CrossMark

Assimilating continental mean temperatures to reconstruct the climate of the late pre-industrial period

Anastasios Matsikaris[1,2] · Martin Widmann[1] · Johann Jungclaus[2]

Abstract An on-line, ensemble-based data assimilation (DA) method is performed to reconstruct the climate for 1750–1850 AD, and the performance is evaluated on large and small spatial scales. We use a low-resolution version of the Max Planck Institute for Meteorology MPI-ESM model and assimilate the PAGES 2K continental mean temperature reconstructions for the Northern Hemisphere (NH). The ensembles are generated sequentially for sub-periods based on the analysis of previous sub-periods. The assimilation has good skill for large-scale temperatures, but there is no agreement between the DA analysis and proxy-based reconstructions for small-scale temperature patterns within Europe or with reconstructions for the North Atlantic Oscillation (NAO) index. To explain the lack of added value in small spatial scales, a maximum covariance analysis (MCA) of links between NH temperature and sea level pressure is performed based on a control simulation with MPI-ESM. For annual values, winter and spring the Northern Annular Mode (NAM) is the pattern that is most closely linked to the NH continental temperatures, while for summer and autumn it is a wave-like pattern. This link is reproduced in the DA for winter, spring and annual means, providing potential for constraining the NAM/ NAO phase and in turn regional temperature variability. It is shown that the lack of actual small-scale skill is likely due to the fact that the link might be too weak, as the NH continental mean temperatures are not the best predictors for large-scale circulation anomalies, or that the PAGES 2K

temperatures include noise. Both factors can lead to circulation anomalies in the DA analysis that are substantially different from reality, leading to unrealistic representation of small-scale temperature variability. Moreover, we show that even if the true amplitudes of the leading MCA circulation patterns were known, there is still a large amount of unexplained local temperature variance. Based on these results, we argue that assimilating temperature reconstructions with a higher spatial resolution might improve the DA performance.

Keywords Climate of the past · Climate modelling · Data assimilation · PAGES 2K

1 Introduction

Climate reconstructions for the pre-industrial period are either based on climate proxy data or on numerical simulations. However, both approaches are associated with substantial uncertainties. In principle, the best state estimates can be expected by employing data assimilation (DA) techniques, which systematically combine the empirical information from proxy data with the representation of the processes that govern the climate system given by climate models. The aim of DA is to develop a reconstruction that captures both the forced and internal climate variability. Here, we simulate the climate for the period 1750–1850 AD by employing an ensemble member selection DA method. We analyse temperatures on the continental scale that has been used in the assimilation, as well as hemispheric and global scales. Moreover, we investigate the performance of the DA on smaller spatial-scale temperature variability within Europe and modes of atmospheric variability, to evaluate the added value on patterns and modes that are not assimilated.

✉ Anastasios Matsikaris
  axm368@bham.ac.uk

[1] University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

[2] Max-Planck-Institute for Meteorology, Hamburg, Germany

_Springer

Even though DA is a mature field in numerical weather prediction, the specific problem in palaeoclimatology is different and the methods cannot be directly transferred (e.g. Goosse et al. 2006; Widmann et al. 2010; Hakim et al. 2013). Different DA approaches have been employed in the recent past. In most cases an ensemble-selection approach has been employed, as pioneered by Goosse et al. (2006). Other approaches include pattern nudging (Widmann et al. 2010) and forcing singular vectors (van der Schrier and Barkmeijer 2005). Some ensemble member selection techniques select a single simulation from an ensemble that is closest to the empirical information of the climate (Goosse et al. 2006; Crespin et al. 2009; Goosse et al. 2010; Matsikaris et al. 2015). In an alternative setup (Goosse et al. 2012; Annan and Hargreaves 2012; Mairesse et al. 2013; Klein et al. 2014), several highly likely members are copied proportionally to their likelihood. Finally, in other ensemble-based DA schemes, Kalman filters or variants of them are used (Dirren and Hakim 2005; Huntley and Hakim 2010; Pendergrass et al. 2012; Bhend et al. 2012; Steiger et al. 2014).

Our simulation period, 1750–1850 AD, is towards the end of the little ice age (LIA, 1500–1850 AD), the spatial and temporal extent of which is often debated (e.g. Jones and Mann 2004; Jansen et al. 2007). This choice of the analysis period has several advantages. A relatively high number of proxies have contributed to the PAGES 2K continental reconstructions used in the assimilation, thus reducing the proxy-induced errors. Additionally, the period is directly before the instrumental period and a relatively high amount of climate data, e.g. early instrumental records and historical documents, are available. Moreover, a major feature of the chosen period is the occurrence of strong tropical volcanic eruptions, which are a major natural driver of the interannual climate variability (Robock 2000; Cole-Dai 2010). The eruptions, such as the tropical eruption VEI-6 (unknown location, 1809 AD), the Tambora (Indonesia, 1815 AD) and Cosiguina (Nicaragua, 1835 AD) eruptions, have various dynamical effects on the ocean and the atmosphere (Zanchettin et al. 2012, 2013). Even though our DA scheme is applied with a time step of 10 years and the peak response to volcanic forcing is on the annual timescale, some forcing signal on the decadal timescale can be expected.

Here, we apply an "on-line" DA technique, which means that the ensembles are generated sequentially for sub-periods based on the best members of previous subperiods, giving the advantage of the temporal consistency of the simulated states. We use the Max Planck Institute for Meteorology Earth System Model (MPI-ESM) and assimilate decadal surface temperature means of the Northern Hemisphere (NH) continents from the PAGES 2K project (PAGES 2K Consortium 2013), as in Matsikaris et al.

(2015). Even though no information about the local temperatures is assimilated, we explore the performance of the DA on small spatial scales, since the assimilation of the NH continental averages might determine to some extent the state of the main modes of circulation variability, such as the Northern Annular Mode (NAM) or the North Atlantic Oscillation (NAO). In principle, prescribing continental temperatures can be expected to constrain the phase of leading circulation modes if these modes have a temperature signal associated with the average continental temperatures. In turn, the temperature signal of these circulation modes can be expected to provide information on temperature variability on sub-continental scales.

The validation of the DA setup is performed against assimilated and independent data, while a maximum covariance analysis (MCA) of links between temperature and pressure in the NH aims to explore the potential of added value in the assimilation run. The focus of the study is on the skill of the reconstruction using DA and the added value that can be achieved compared to simulations without DA. The structure of the paper is as follows: in Sect. 2, we review the characteristics of the model, proxy and instrumental datasets, followed by the details of the experimental design. Section 3 validates the assimilation approach for large-scale temperatures by comparing the simulation with the assimilated and independent data, and other simulations. The circulation and temperature variability in the North Atlantic-European sector for the assimilation run is examined in Sect. 4. Finally, in Sect. 5, we summarize and draw the main conclusions.

2 Methodology

2.1 Model, proxies and instrumental data

The assimilation has been performed with the Max Planck Institute for Meteorology Earth System Model (MPI-ESM). The atmospheric component is the General Circulation Model (GCM) ECHAM6 (Stevens et al. 2013) and the ocean component is MPIOM (Marsland et al. 2003). ECHAM6 was run at T31 horizontal resolution ($3.75° \times 3.75°$), with 31 vertical levels, resolving the atmosphere up to 10 hPa, while MPIOM was run at a horizontal resolution of $3.0°$ (GR30) and 40 vertical levels. The ocean and the atmosphere were coupled daily without flux corrections with the OASIS3 coupler. MPI-ESM also includes the land surface model JSBACH (Raddatz et al. 2007), while no ocean biogeochemistry model was employed. The model is a coarse-resolution version of the model used for the Coupled Model Intercomparison Project Phase 5 (CMIP5) simulations, and hereafter is referred to as MPI-ESM-CR. The configuration used here follows

the configuration for palaeo-applications (MPIESM-P) described in Jungclaus et al. (2014).

The simulations follow the "past1000" protocol of the Paleoclimate Modelling Intercomparison Project (PMIP3) Phase 3 (Schmidt et al. 2011). The "past1000" simulation has been started after a 700-year long spin-up with constant 850 AD boundary conditions. The DA simulations consist of 20 ensemble members for each decade between 1750 and 1850 AD, and include all natural and anthropogenic forcing. In particular, prescribed external forcings are reconstructed variations of total solar irradiance (Vieira et al. 2011), volcanic aerosols (Crowley and Unterman 2012), concentrations of the most important greenhouse gases (Schmidt et al. 2011) and anthropogenic land-cover changes (Pongratz et al. 2008). Two different types of simulations using the MPI-ESM-CR are examined, namely simulations with and without DA. In addition to the forced simulations, a 1000-year-long control run was performed and used for a MCA of links between temperature and pressure in the NH due to internal variability.

The DA simulations are constrained to follow the "2K Network" of the IGBP Past Global Changes (PAGES) datasets, which provide proxy-based temperature reconstructions for seven continental-scale regions and have annual resolution, apart from North America, which is resolved in 10- and 30-year periods (PAGES 2K Consortium 2013). Only the reconstructions of the NH continents were assimilated. The datasets have been produced by nine regional working groups, who identified the best proxy climate records within their region and followed either the "composite plus scale" approach for the adjustment of the mean and variance of a predictor composite to an instrumental target (e.g. Moberg et al. 2005), or regression-based techniques for the predictors (e.g. Mann et al. 2009). 511 time series of individual proxies have been employed, and include ice cores, tree rings, pollen, speleothems, corals, lake and marine sediments as well as historical documents of changes in biological and physical processes. The proxy data have been used to reconstruct annual means for the Arctic (60–90N, 180W–180E), summer (JJA) means for Asia (23.5–55N, 60–160E), summer (JJA) means for Europe (35–70N, 10W–40E), and decadal means of annual values for North America (30–55N, 130–75W).

The skill of the the DA simulations is assessed against the European seasonal surface air temperature reconstruction from Luterbacher et al. (2004). This statistical reconstruction provides gridded (0.5° × 0.5° resolution) monthly (back to 1659 AD) and seasonal (from 1500 AD to 1658 AD) temperatures for European land areas (25W–40E, 35–70N). It is based on homogenized and quality-checked instrumental data, reconstructed sea-ice and temperature indices derived from documentary records and seasonally resolved proxy temperature reconstructions from

Greenland ice cores and tree rings. The DA simulations are also compared with the early instrumental record; specifically, the Berkeley Earth Surface Temperature (BEST) dataset (Rohde et al. 2012), which uses temperature observations from a large collection of weather stations in order to estimate the underlying global land temperatures. Temperatures are reported as anomalies relative to the 1951–1980 AD average, along with their uncertainties, which represent the 95 % confidence interval for statistical noise and spatial undersampling effects. The uncertainties are larger in the earlier reconstructions and account for the effects of random noise as well as random biases affecting station trends and random shifts in station baselines. The BEST framework is expected to be robust against most forms of bias.

2.2 DA experimental design

The method we employ for assimilating the PAGES 2K proxy-based reconstructions is similar to the ones followed in recent ensemble-based DA studies (e.g. Goosse et al. 2006; Crespin et al. 2009; Matsikaris et al. 2015) and is based on a degenerated particle filter. We set up our DA by taking the last day of the year 1749 AD from a transient forced simulation starting in 850 AD (the "past1000" simulation) as the initial conditions of the ensemble. We generate 20 ensemble members by introducing small perturbations in an atmospheric diffusion parameter for the first year of the simulations, 1750 AD. After 10 years of simulations, a root mean square error-based cost function is used to compare the simulated decadal mean temperatures of the NH continents with the PAGES 2K continental proxy-based reconstructions. The member that minimizes the cost function is selected as the best member for that sub-period and is used as the initial condition for the subsequent 10-year simulation. A new ensemble consisting of twenty members is performed for the second decade, using the same method as before. The procedure is repeated sequentially until the end of the simulation period, 1850 AD. We combine each decade's best member to form the "DA analysis". The high computational cost did not allow a larger number of ensemble members, however the ensemble is twice as large as that in the previous study of Matsikaris et al. (2015).

The cost function used in the assimilation process to select the best simulation of the ensemble for each decade is:

$$CF(t) = \sqrt{\sum_{i=1}^{4} \left(T_{mod}^i(t) - T_{prx}^i(t) \right)^2} \qquad (1)$$

where i denotes the NH continents (Arctic, Asia, Europe and North America), $T_{mod}^i(t)$ is the standardized modelled decadal mean of the temperatures in each continent and

$T_{prx}^i(t)$ is the standardized proxy-based reconstruction for the decadal mean of the temperatures in each continent. In order to remove biases in means and variances, we standardise the model output and the proxy-based reconstructions by subtracting the 850–1849 AD means from the 1750–1849 AD raw model output and PAGES 2K reconstructions respectively, and dividing by the respective standard deviations, based on the decadal averages for the 850–1849 AD period. More details for the specific choices of the cost function are given in Matsikaris et al. (2015).

The proxy-based reconstructions are affected by various types of errors, which in turn affect the cost function and the member selection directly. More specifically, the statistical methods followed influence the reconstructions, the seasonal representativity is different in different proxies, non-climatic factors influence the proxies, while the poor spatial coverage induces large uncertainties in hemispheric or continental means, such as the ones we assimilate (e.g. Jones and Mann 2004; Jansen et al. 2007). As a result of these errors, there is a high chance that the selected ensemble member is not in agreement with reality. This problem is often tackled by either weighing the cost function according to the errors of the proxy data sets, or by retaining several members that are close to the proxy-based reconstructions (e.g. Goosse et al. 2012; Annan and Hargreaves 2012). Our simple assimilation scheme does not take the proxy errors into account, because they are not directly comparable for the different continents due to the different methods followed by each of the PAGES 2K groups.

## 3 Validation for continental to global-scale temperatures

The DA simulations are validated against empirical evidence for large-scale temperature variability, namely continental, hemispheric and global scales, in three ways. Firstly, we validate the simulations against the proxy-based reconstructions used during the assimilation, i.e. the PAGES 2K data, to check the extent to which the two are consistent. Secondly, the assimilation results are tested against other proxy-based temperature reconstructions, including Luterbacher et al. (2004). The independence between those and the assimilated reconstructions is often not clear, as predictors used in different proxy-based reconstructions may be common. Nevertheless, proxy data from Luterbacher et al. (2004) have not been used by the PAGES 2K groups in their reconstructions. Thirdly, the DA analysis is compared with early instrumental records (BEST dataset). We also evaluate the accordance of the DA analysis with simulations that do not perform DA.

3.1 Comparison of simulations with the assimilated proxy data

We compare the simulated NH continental temperature series with the assimilated PAGES 2K proxy-based reconstructions. We initially compare the decadal means, which is the timescale used in the DA. Figure 1 shows the NH continents' decadal mean temperature anomalies for 1750–1850 AD w.r.t. the 850–1849 AD mean, for the DA analysis, the different ensemble members, the ensemble mean, the simulation without DA ("past1000") and the proxy-based reconstructions. The simulated data are for the same regions as described by the PAGES 2K reconstructions. The analysis follows the assimilated reconstructions well, which is a prerequisite for a skilful DA method and indicates a sufficiently large ensemble size. However, in decades with strong volcanic forcing, namely 1810–1820 AD and 1830–1840 AD, some agreement also stems from a common response to the forcings. Correlations between analysis and proxy-based reconstructions in most continents (apart from Asia) are higher than the respective correlations found in Matsikaris et al. (2015) (0.93 instead of 0.79 for the Arctic, 0.64 instead of 0.76 for Asia, 0.95 instead of 0.79 for Europe and 0.96 instead of 0.81 for North America). This may be due to the increase in ensemble size from 10 to 20 members, or indicate a more realistic or stronger forcing in the 1750–1850 AD period compared with the earlier period, 1600–1700 AD, analysed in Matsikaris et al. (2015). The correlations of the "past1000" simulation with the PAGES 2K reconstruction are much lower (0.70, 0.45, 0.57 and 0.43 respectively) than those of the DA analysis, showing that DA improves the skill on the continental scale.

The "past1000" simulation includes the forcing signal and random internal variability. To focus on the forced variability we also investigate the ensemble mean for the DA simulations. However, it should be noted that the ensemble is generated sequentially, which means that all ensemble members are to some extent influenced by the assimilated data. How strong this effect is depends on how long the empirical information in the initial state for each decade is retained for in the system. It is also noteworthy that the ensemble mean is not a physically consistent state. The correlations between the DA ensemble mean and the PAGES 2K reconstructions are 0.73 for the Arctic, 0.52 for Asia, 0.92 for Europe and 0.71 for North America. These values are higher than the correlations of the PAGES 2K reconstruction with the "past1000" simulation, as the latter includes more random internal variability than the ensemble mean. The correlations of the DA analysis with the proxy-based reconstruction are highest. However, in the decades with strong influence of the volcanic forcing (1810–1820 AD and 1830–1840 AD), the ensemble mean is closer to the PAGES 2K data than the DA analysis. This

Fig. 1 Continental decadal mean temperature anomalies for 1750–1850 AD w.r.t. the 850–1849 AD mean in the NH, for the DA analysis (*blue line*), the PAGES 2K proxy-based reconstructions (*green line*), the ensemble mean (*magenta line*), the simulation without DA (*black line*) and the individual ensemble members (*yellow lines*)
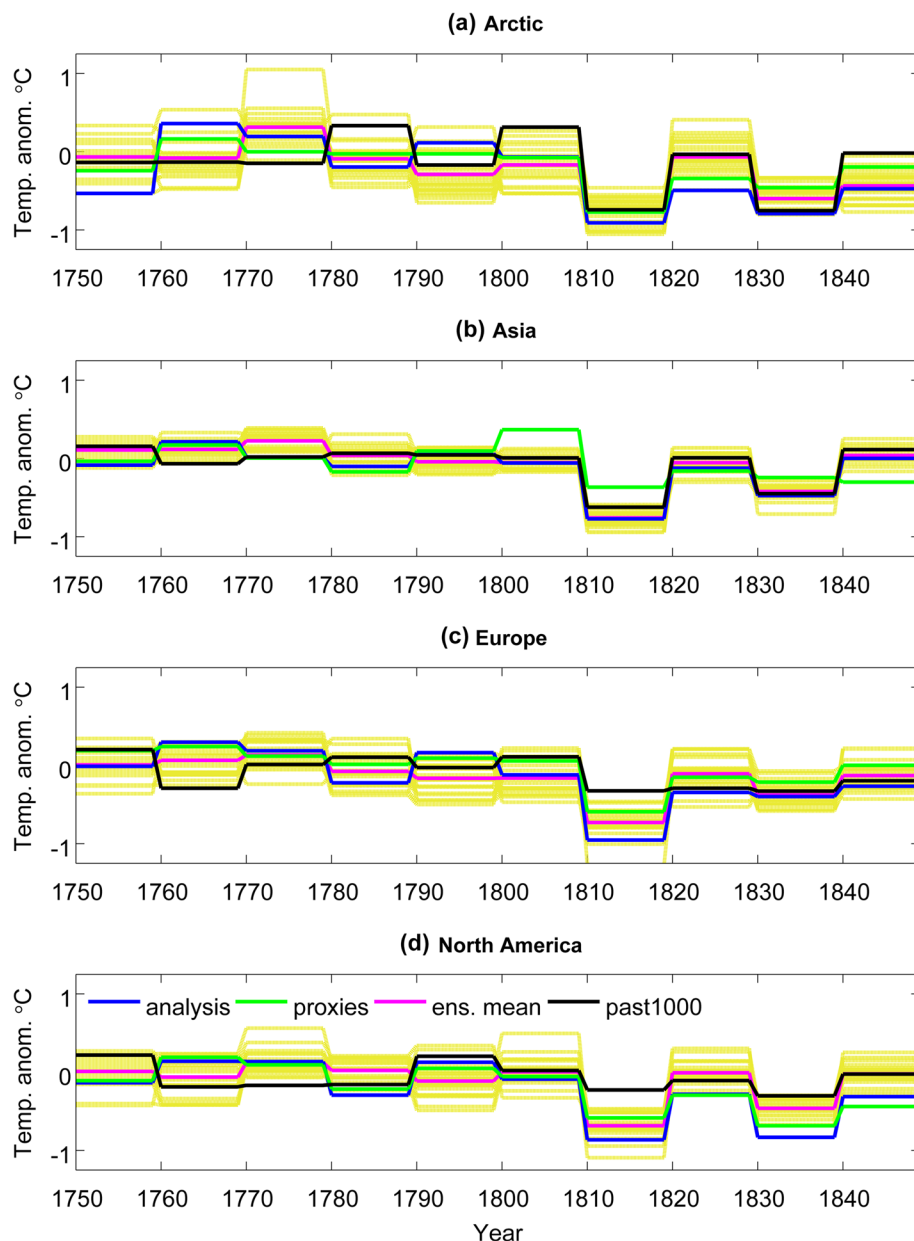

is likely to be due to the relatively small ensemble size, which does not always allow the best member to capture the true internal variability. For a DA analysis that includes the forcing signal and the wrong internal variability, the cost function can have a higher value than that of a simulation with the forcing signal and much lower internal variability, such as the ensemble mean.

We now examine the skill of the DA on shorter timescales (5-year running mean) in the NH continents, apart from North America for which the PAGES 2K reconstruction has decadal resolution (Fig. 2). Agreement between the DA analysis and the proxy-based reconstructions for the individual annual values is not expected, as only decadal average temperatures have been assimilated. However,

some overall agreement might be caused by the fact that the decadal averages are formed by the individual annual averages. The main possible source of consistency is the response of the model and the proxies to the forcings. Figure 2 indicates some agreement between the DA analysis and the assimilated temperatures, with moderate positive correlations (0.72 for the Arctic, 0.64 for Asia and 0.62 for Europe). These correlations are again higher than the correlations of the "past1000" simulation with the PAGES 2K reconstructions (0.63, 0.52 and 0.40 respectively). Thus, the response to the forcings, in particular to the volcanic eruptions, provides the main contribution to the correlations of the shorter timescales, but some small added value is obtained from DA.

Fig. 2 Continental temperature anomalies for 1750–1850 AD w.r.t. the 850–1849 AD mean, smoothed with a 5-year running mean, in Arctic, Asia and Europe, for the DA analysis (*blue line*), the PAGES 2K proxy-based reconstructions (*green line*) and the simulation without DA (*black line*). Brown *vertical lines* denote the major volcanic eruptions. *Yellow horizontal lines* indicate the control experiment mean anomalies and the one standard deviation range
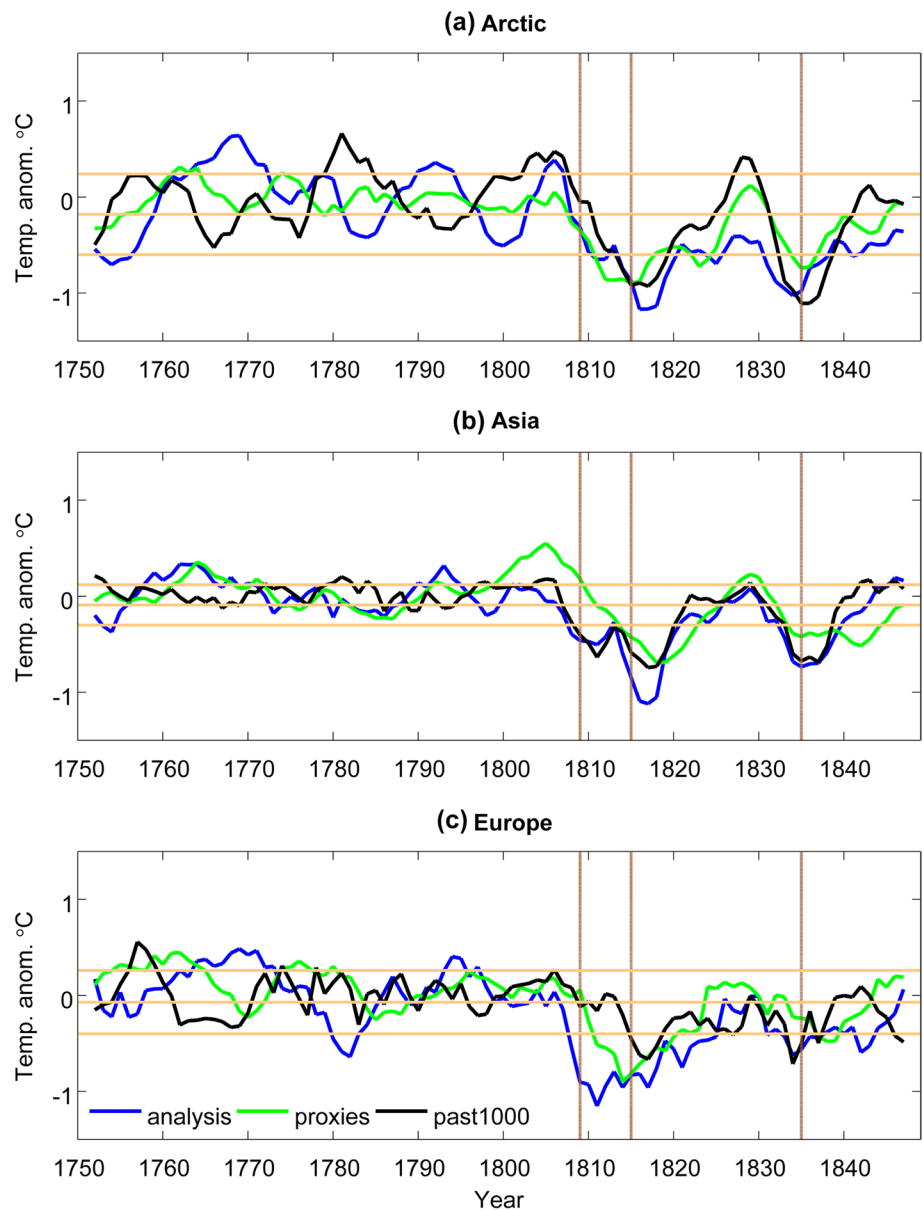
The control experiment mean anomalies and the one standard deviation range in Fig. 2 give an estimate of the internal variability. The differences between the DA analysis and the proxy-based reconstructions are smaller than the range of natural variability in some periods, showing the combined effect of the forcings and of the DA. The volcanic events of 1809–1815 AD leave a similar imprint in all continents, while the smaller scale event of 1835 AD greatly affects North America, where the volcano is located, and to a lesser extent the other continents. The eruption of Mount Tambora in April 1815 AD is the largest known historical strong tropical volcanic eruption (Oppenheimer 2003; Cole-Dai 2010) and followed the 1809 AD VEI-6 tropical eruption of unknown location (Cole-Dai et al. 2009), which occurred during the 1790–1830 AD Dalton solar minimum.

The fact that the decade 1810–1819 AD was the coldest during at least the past 500 years in the NH and the tropics (Cole-Dai et al. 2009) is attributed to the combined effects of the 1809 AD and Tambora eruptions. Approximately 10 years after the major eruptions, most regions revert back to the natural variability range.

3.2 Comparison with independent proxy data

Proxy-based temperature reconstructions for the NH for the last millennium differ significantly among each other and show discrepancies with simulations (Jansen et al. 2007). The period we examine, however, is more recent and experienced strong forcing variations, thus the agreement among the proxy-based reconstructions is higher than
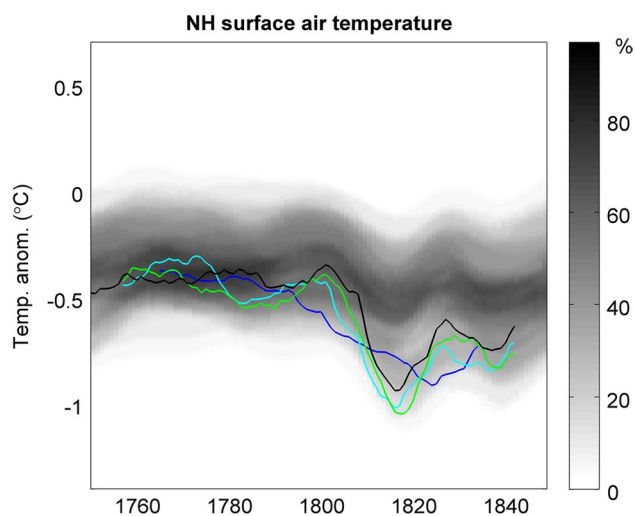
Fig. 3 NH 2 m temperature anomalies for 1750–1850 AD w.r.t. the 1961–1990 AD mean for the DA analysis (*blue line* for 31-year running mean and *cyan line* for 15-year running mean) in comparison with the simulation without DA (15-year running mean, *black line*) and the range of reconstructions (*grey scale*), redrawn from Jansen et al. (2007). Data time series and the *shaded* representation of overlap of proxy-based reconstructions (consensus) were obtained from: http://www.cru.uea.ac.uk/datapages/ipccar4.htm. The PAGES 2K direct average of the NH is also shown (15-year running mean, *green line*)

in previous times. Comparing the DA analysis with independent proxy-based reconstructions allows to detect errors due to unrealistic assimilated temperatures. Moreover, if the validation data include information from areas that have not been assimilated, e.g. hemispheric means, the comparison with independent data provides some evaluation of information propagation in space.

Figure 3 presents the NH near-surface (2 m) air temperature anomalies w.r.t. the 1961–1990 AD mean for the DA analysis in relation to the range of reconstructions redrawn from Jansen et al. (2007). The grey shading includes published multi-decadal timescale uncertainty ranges of all temperature reconstructions identified in Table 6.1 (except for RMO2005 and PS2004) of Jansen et al. (2007). The proxy-based time series are smoothed with a 31-year running mean. The reconstruction data used are those featured in Fig. 6.10 of the IPCC Fourth Assessment Report (Jansen et al. 2007). The DA analysis is presented in two ways; with a 31-year running mean time series, to be directly comparable with the overlap of the proxy-based reconstructions, and with a 15-year running mean series, to show the variability on shorter timescales. The 15-year running mean DA analysis falls outside the range of the proxy-based reconstructions, which have averaged all the short timescale variations out, but is in good agreement with the PAGES 2K direct average for the NH, also presented with a 15-year running mean (correlation 0.95).

The correlation of the run without DA with the PAGES 2K direct average is about the same as the latter (0.96). The DA skill on the hemispheric scale is thus very good, but no additional skill compared with the simulation without DA is gained. The 31-year running mean DA analysis lies well within the range of the proxy-based reconstructions. This is not a strict validation though, as the grey band is quite wide. We use the running mean filter to be consistent with the reconstructions, however in other parts of the study we prefer to use the Hamming window, which has better filter characteristics. The coldest anomalies in the simulation are between 1810–1820 AD and agree well with the consensus of reconstructions. The severe volcanic eruptions of 1809, 1815 and 1835 AD cause a sharp drop in the mean temperature during the years that follow, as well as a long-lasting effect on the NH climate, recorded in both proxies and the simulation. We note that although the proxy-based reconstructions of Fig. 3 seem to exhibit more multi-decadal variability than the DA, which is counter-intuitive given the tendency of regression-based reconstructions to reduce variance, this is not necessarily the case, as the grey shading shows the concentration of overlapping NH proxy-based reconstructions, which is different from individual trajectories. Additionally, not all of the proxy-based reconstructions are regression-based, as some of them have used the "composite plus scale" methodology, which does not affect the variance.

The DA analysis is also validated on the decadal timescale against the proxy-based European mean land temperature reconstruction by Luterbacher et al. (2004) (Fig. 4). The spatial patterns of this reconstruction are also available and are compared with the DA analysis later. Agreement between the DA analysis and the Luterbacher et al. (2004) independent reconstruction is higher in summer, but the DA analysis shows a stronger response to the volcanic eruptions than the proxies and larger variability in both seasons. The correlations are 0.73 for summer and 0.61 for winter. The upper panel also includes the summer decadal means of the PAGES 2K reconstruction. The correlation between the two reconstructions is 0.70. The correlations of the "past1000" simulation with Luterbacher et al. (2004) are 0.75 summer and −0.24 for winter, indicating again an additional skill of DA on the continental scale, at least in winter. Validation against independent data was also performed using the annual time series. Moderate positive correlations were found as expected, due to the forcings, but they were lower than the ones in the decadal case.

3.3 Comparison with instrumental data

An advantage of the investigated period is that some early instrumental records are available, which can be used to evaluate the assimilation method. Here we validate the

Fig. 4 European summer
(**a**) and winter (**b**) decadal
mean temperature anomalies
for 1750–1850 AD w.r.t. the
850–1849 AD mean, for the
DA analysis (*blue line*), the
simulation without DA (*black
line*) and the Luterbacher et al.
([2004](#)) proxy-based reconstruc-
tion (*green line*). The PAGES
2K proxy-based reconstruction
(*magenta line*) for the summer
period is also shown
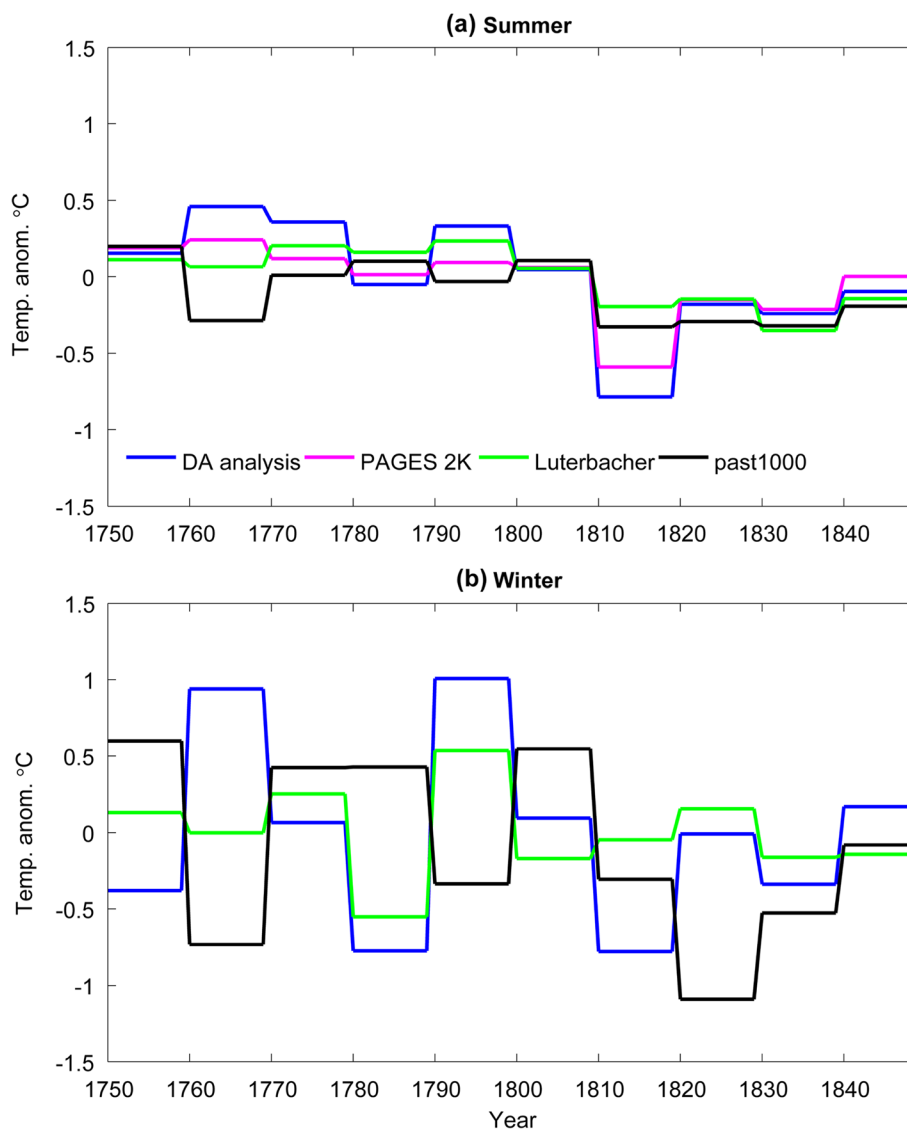
DA analysis against the instrumental BEST reconstruc-
tion (Rohde et al. [2012](#)) for global mean tempaeratures.
Figure [5](#) shows the global land 2 m air temperatures (anom-
alies w.r.t. the 1951–1980 AD mean), smoothed with a 9
point Hamming window (approximately corresponding to
a 5-year running mean), as simulated with and without DA,
compared with the reconstruction from the BEST dataset.
All three time series remain relatively constant over the
period, interrupted by the cooling induced by the volcanic
eruptions. The simulated DA time series includes minima
in the 1810s and 1830s temperatures and maxima in the
1770s and 1800s. The correlation between the DA analy-
sis and the BEST dataset is high (0.81), and the correlation
of the BEST dataset with the "past1000" run without DA
(0.79) is only marginally lower. The correlation of the DA
analysis with the "past1000" simulation without DA is very
high (0.88).

Overall, the performance of the DA scheme on the
global scale is very good, but DA does not offer additional
skill compared with the simulations without DA. This is
similar to what has been found on the hemispheric scale
examined earlier. The lack of added value of DA is likely
due the influence of the strong forcings and the high ratio
of forced to internal variability on large spatial scales. In
principle, if the forced response in a model is unrealistic,
DA can select an ensemble member that, due to internal
variability, is closer to reality. Therefore, DA can correct to
some extent for unrealistic response to forcing. However, if
the response to the forcing is close to reality and the con-
tribution of internal variability is low, then no systematic
change to the response to the forcing can be expected from
DA. On the continental scale, the contribution of interan-
nual variability is larger, and if a DA scheme works prop-
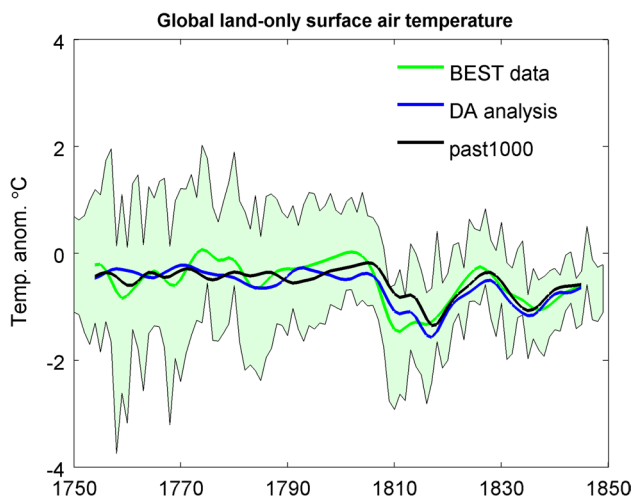erly, it can give additional skill compared with a simulation

Fig. 5 Global land surface air temperatures for 1750–1850 AD (anomalies w.r.t. the 1951–1980 AD mean), smoothed with a nine point Hamming window, for the DA analysis (*blue line*), the simulation without DA (*black line*), and the BEST instrumental dataset (*green line*). The green shading shows the 95 % confidence interval of the BEST estimate, representing statistical and spatial undersampling uncertainties

Fig. 6 Evolution of the NH surface air temperature anomalies for 1750–1850 AD w.r.t. the 1961–1990 AD mean, smoothed with a nine point Hamming window, for the DA analysis (*blue line*) and other GCM simulations without DA. The PAGES 2K direct average of the NH (*green line*) is also included (for Arctic, Asia and Europe, *green line*)

without DA. The signal (response to the forcing) to noise (internal variability) ratio increases with the spatial scale. Therefore, forcings are sufficient to provide realistic simulations on hemispheric and global means, which have a relatively small contribution from internal variability.

3.4 Comparison with other GCMs

To further examine whether DA affects the long timescale (decadal to centennial) variability, we compare the NH surface air temperature anomalies of the DA simulations with other GCM simulations that have not used DA (Fig. 6). This aims to investigate the behaviour of DA with respect to the forcings and the internal variability, in comparison to simulations without DA. The simulations were part of the CMIP5 or the PMIP3. We display the simulations performed with the models CCSM4 (Community Earth System Model, developed at NCAR, USA), HadCM3 (Hadley Centre Coupled Model, version 3, developed by the Hadley Centre, UK), IPSL-CM5A-LR (Institute Pierre Simon Laplace model, France), and MPI-ESM-P (Max Planck Institute PMIP3 simulations). The simulation performed with the model we use in this study, MPI-ESM-CR, without DA, is also shown.

The correlation of the DA analysis hemispheric means with the PAGES 2K NH direct average for filtered (nine point Hamming window) temperatures is 0.85. Data for North America could not be included in the direct average because of their decadal resolution. The DA simulation
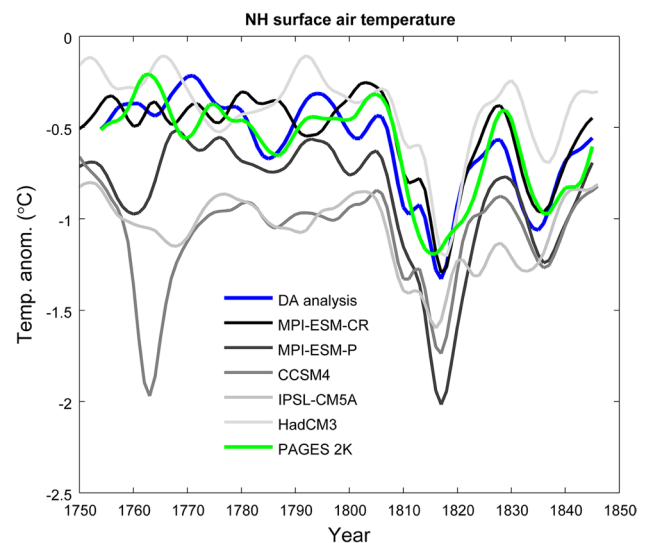
is not closer to the proxy-based reconstruction than the non-assimilated MPI-ESM-CR and MPI-ESM-P simulations (correlations with PAGES 2K are 0.86 and 0.83 respectively). This shows again that the forcings play the dominant role in giving the skill on the hemispheric scale. The correlation of the DA analysis with the proxy-based reconstruction is higher than the correlations of the CCSM4 (0.24), IPSL-CM5A (0.64) and HadCM3 (0.80) simulations. The DA analysis arrives at slightly lower temperatures compared with the MPI-ESM-CR without DA towards the end of the simulation period. Multi-decadal variability is similar in most simulations.

4 Circulation and temperature variability in the North Atlantic-European sector

One of the main reasons for performing DA is to obtain knowledge about variables that are not assimilated. We now investigate whether this is the case with respect to small-scale temperature variability in Europe and leading modes of circulation variability. The link between continental mean temperatures, which have been used in our assimilation, and large-scale circulation, is crucial for providing this added value. We examine the temperature-sea level pressure (SLP) link in the control simulation using MCA, and check whether it is also visible when performing the assimilation.

4.1 European temperature patterns

In our DA scheme, only continental average temperatures have been assimilated. Hence, it is not guaranteed that skill on smaller spatial structures exists, since no information about the local temperatures was inserted in the model. The assimilation of the NH continental averages might however provide information that could potentially determine the state of leading circulation modes, which in turn may lead to skill in reconstructing the smaller-scale spatial patterns. Two possible reasons may lead to added value of DA on these scales: firstly, leading modes of variability, such as the NAM or the NAO, may have a link to the NH continental mean temperatures and be partially captured correctly with DA. Secondly, certain temperature mean values may be associated with tendencies for specific spatial patterns, in which case no explicit atmospheric circulation information is required. This would be the case if, for example, a low mean European temperature was associated with very low Northern European temperatures, less cold Southern European conditions, mild Eastern Europe temperatures etc. This hypothesis is not examined in the current study. In addition, another possibility for agreement on the small spatial scales could stem from the response of the circulation to forcing, such as the strong volcanic eruptions.

We compare the DA-simulated European temperature patterns with the patterns from the Luterbacher et al. (2004) reconstruction. Figures 7 and 8 show the European decadal 2 m temperature maps (anomalies w.r.t. the 1750–1849 AD mean) for the DA analysis and the reconstructions by Luterbacher et al. (2004) for the summer and winter periods of 1750–1850 AD respectively. The comparison of the patterns in most decades exhibits no agreement. If the Luterbacher et al. (2004) reconstructions are skilful, we can conclude that although the DA scheme is skilful in reconstructing the large-scale temperatures, it has no skill in capturing the spatial patterns within Europe correctly. The spatial correlations found for each decade are shown in Table 1. The mean correlations over all the decades are negligible (0.03 for summer and −0.03 for winter). We note that regression generally underestimates variability, so the stronger variability appearing in the simulations than in the proxy-based reconstructions is to be expected.

4.2 NH Modes of variability

The atmospheric circulation variability in the North Atlantic-European sector is linked to modes of variability such as the NAO or the NAM. The NAO is one of the major modes of atmospheric variability in the NH and strongly influences the wintertime temperature over much of Europe. It is related to the pressure difference between Iceland and Azores, which generates the westerly winds that characterize the atmospheric circulation in the North Atlantic at mid-latitudes. The NAM, which is the dominant mode of the extratropical atmospheric circulation in the NH, is closely related to the NAO. We compare the decadal mean values of the simulated winter (December–March) NAO index with the available NAO reconstructions, namely the proxy-based reconstruction by Luterbacher et al. (2002) and the UEA Climatic Research Unit instrumental-based reconstruction by Jones et al. (1997) (Fig. 9).

Luterbacher et al. (2002) estimated the NAO index based on a Principal Component Regression Analysis, as the standardized (1901–1980) difference between the SLP over the Azores and over Iceland. Jones et al. (1997) used early instrumental data to calculate the NAO index as the difference between the normalised SLP over Gibraltar and the normalised SLP over Southwest Iceland, and extended this index back to 1823 AD. Our NAO index (difference between the normalised SLP over Gibraltar and over Southwest Iceland), is not compatible with the two reconstructions, partly explaining the inconsistency between the simulated European temperature patterns and the reconstructions by Luterbacher et al. (2004). The correlation with Luterbacher et al. (2002) is −0.09. The annual NAO index (not shown) does not show any agreement either (correlation 0.21). We did not analyse the summer circulation as the NAO is less pronounced in this season. Zanchettin et al. (2012), using European climate reconstructions, found that the dynamic response to the volcanic eruptions is a decadal-scale positive phase of the winter NAO, accompanied by winter warming over Europe peaking approximately one decade after a major eruption. There is no consensus among the NAO reconstructions however. Luterbacher et al. (2002) and Jones et al. (1997) reconstructions are also very different to each other, showing the limitation of proxy-based reconstructions, which affects the skill of DA too.

4.3 Link between continental temperatures and large-scale circulation

To explain the disagreement between the DA analysis and the proxy-based reconstructions regarding the temperature patterns over Europe and the NAO index, we use MCA to investigate the link between the NH continental mean temperatures and the SLP field. We analyse 1000 years of the control MPI-ESM-CR simulation to focus on internal variability and to avoid the influence of forcings, because the main purpose of DA is to capture the internal variability. We use simulated decadal temperatures that represent the same continental regions and seasons as the PAGES 2K reconstructions, and NH decadal averages for the SLP field, based on annual and seasonal values. Because the PAGES
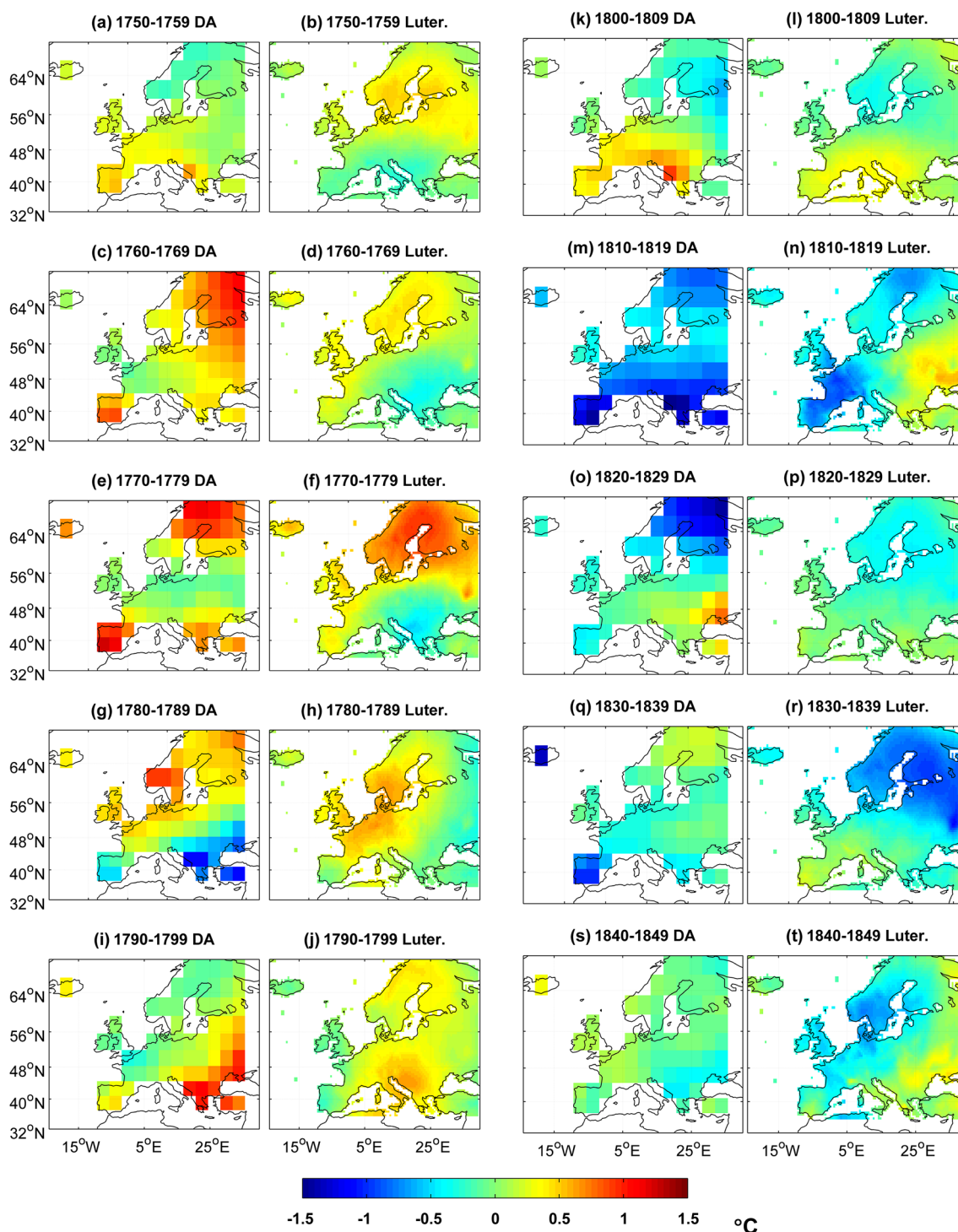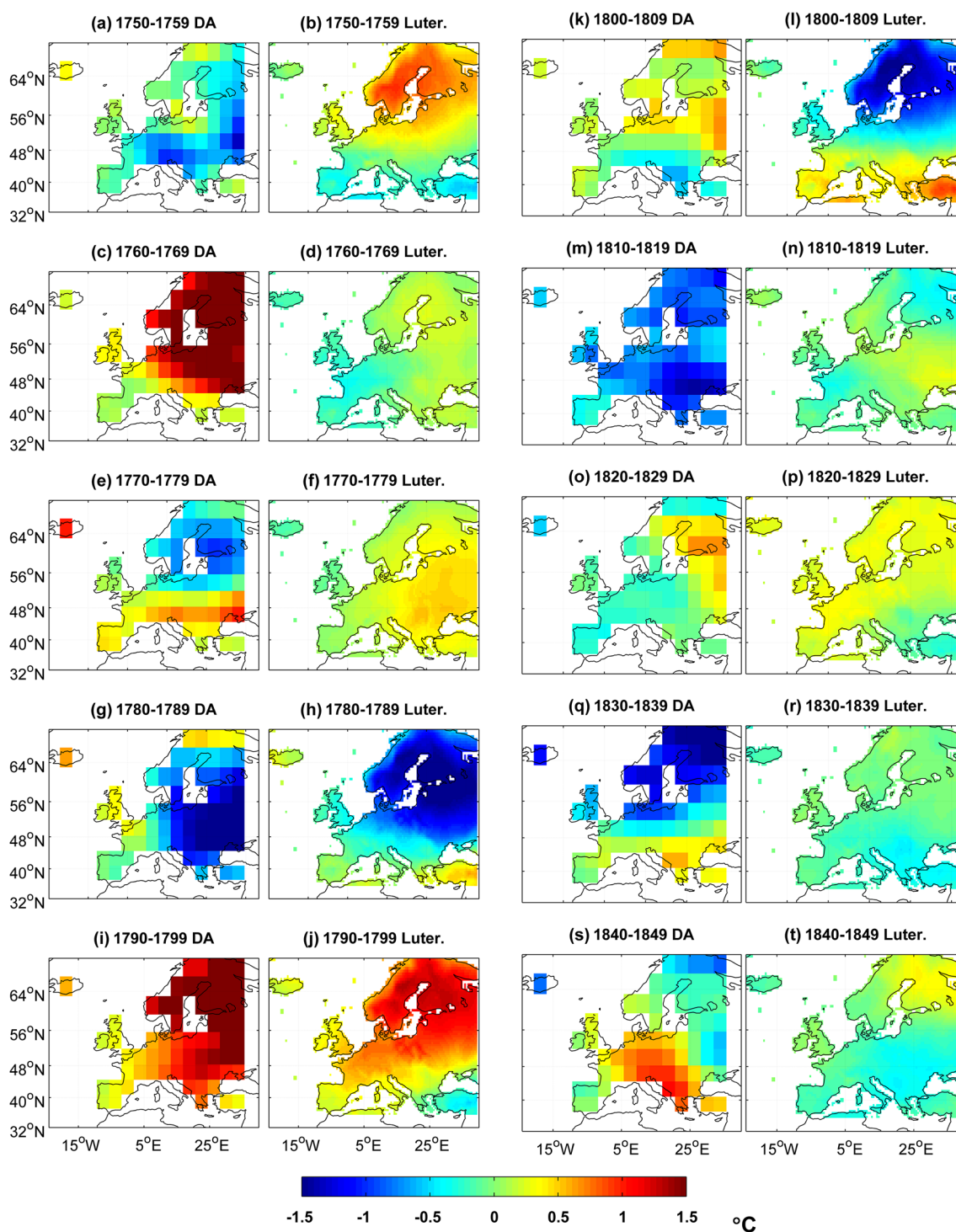
Fig. 7 European decadal surface air temperatures (anomalies w.r.t. the 1750–1849 AD mean) for the DA analysis and the Luterbacher et al. (2004) reconstruction, for the summer of 1750–1850 AD

2K temperatures are seasonally mixed, it is not clear in which seasons the link of the seasonally mixed temperatures to the atmospheric flow is strongest, therefore we separate our analysis for the different seasons and the annual means. In all five cases the temperature data are the same.

MCA applies a Singular Value Decomposition (SVD) to the cross-covariance matrix between two data sets to find pairs of patterns whose time expansion coefficients (TECs) have maximum covariance (Bretherton et al. 1992). In our case, the dimension of the cross-covariance matrix

Fig. 8 European decadal surface air temperatures (anomalies w.r.t. the 1750–1849 AD mean) for the DA analysis and the Luterbacher et al. (2004) reconstruction, for the winter of 1750–1850 AD

is 4×2304 (number of continental temperature means $\times$ number of grid-cells for the SLP field). To account for the different size of the grid-cells, the SLP anomalies field (w.r.t. to the 1000-year control average), SLP_i, has been weighted with weights w_i proportional to the square root of

the grid-cell size, prior to calculating the cross-covariance matrix. The weights have also been used for calculating the TECs. For the continental temperature anomalies, T_i, no area weights have been used. The temperature and SLP singular vectors, which are the MCA coupled patterns, are

Table 1 Spatial correlations between the DA analysis and the Luterbacher et al. (2004) reconstructions for European summer and winter, for the decades 1750–1759 AD (1st) to 1840–1849 AD (10th)

Decade	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
Summer	−0.62	0.19	0.21	0.59	0.02	0.70	0.09	0.40	−0.56	−0.70
Winter	−0.05	0.69	−0.04	0.48	0.77	−0.63	−0.31	0.02	−0.70	−0.54
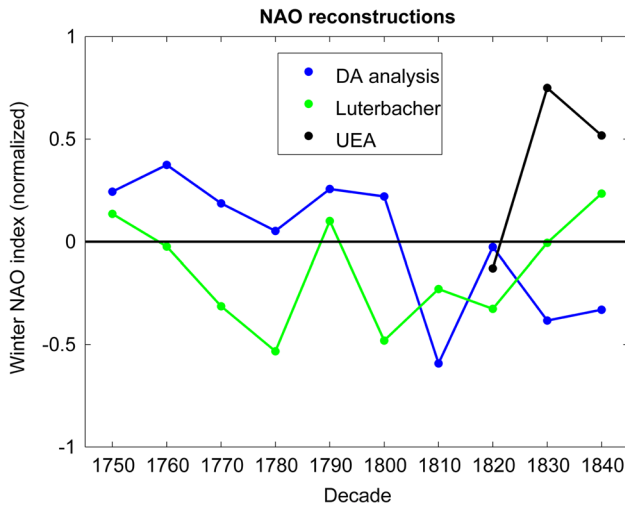
Fig. 9 Normalized decadal mean winter NAO index for 1750–1850 AD, for the DA analysis (*blue dots*), the proxy-based reconstruction by Luterbacher et al. (2002) (*green dots*) and the instrumental reconstruction by Jones et al. (1997) (*black dots*)

denoted by the vectors u_k and v_k respectively. The TECs (a_k and b_k) are given through area-weighted orthogonal projections of the data anomalies (T_i and SLP_i) onto the patterns, such that:

$$a_k(t_j) = \sum_{i=1}^{4} T_i(t_j)\, u_{ik} \tag{2}$$

$$b_k(t_j) = \sum_{i=1}^{2304} w_i\, SLP_i(t_j)\, v_{ik} \tag{3}$$

Figures 10 and 11 show the first coupled MCA patterns of the NH continental mean temperatures and the NH SLP respectively. The patterns are displayed as the singular vectors multiplied by the standard deviation of their TECs in order to include information about their amplitude. For the annual means, winter and to a lesser extent for spring, the SLP MCA patterns look similar to the leading Empirical

Fig. 10 MCA between NH continental mean temperatures (seasons defined by PAGES 2K) and NH SLP (annual and seasonal means) from the MPI-ESM-CR control simulation: first temperature patterns for the Arctic, Asia, Europe and North America
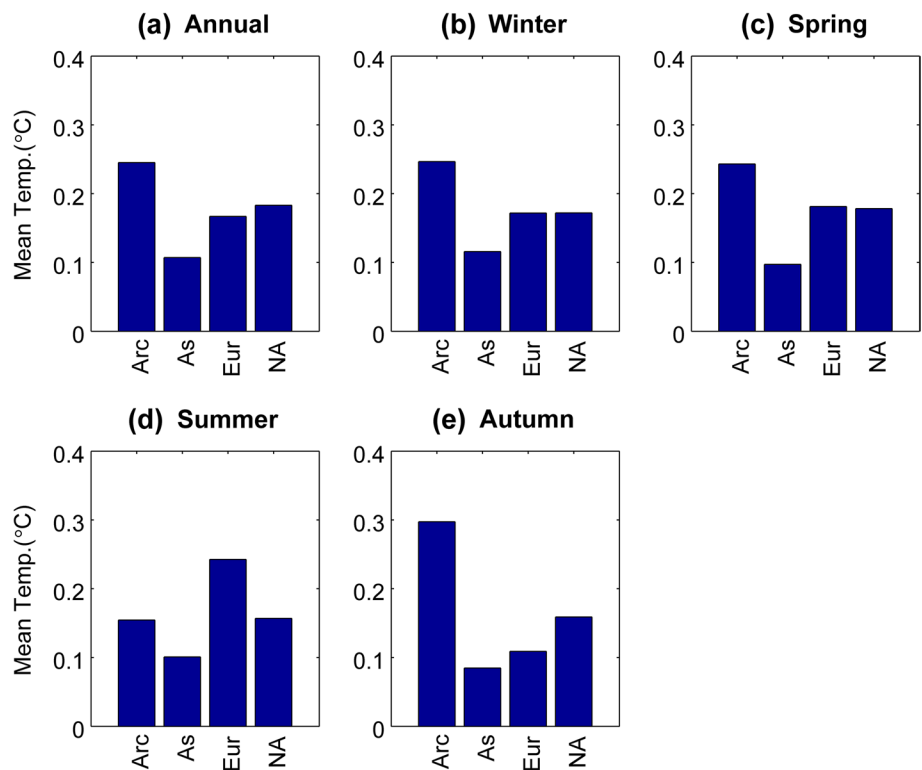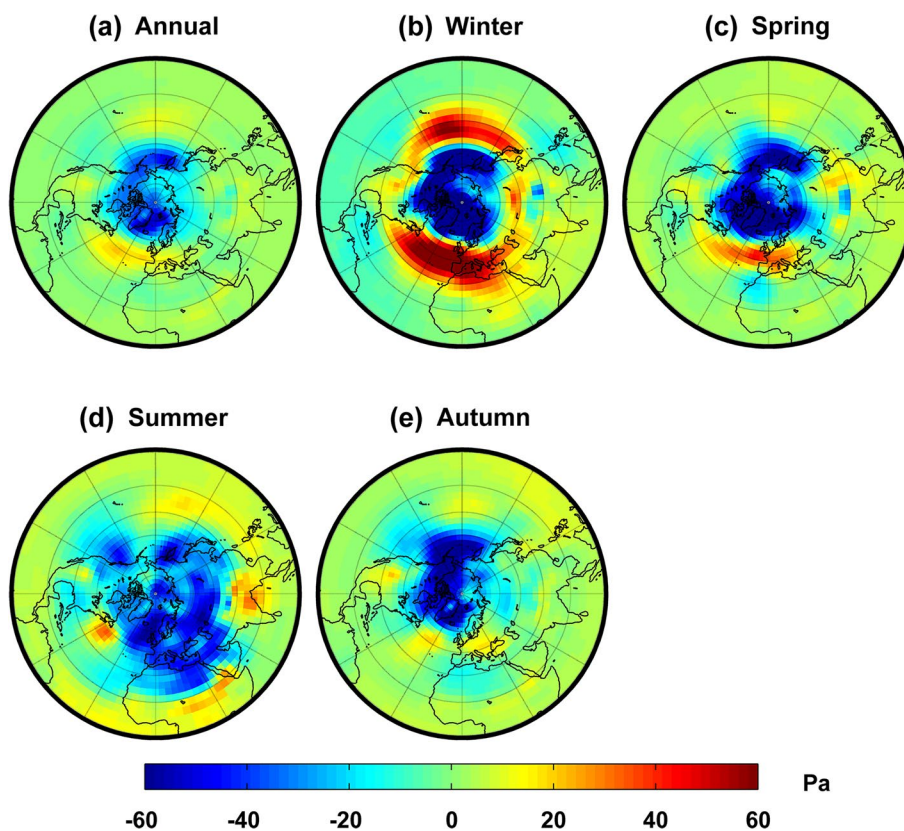
Fig. 11 MCA between NH continental mean temperatures (seasons defined by PAGES 2K) and NH SLP (annual and seasonal means) from the MPI-ESM-CR control simulation: first SLP patterns

Orthogonal Function (EOF) of the NH SLP field in the control run, shown in Fig. 12. For summer and autumn the SLP MCA patterns are different from the leading EOF. Looking at the EOF structures, in all seasons except summer, the SLP leading EOF patterns resemble the annular mode patterns. This mode is basically the NAM, but it should be noticed that our EOF analysis is based on decadal SLP anomalies in the region 0N to 90N, whereas the standard definition of the NAM is based on the first principal component of monthly SLP anomalies in the region 20N to 90N. In summer, the SLP leading EOF shows a pronounced land-sea contrast, which is presumably linked to the intensity of the land-sea temperature contrast. From the above, we can conclude that for the annual means, winter and to a lesser extent for spring, the NAM is the pattern that is most closely linked to the NH continental temperatures. Hence, by assimilating continental temperatures, the state of the NAM should be determined to some extent. For summer and autumn, the MCA patterns do not resemble the NAM but have a wavelike structure. The correlations between the temperature and SLP TECs are 0.73 for the annual case, 0.48 for winter, 0.51 for spring, 0.36 for summer and 0.54 for autumn. Given the correct continental mean temperatures, the explained variance for the SLP TECs is 0.73^2 for the annual means, i.e. 52 %, and much less for the seasonal means. In our case the estimate for the

SLP TECs will actually explain <52 % variance because the PAGES 2K reconstructions are different from reality, as they include noise. We also analysed higher coupled patterns. The correlations between their TECs were lower, and thus we focus on the first pair of coupled patterns.

To test whether the correlations between the two sets of TECs are significant, we applied a Monte Carlo method to calculate the distribution of correlations in the case of no link. We created a 1000-year random temperature array for the four continents, 300 times, and performed a MCA analysis of the link between these random temperatures and the control run's SLP field. In the annual case, the mean TEC correlation of the randomly sampled temperatures with the SLP field is 0.39, ranging between 0.25 and 0.55. These correlations are lower than the value of 0.73 of the case in which the simulated temperatures were used for the MCA. Similar results are found for all four seasons, but the largest difference is noticed for the annual means. In all cases, the correlations of the simulated control run temperature TECs were either outside or at the very end of the randomly sampled distribution. The test shows that the link found between the NH continental mean temperatures and the SLP TEC in the control run is significant. The same Monte Carlo experiment was performed for higher coupled patterns, the TEC correlations of which remain relatively significant but are gradually reducing.

4.4 Link between temperatures and large-scale circulation in the assimilation run

We now investigate whether the link between the NH continental average temperatures and the SLP field that has been found in the control run is also visible when performing the DA. The link between temperatures and SLP can be caused by two processes: (1) for a given SLP anomaly the temperature is influenced through advection; and (2) given temperature anomalies might directly influence the pressure field, for instance through changes in the thickness of atmospheric layers and the associated upper-level flow anomalies which in turn affect the surface pressure. The former process can be expected to be dominant. Independent of the type of process that is relevant, the DA may reproduce the link because it is just selecting states rather than prescribing temperature anomalies, and the model captures both process types.

To examine whether the DA actually reproduces the link, we project the DA decadal NH temperature and SLP anomalies of the period 1750–1850 AD onto the control experiment's temperature (u_k) and SLP (v_k) MCA patterns respectively, to get temperature and SLP TECs for the assimilation period, and calculate the correlation of the TECs. The anomalies were calculated w.r.t. the means of the control run to be consistent with the previous analysis.

The correlations are 0.81 for the annual means, 0.82 for winter and spring, 0.17 for summer and −0.01 for autumn. After detrending the TEC time series, the correlations were only slightly reduced (0.70 for the annual means, 0.73 for winter, 0.70 for spring, 0.02 for summer and −0.21 for autumn). The fact that the difference between detrended and non-detrended correlations is very small shows that the correlations are mainly due to the decadal variations. When the simulated temperatures in the analysis are replaced by the PAGES 2K temperatures, which have been assimilated, very similar correlations are obtained. This can be expected because, as shown earlier, the continental mean temperatures in the analysis follow closely the assimilated temperatures. The above analysis shows that the link between the NH continental mean temperatures and the SLP field in the assimilation run exists in winter, spring and the annual means, but not in summer and autumn.

A question that arises from the analysis is why the link that we found between temperatures and SLP in the control run is not reproduced in summer and autumn in the assimilation run. To investigate these seasonal differences, we take a closer look at the SLP MCA patterns (Fig. 11). We note that the link in the assimilation run appears in the seasons when the SLP MCA patterns are similar to the leading SLP EOFs, which in turn resemble in these seasons the NAM. The link does not appear when the MCA pattern

Fig. 12 Leading EOF of the NH SLP decadal means in the MPI-ESM-CR control simulation, for the different seasons

(a) Annual (b) Winter (c) Spring

(d) Summer (e) Autumn

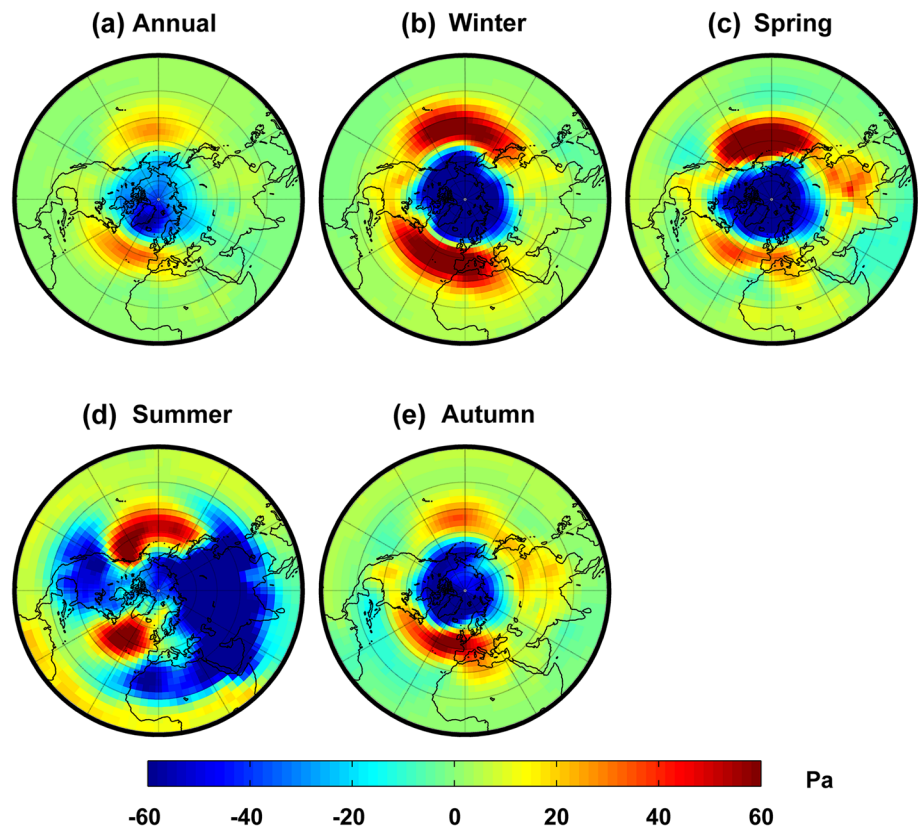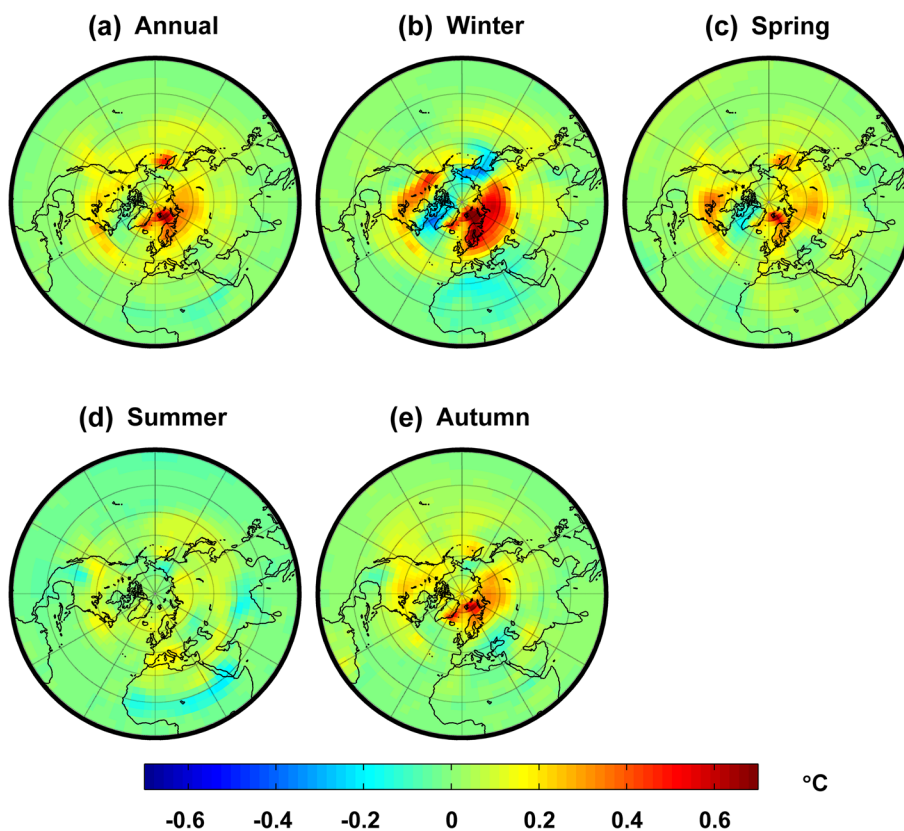-60 -40 -20 0 20 40 60 Pa

Fig. 13 Regression maps of the MPI-ESM-CR control simulation's grid-point NH temperatures onto the SLP MCA TECs, multiplied by one standard deviation of the SLP MCA TECs, for the different seasons
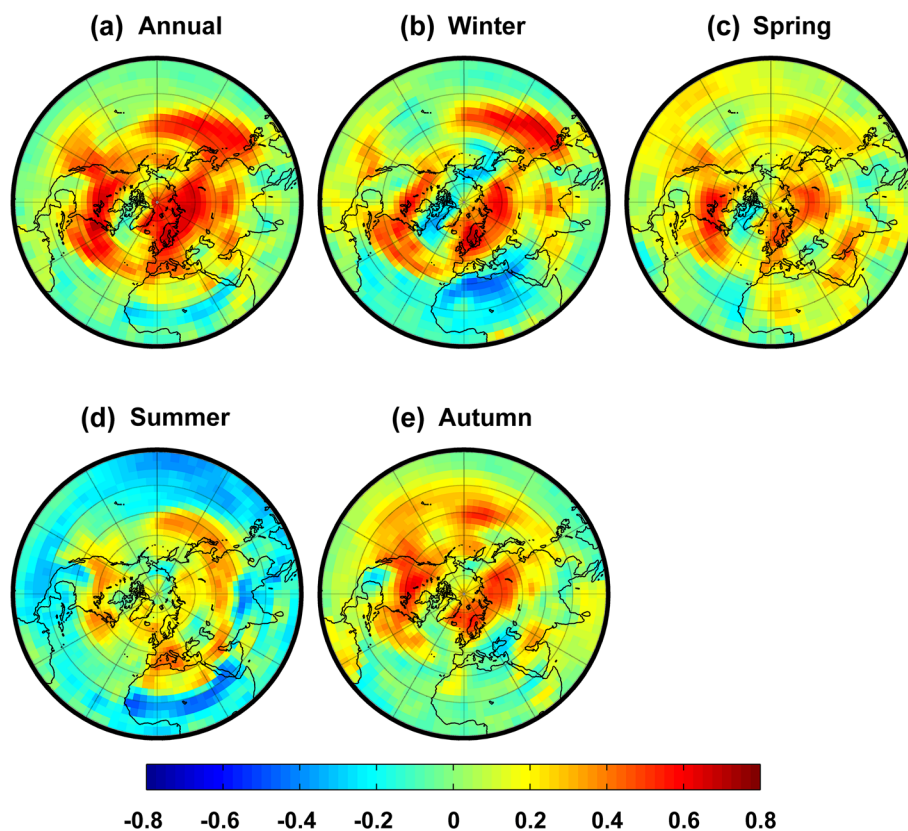

does not look like the leading EOF. This observation hints to a potential explanation; it seems plausible that it is easier in a small ensemble to find analogues for circulation and temperature anomalies that resemble the dominant mode of variability than finding good analogues for anomalies that are different from the leading EOFs. The reason for this might be that the simulated anomalies have a tendency to look like the leading EOF patterns or linear combination of these, while they are usually not similar to patterns that are not in the space of the leading EOFs. Moreover, the link in summer is weaker than the other seasons, as the TECs correlations reveal.

The MCA thus shows that the NH continental mean temperatures for the PAGES 2K regions are most closely linked to the NAM in winter, spring and the annual means, and that this link is reproduced in the DA for the same seasons. However, this does not imply that the continental temperatures are the best predictors for the NAM. To investigate this further, we have regressed the control run's grid-point NH temperatures onto the first SLP TECs. As shown in Widmann (2005), these regression coefficients are proportional to the weights for the temperature field that lead to an optimal linear estimation of the SLP MCA TEC in a one-dimensional MCA. Therefore, temperature information from areas with high values are needed for obtaining good estimates for the SLP MCA TEC. Figure 13 shows

the regression maps of the control run NH temperatures onto the first SLP MCA TECs, multiplied by one standard deviation of the SLP TECs. It shows that this TEC, which is similar to the NAM index for annual data, winter and spring, has a strong temperature signal only at the very high latitudes, e.g. Scandinavia for the case of Europe. Similarly, the temperature signal of the TEC is larger at the northern parts of North America and Asia.

The correlation maps between the control run's local NH temperatures and the leading SLP MCA TECs for the different seasons (Fig. 14), show that only at the very high latitudes, the SLP TECs have relatively high correlations (up to 0.7) with the local temperatures. The correlations at the mid-latitudes range from 0.2 to 0.4 for the annual means, and even lower values for the seasonal means. This means that a large amount of the local temperature variance cannot be explained by the leading SLP TEC. For instance, only 49 % of the local temperature variance can be explained at the very high latitudes given a perfect estimation of the SLP TEC, which is similar to the NAM index in the annual case, and almost no variance at the low and mid latitudes. This designates that even if the true amplitude of the leading MCA circulation pattern were simulated, the local temperatures would still not be strongly confined in many areas. It can further be noticed that the local temperature correlation and regression coefficients over Europe have

Fig. 14 Correlation maps between the MPI-ESM-CR control simulation's grid-point NH temperatures and the SLP MCA TECs, for the different seasons

(a) Annual   **(b) Winter**   **(c) Spring**

(d) Summer   **(e) Autumn**

-0.8   -0.6   -0.4   -0.2   0   0.2   0.4   0.6   0.8

the same sign everywhere, which means that the variability in the SLP TEC only leads to limited variability in temperature gradients across Europe.

To summarise, the continental temperatures for the NH PAGES 2K regions are not optimal for confining the amplitude of the leading MCA SLP pattern, which for annual values, winter and spring are similar to the NAM. The reason is that the temperature signals of the SLP MCA patterns are mostly concentrated in high latitudes, hence the continental mean temperatures include information from regions not related to these circulation anomalies. This limits the correlation between the temperature TECs and SLP TECs. In the DA this can be expected to lead to circulation states that are not strongly constrained by the assimilated continental temperatures, and in turn to a limited potential for realistically simulating small-scale temperature variability. Moreover, for summer and autumn the link between NH continental temperatures and the SLP field identified by MCA is not reproduced in the DA. This might be due to the fact that the leading SLP MCA pattern is not similar to the leading SLP EOF in these seasons, which in turn may lead to sampling problems if a small ensemble size is used. We have also shown that even if the amplitudes (i.e. the TECs) of the MCA SLP pattern were perfectly captured in the DA the local explained temperature variance would be limited.

5 Summary and conclusions

We have assimilated the NH continental decadal mean temperatures from the PAGES 2K reconstructions using the MPI-ESM model to get a climate reconstruction for 1750–1850 AD. The skill of the DA-based reconstruction was firstly evaluated against the assimilated data and other proxy-based temperature reconstructions on large spatial scales. We also validated the DA analysis for smaller-scale temperatures for Europe and for the NAO, to see whether there is added value from the DA. The assimilation showed good skill for large-scale temperatures. With respect to the NH continental means, the DA analysis followed closely the assimilated proxy-based reconstructions on the decadal timescale. Skill was also found on shorter timescales due to the strong forcings, especially the major volcanic eruptions. The DA was shown to provide added value on the continental scale, especially on the decadal, but also on shorter timescales, when compared with simulations without DA. The DA analysis was in agreement with independent proxy data and early instrumental records for global and hemispheric-scale temperatures too, but no additional skill of the assimilation was found due to the dominance of forcings on these scales. Although no information about the smaller-scale spatial temperature anomalies was assimilated, the assimilation of the NH continental average

temperatures might lead to some skill in simulating small-scale spatial temperature patterns because of the potential capturing of information that may determine the state of leading circulation modes, such as the NAM or the NAO. However, no agreement between the DA analysis and the proxy-based European temperature reconstruction of Luterbacher et al. (2004) or the NAO index reconstruction of Luterbacher et al. (2002) was found.

To examine the reasons why the DA does not provide added value on smaller spatial scales, we calculated the link between NH continental temperatures and the NH SLP field using MCA and decadal means from the MPI-ESM control simulation. The temperature data for the different continents represented different seasons or annual data in accordance with the seasons that had been used for the PAGES 2K reconstructions, while the SLP data were taken for each of the seasons and for annual data, and the MCA was performed separately for the five cases. The analysis showed that for annual, winter and spring SLP, the circulation pattern that is most closely linked to the NH continental mean temperatures strongly resembles the NAM, while for summer and autumn it is a wave-like pattern. The correlations between the temperature and SLP TECs are 0.73 for the annual data, 0.48 for winter, 0.51 for spring, 0.36 for summer and 0.54 for autumn. These correlations indicate potential for constraining to some extent the large-scale circulation if the continental temperature means are known, which in turn might lead to some skill for small-scale temperature variability in the DA simulations. However, the validation results showed that this potential was not achieved.

To further investigate the lack of small-scale skill in the DA simulation we also checked whether the link found in the control run is visible in the DA simulations and found this to be the case for annual, winter and spring SLP, but not for the other seasons. This is possibly due to the fact that for the former, the NH continental temperatures are linked to SLP anomalies that resemble the first SLP EOF (which in these seasons is similar to the NAM), which might be better captured in DA with small ensemble sizes than variability patterns that are less dominant. It was also shown that even if the amplitudes of the leading MCA SLP pattern were perfectly known, the local explained temperature variance would still be less than 50 % in the best case and much lower in many areas.

In summary the analysis suggests the following potential reasons for a lack of skill of the DA in simulating small-scale temperature variability: (1) the link between NH continental temperatures and large-scale atmospheric circulation anomalies might be too weak to sufficiently constrain the large-scale circulation; (2) the small ensemble size might make it difficult to find ensemble members that resemble the real circulation state that led to the PAGES 2K continental

temperature anomalies; (3) noise and errors in the PAGES 2K temperature reconstructions contribute to unrealistic large-scale circulation states in the DA simulations; and (4) the local temperature variance explained by the large-scale circulation anomalies that can be estimated from the NH continental temperatures is substantially limited.

These potential contributing factors to a lack of small-scale skill give some guidance for improvements of the DA method. Firstly, the link between assimilated proxy-based temperature reconstructions and the large-scale circulation can in general be expected to get stronger if local or regional rather than continental-scale reconstructions are used, provided they cover suitable areas. However, it should be noted that with decreasing spatial scale of the reconstructions their error can be expected to increase, as there is less cancellation of random errors. Determining the optimal scale for the temperature data to be assimilated is thus a challenge. In this context we also point out that the location of most of the proxy records that have been used for the PAGES 2K reconstructions are not in the high-latitude regions that have the strongest NAM signal, and thus the reconstructed continental temperature variability may have a disproportionally low contribution from these areas. For example, the European PAGES 2K mean temperature is reconstructed from proxies mostly in the southern and central parts of the continent, which are not sensitive to NAM variations, and fewer proxies in Scandinavia and Northern Europe. Using temperature reconstructions that provide a stronger constraint for the atmospheric circulation may also lead to constraining several modes of circulation variability and thus increase the local explained temperature variance. Secondly an increase in ensemble size will make it more likely to find simulated climate anomalies that are similar to the real world. The required sample size to have a good chance for finding close analogues depends on the dimensionality of the state space in which the analogues are defined, with the required sample size increasing very strongly with the dimensionality. This is an argument for using low-dimensional state spaces, such as the four continental means we have used, rather than temperature reconstructions with a high spatial resolution. However, other methods for reducing dimensionality such as principal component analysis might be an alternative to using continental averages and might lead to a stronger link with the atmospheric circulation. Finally, retaining more than one ensemble member, i.e. using a non-degenerated particle filter as in Goosse et al. (2012) might improve the performance of the DA.

References

Annan JD, Hargreaves JC (2012) Identification of climatic state with limited proxy data. Clim Past 8:1141–1151

Bhend J, Franke J, Folini D, Wild M, Broennimann S (2012) An ensemble-based approach to climate reconstructions. Clim Past 8:963–976

Bretherton CS, Smith C, Wallace JM (1992) An intercomparison of methods for finding coupled patterns in climate data. J Clim 5:541–560

Cole-Dai J (2010) Volcanoes and climate. Wiley Interdiscip Rev Clim Change 1:824–839. doi:10.1002/wcc.76

Cole-Dai J, Ferris D, Lanciki A, Savarino J, Baroni M, Thiemens MH (2009) Cold decade (AD 1810–1819) caused by Tambora (1815) and another (1809) stratospheric volcanic eruption. Geophys Res Lett. doi:10.1029/2009gl040882

Crespin E, Goosse H, Fichefet T, Mann ME (2009) The 15th century Arctic warming in coupled model simulations with data assimilation. Clim Past 5:389–401

Crowley TJ, Unterman MB (2012) Technical details concerning development of a 1200-yr proxy index for global volcanism. Earth Syst Sci Data Discuss 5:1–28. doi:10.5194/essdd-5-1-2012

Dirren S, Hakim GJ (2005) Toward the assimilation of time-averaged observations. Geophys Res Lett 32:L04 804. doi:10.1029/2004GL021444

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, Crespin E, de Montety A, Mann ME, Renssen H, Timmermann A (2010) Reconstructing surface temperature changes over the past 600 years using climate model simulations with data assimilation. J Geophys Res Atmos 115:D09 108. doi:10.1029/2009jd012737

Goosse H, Crespin E, Dubinkina S, Loutre MF, Mann ME, Renssen H, Sallaz-Damaz Y, Shindell D (2012) The role of forcing and internal dynamics in explaining the "Medieval Climate Anomaly". Clim Dyn 39:2847–2866

Hakim GJ, Annan J, Bronnimann S, Crucifix M, Edwards T, Goosse H, Paul A, van der Schrier G, Widmann M (2013) Overview of data assimilation methods. PAGES News 21:72–73

Huntley HS, Hakim GJ (2010) Assimilation of time-averaged observations in a quasi-geostrophic atmospheric jet model. Clim Dyn 35:995–1009

Jansen E, Overpeck J, Briffa K, Duplessy J-C, Joos F, Masson-Delmotte V, Olago D, Otto-Bliesner B, Peltier W, Rahmstorf S, Ramesh R, Raynaud D, Rind D, Solomina O, Villalba R, Zhang D (2007) Palaeoclimate. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt K, Tignor M, Miller H (eds) Climate change 2007: the physical science basis. Contribution of working group 1 to the fourth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge, pp 433–497

Jones PD, Mann ME (2004) Climate over past millennia. Rev Geophys 42:RG2002. doi:10.1029/2003RG000143

Jones PD, Jonsson T, Wheeler D (1997) Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibraltar and south-west Iceland. Int J Climatol 17:1433–1450

Jungclaus JH, Lohmann K, Zanchettin D (2014) Enhanced 20th-century heat transfer to the Arctic simulated in the context of climate variations over the last millennium. Clim Past 10:2201–2213. doi:10.5194/cp-10-2201-2014

Klein F, Goosse H, Mairesse A, de Vernal A (2014) Model-data comparison and data assimilation of mid-Holocene Arctic sea ice concentration. Clim Past 10:1145–1163. doi:10.5194/cp-10-1145-2014

Luterbacher J, Xoplaki E, Dietrich D, Jones PD, Davies TD, Portis D, Gonzalez-Rouco JF, von Storch H, Gyalistras D, Casty C, Wanner H (2002) Extending North Atlantic Oscillation reconstructions back to 1500. Atmos Sci Lett 2:114–124. doi:10.1006/asle.2001.0044

Luterbacher J, Dietrich D, Xoplaki E, Grosjean M, Wanner H (2004) European seasonal and annual temperature variability, trends, and extremes since 1500. Science 303:1499–1503. doi:10.1126/science.1093877

Mairesse A, Goosse H, Mathiot P, Wanner H, Dubinkina S (2013) Investigating the consistency between proxy-based reconstructions and climate models using data assimilation: a mid-Holocene case study. Clim Past 9:2741–2757. doi:10.5194/cp-9-2741-2013

Mann ME, Zhang ZH, Rutherford S, Bradley RS, Hughes MK, Shindell D, Ammann C, Faluvegi G, Ni FB (2009) Global signatures and dynamical origins of the little ice age and medieval climate anomaly. Science 326:1256–1260, iSI Document Delivery No.: 524BD

Marsland SJ, Haak H, Jungclaus JH, Latif M, Roske F (2003) The Max-Planck-Institute global ocean/sea ice model with orthogonal curvilinear coordinates. Ocean Model 5:91–127. doi:10.1016/s1463-5003(02)00015-x

Matsikaris A, Widmann M, Jungclaus J (2015) On-line and off-line data assimilation in palaeoclimatology: a case study. Clim Past 11:81–93. doi:10.5194/cp-11-81-2015

Moberg A, Sonechkin DM, Holmgren K, Datsenko NM, Karlen W, Lauritzen SE (2005) Highly variable Northern Hemisphere temperatures reconstructed from low- and high-resolution proxy data (vol 433, pg 613, 2005). Nature 439:1014. doi:10.1038/nature04575

Oppenheimer C (2003) Climatic, environmental and human consequences of the largest known historic eruption: Tambora volcano (Indonesia) 1815. Prog Phys Geogr 27:230–259. doi:10.1191/0309133303pp379ra

PAGES 2K Consortium (2013) Continental-scale temperature variability during the past two millennia. Nat Geosci 6:339–346

Pendergrass AG, Hakim GJ, Battisti DS, Roe G (2012) Coupled air-mixed layer temperature predictability for climate reconstruction. J Clim 25:459–472

Pongratz J, Reick C, Raddatz T, Claussen M (2008) A reconstruction of global agricultural areas and land cover for the last millennium. Glob Biogeochem Cycles 22:Gb3018. doi:10.1029/2007gb003153

Raddatz TJ, Reick CH, Knorr W, Kattge J, Roeckner E, Schnur R, Schnitzler KG, Wetzel P, Jungclaus J (2007) Will the tropical land biosphere dominate the climate-carbon cycle feedback during the twenty-first century? Clim Dyn 29:565–574. doi:10.1007/s00382-007-0247-8

Robock A (2000) Volcanic eruptions and climate. Rev Geophys 38:191–219. doi:10.1029/1998rg000054

Rohde R, Muller R, Jacobsen R, Muller E, Perlmutter S, Rosenfeld A, Wurtele J, Groom D, Wickham C (2012) A new estimate of the average earth surface land temperature spanning 1753 to 2011. Geoinform Geostat Overv 1:1. doi:10.4172/2327-4581.1000101

Schmidt GA, Jungclaus JH, Ammann CM, Bard E, Braconnot P, Crowley TJ, Delaygue G, Joos F, Krivova NA, Muscheler R, Otto-Bliesner BL, Pongratz J, Shindell DT, Solanki SK, Steinhilber F, Vieira LEA (2011) Climate forcing reconstructions for use in PMIP simulations of the last millennium (v1.0). Geosci Model Dev 4:33–45. doi:10.5194/gmd-4-33-2011

Steiger NJ, Hakim GJ, Steig EJ, Battisti DS, Roe GH (2014) Assimilation of time-averaged pseudoproxies for climate reconstruction. J Clim 27:426–441. doi:10.1175/jcli-d-12-00693.1

Stevens B, Giorgetta M, Esch M, Mauritsen T, Crueger T, Rast S, Salzmann M, Schmidt H, Bader J, Block K, Brokopf R, Fast I, Kinne S, Kornblueh L, Lohmann U, Pincus R, Reichler T, Roeckner E (2013) Atmospheric component of the Earth System Model: ECHAM6. J Adv Model Earth Syst 5:146–172. doi:10.1002/jame.20015

van der Schrier G, Barkmeijer J (2005) Bjerknes' hypothesis on the coldness during AD 1790–1820 revisited. Clim Dyn 25:537–553. doi:10.1007/s00382-005-0053-0

Vieira LEA, Solanki SK, Krivova NA, Usoskin I (2011) Evolution of the solar irradiance during the Holocene. Astron Astrophys 531:A6. doi:10.1051/0004-6361/201015843

Widmann M (2005) One-dimensional CCA and SVD, and their relationship to regression maps. J Clim 18:2785–2792. doi:10.1175/jcli3424.1

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

Zanchettin D, Timmreck C, Graf HF, Rubino A, Lorenz S, Lohmann K, Kruger K, Jungclaus J (2012) Bi-decadal variability excited in the coupled ocean–atmosphere system by strong tropical volcanic eruptions. Clim Dyn 39:419–444, iSI Document Delivery No.: 965CZ5

Zanchettin D, Timmreck C, Bothe O, Lorenz SJ, Hegerl G, Graf HF, Luterbacher J, Jungclaus JH (2013) Delayed winter warming: a robust decadal response to strong tropical volcanic eruptions? Geophys Res Lett 40:204–209. doi:10.1029/2012gl054403