10.1007/s00704-015-1634-4

Christiansen B (2011) Reconstructing the NH mean temperature: can underestimation of trends and variability be avoided? J Clim 24:674–692

Christiansen B, Ljungqvist FC (2011) Reconstruction of the extratropical NH mean temperature over the last millennium with a method that preserves low-frequency variability. J Clim 24:6013–6034

Christiansen B, Schmith T, Thejll P (2009) A surrogate ensemble study of climate reconstruction methods: stochasticity and robustness. J Clim 22:951–976

Cook ER, Meko DM, Stahle DW et al (1999) Drought reconstructions for the continental United States. J Clim 12:1145–1162

Cook ER et al (2010) Asian monsoon failure and megadrought during the last millennium. Science 328:486–489

Cook ER et al (2013) Tree-ring reconstructed summer temperature anomalies for temperate East Asia since 800 CE. Clim Dyn 41:2957–2972

D'Arrigo R, Wilson R, Anchukaitis KJ (2013) Volcanic cooling signal in tree ring temperature records for the past millennium. J Geophys Res 118:9000–9010

Deser C, Phillips AS, Bourdette V, Teng H (2012) Uncertainty in climate change projections: the role of internal variability. Clim Dyn 38:527–546

Deser C, Phillips AS, Alexander MA, Smoliak BV (2014) Projecting North American Climate over the next 50 years: uncertainty due to internal variability. J Clim 27:2271–2296

Driscoll S, Bozzo A, Gray LJ, Robock A, Stenchikov G (2012) Coupled Model Intercomparison Project 5 (CMIP5) simulations of climate following volcanic eruptions. J Geophys Res 117:127–135

Esper J, Cook ER, Schweingruber FH (2002) Low-frequency signals in long tree-ring chronologies for reconstructing past temperature variability. Science 295:2250–2253

Evensen G (2003) The ensemble Kalman filter: theoretical formulation and practical implementation. Ocean Dyn 53:343–367

Fernández-Donado L, González-Rouco JF, Raible CC et al (2013) Large-scale temperature response to external forcing in simulations and reconstructions of the last millennium. Clim Past 9:393–421

Franzke C (2012) Nonlinear trends, long-range dependence, and climate noise properties of surface temperature. J Clim 25:4172–4183

Franzke C (2014) Warming trends: nonlinear climate change. Nat Climate Change 4:423–424

Gandin LS, Hardin R (1965) Objective analysis of meteorological fields. Vol. 242, Israel program for scientific translations, Jerusalem

Gao C, Robock A, Ammann CM (2008) Volcanic forcing of climate over the past 1500 years: an improved ice core-based index for climate models. J Geophys Res 113:D23111

Goosse H (2016) Reconstructed and simulated temperature asymmetry between continents in both hemispheres over the last centuries. Clim Dyn. doi:10.1007/s00382-016-3154-z

Goosse H, Renssen H, Timmermann A, Bradley RS (2005) Internal and forced climate variability during the last millennium: a model-data comparison using ensemble simulations. Quaternary Sci Rev 24:1345–1360

Goosse H, Renssen H, Timmermann A, Bradley RS, Mann ME (2006) Using paleoclimate proxy-data to select optimal realisations in an ensemble of simulations of the climate of the past millennium. Clim Dyn 27:165–184

Goosse H, Guiot J, Mann ME, Dubinkina S, Sallaz-Damaz Y (2012) The medieval climate anomaly in Europe: comparison of the summer

and annual mean signals in two reconstructions and in simulations with data assimilation. Glob Planet Chang 84:35–47

Harris I, Jones PD, Osborn T, Lister D (2014) Updated high-resolution grids of monthly climatic observations–the CRUTS3.10 Dataset. Int J Climatol 34:623–642

Hasler A, Gruber S, Haeberli W (2011) Temperature variability and offset in steep alpine rock and ice faces. Cryosphere 5:977–988

Holland MM, Bitz CM (2003) Polar amplification of climate change in coupled models. Clim Dyn 21:221–232

Huang NE, Wu Z (2008) A review on Hilbert-Huang transform: method and its applications to geophysical studies. Rev Geophys 46:RG2006

Huber M, Knutti R (2014) Natural variability, radiative forcing and climate response in the recent hiatus reconciled. Nat Geosci 4:651–656

Ji F, Wu Z, Huang J, Chassignet EP (2014) Evolution of land surface air temperature trend. Nat Climate Change 4:462–466

Jones PD, Mann ME (2004) Climate over past millennia. Rev Geophys 42:99–103

Jones PD, Briffa KR, Barnett TP et al (1998) High-resolution palaeoclimatic records for the last millennium: interpretation, integration and comparison with general circulation model control-run temperatures. The Holocene 8:455–471

Jones PD, Briffa KR, Osborn TJ et al (2009) High-resolution palaeoclimatology of the last millennium: a review of current status and future prospects. The Holocene 19:3–49

Juckes MN, Allen MR, Briffa KR et al (2007) Millennial temperature reconstruction intercomparison and evaluation. Clim Past 3:591–609

Jungclaus JH, Lohmann K, Zanchettin D (2014) Enhanced 20th century heat transfer to the Arctic simulated in context of climate variations over last millennium. Clim Past 10:2201–2213

Landrum L, Otto-Bliesner BL, Wahl ER et al (2013) Last millennium climate and its variability in CCSM4. J Clim 26:1085–1111

Lee TC, Zwiers FW, Tsao M (2008) Evaluation of proxy-based millennial reconstruction methods. Clim Dyn 31:263–281

Lookingbill TR, Urban DL (2003) Spatial estimation of air temperature differences for landscape-scale studies in montane environments. Agric For Meteorol 114:141–151

Mann ME, Bradley RS, Hughes MK (1998) Global-scale temperature patterns and climate forcing over the past six centuries. Nature 392:779–787

Mann ME et al (2005) Testing the fidelity of methods used in proxy-based reconstructions of past climate. J Clim 18:4097–4107

Mann ME et al (2007) Robustness of proxy-based climate field reconstruction methods. J Geophys Res 112:1103–1118

Mann ME et al (2008) Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia. Proc Natl Acad Sci 105:13252–13257

Mann ME et al (2009) Global signatures and dynamical origins of the Little Ice Age and Medieval Climate Anomaly. Science 326:1256–1260

Mann ME, Fuentes JD, Rutherford S (2012) Underestimation of volcanic cooling in tree-ring-based reconstructions of hemispheric temperatures. Nat Geosci 5:202–205

Matsikaris A, Widmann M, Jungclaus JH (2015) Assimilating continental mean temperatures to reconstruct the climate of the late pre-industrial period. Clim Dyn 46:3547–3566

Mitchell TD, Jones PD (2005) An improved method of constructing a database of monthly climate observations and associated high-resolution grids. Int J Climatol 25:693–712

Otto-Bliesner BL, Brady EC, Fasullo J et al (2015) Climate variability and change since 850 CE: an ensemble approach with the Community Earth System Model (CESM). Bull Amer Meteor Soc. doi:10.1175/BAMS-D-14-00233.1

Pithan F, Mauritsen T (2014) Arctic amplification dominated by temperature feedbacks in contemporary climate models. Nat Geosci 7:181–184

Rohde R et al (2013) A new estimate of the average Earth surface land temperature spanning 1753 to 2011. Geoinfor Geostat An Overview 1(1):1–7

Schneider T (2001) Analysis of incomplete climate data: estimation of mean values and covariance matrices and imputation of missing values. J Clim 14:853–871

Shi F (2015) A multi-proxy reconstruction of spatial and temporal variations in Asian summer temperatures over the last millennium. Clim Chang 131:1–14

Shi F, Yang B, Gunten L, Qin C, Wang Z (2011) Ensemble empirical mode decomposition for tree-ring climate reconstructions. Theor Appl Climatol 109:233–243

Shi F, Yang B, Gunten L (2012) Preliminary multiproxy surface air temperature field reconstruction for China over the past millennium. Sci China: Earth Sci 55:2058–2067

Solomon A (2011) Distinguishing the roles of natural and anthropogenically forced decadal climate variability: implications for prediction. Bull Amer Meteor Soc 92:141–156

Steiger NJ, Hakim GJ (2016) Multi-time scale data assimilation for atmosphere-ocean state estimates. Clim Past 12:1375–1388

Steiger NJ, Hakim GJ, Steig EJ et al (2014) Assimilation of time-averaged pseudoproxies for climate reconstruction. J Clim 27:426–441

Stine AR, Huybers P (2014) Arctic tree rings as recorders of variations in light availability. Nat Commun. doi:10.1038/ncomms4836

von Storch H et al (2000) Combining paleoclimatic evidence and GCMs by means of Data Assimilation Through Upscaling and Nudging. pp28–31, Proc. 11th Symposium on Global Climate Change Studies, AMS Long Beach, CA, 2000

Tingley MP, Huybers P (2010a) A bayesian algorithm for reconstructing climate anomalies in space and time. Part I: development and applications to paleoclimate reconstruction problems. J Clim 23:2759–2781

Tingley MP, Huybers P (2010b) A Bayesian algorithm for reconstructing climate anomalies in space and time. Part II: comparison with the regularized expectation-maximization algorithm. J Clim 23:2782–2800

Tingley MP, Stine AR, Huybers P (2014) Temperature reconstructions from tree-ring densities overestimate volcanic cooling. J Geophys Res 41:7838–7845

Vavrus SJ (2012) Twenty-first-century Arctic climate change in CCSM4. J Clim 25:2696–2710

Wang G, Chen XY, Qiao FL, Wu Z, Huang NE (2010) On intrinsic mode function. Adv Adapt Data Anal 2:277–293

Widmann M, Goosse H, van der Schrier G, Schnur R, Barkmeijer J (2010) Using data assimilation to study extratropical Northern Hemisphere climate over the last millennium. Clim Past 6:627–644

Wu Z, Huang NE (2009) Ensemble empirical mode decomposition: a noise-assisted data analysis method. Adv Adapt Data Anal 1:1–41

Xing P, Chen X, Luo Y et al (2016) The extratropical Northern Hemisphere temperature reconstruction during the last millennium based on a novel method. PLoS One. doi:10.1371/journal.pone.0146776

Zhang R (2015) Mechanisms for low-frequency variability of summer Arctic sea ice extent. Proc Natl Acad Sci 112:4570–4575

Zhang Z, Mann ME, Cook ER (2004) Alternative methods of proxy-based climate field reconstruction: application to summer drought over the conterminous United States back to AD 1700 from tree-ring data. The Holocene 14:502–516

Zheng J, Liu Y, Hao Z (2015) Annual temperature reconstruction by signal decomposition and synthesis from multi-proxies in Xinjiang, China, from 1850 to 2001. PLoS One. doi:10.1371/journal.pone.0144210