Zorita, E.; Moberg, A.; Leijonhufvud, L.; Wilson, R.; Brazdil, R.;
Dobrovolny, P.; Luterbacher, J.; Boehm, R.; Pfister, C.; Riemann, D.;
Glaser, R.; Soederberg, J.; Gonzalez-Rouco, F.:

European temperature records of the past five centuries based on documentary/instrumental information compared to climate simulations

In: Climatic Change (2010) Springer

DOI: 10.1007/s10584-010-9824-7

Climatic Change manuscript No.
(will be inserted by the editor)

European temperature records of the past five centuries based on documentary information compared to climate simulations

Eduardo Zorita · Anders Moberg ·
Lotta Leijonhufvud · Rob Wilson ·
Rudolf Brázdil · Petr Dobrovolný · Jürg
Luterbacher · Reinhard Böhm · Christian
Pfister · Rüdiger Glaser · Johan Söderberg ·
Fidel González-Rouco

Received: date / Accepted: date

This study is framed by the EU project Millennium Contract no: 017008 (GOCE)

Eduardo Zorita
GKSS Research Centre, Geesthacht, Germany
E-mail: eduardo.zorita@gkss.de

Anders Moberg and Lotta Leijonhufvud
Department of Physical Geography and Quaternary Geology, Stockholm University, Sweden
E-mail: anders.moberg@natgeo.su.se;Lotta.Leijonhufvud@natgeo.su.se

Rob Wilson
School of Geography & Geosciences, University of St Andrews, Scotland
E-mail: rjsw@st-andrews.ac.uk

Rudolf Brázdil and Petr Dobrovolný
Institute of Geography, Masaryk University, Brno, Czech Republic
E-mail: brazdil@sci.muni.cz; dobro@sci.muni.cz

Jürg Luterbacher
Oeschger Centre for Climate Change Research and Institute of Geography Climatology and Meteorology
University of Bern, Switzerland
E-mail: juerg@giub.unibe.ch

Reinhard Böhm
Zentralanstalt für Meteorologie und Geodynamik , Vienna, Austria
E-mail: reinhard.boehm@zamg.ac.at

Christian Pfister
Oeschger Centre for Climate Change Research and Institute of History, University of Bern, Switzerland
E-mail: pfister@hist.unibe.ch

Rüdiger Glaser
Institute of Physical Geography, University of Freiburg, Germany
E-mail: ruediger.glaser@geographie.uni-freiburg.de

Johan Söderberg
Department of Economic History,Stockholm University,Sweden
E-mail: Johan.Soderberg@ekohist.su.se

Fidel González-Rouco

Abstract Two European temperature records for the past half-millennium, January-to-April air temperature for Stockholm (Sweden) and seasonal temperature for a Central European region, both derived from the analysis of documentary sources combined with long instrumental records, are compared with the output of forced (solar, volcanic, greenhouse gases) climate simulations with the model ECHO-G. The analysis is complemented with the long (early)-instrumental record of Central England Temperature (CET). Both approaches to study past climates (simulations and reconstructions) are burdened with uncertainties. The main objective of this comparative analysis is to identify robust features and weaknesses that may help to improve models and reconstruction methods.

The results indicate a general agreement between simulations and the reconstructed Stockholm and CET records regarding the long-term temperature trend over the recent centuries, suggesting a reasonable choice of the amplitude of the solar forcing in the simulations and sensitivity of the model to the external forcing. However, the Stockholm reconstruction and the CET record also show a long and clear multi-decadal warm episode peaking around 1730, which is absent in the simulations. The uncertainties associated with the reconstruction method or with the simulated internal climate variability cannot easily explain this difference. Regarding the interannual variability, the Stockholm series displays in some periods higher amplitudes than the simulations but these differences are within the statistical uncertainty and further decrease if output from a regional model driven by the global model is used.

The long-term trends in the simulations and reconstructions of the Central European temperature agree less well. The reconstructed temperature displays, for all seasons, a smaller difference between the present climate and past centuries than the simulations. Possible reasons for these differences may be related to a limitation of the traditional technique for converting documentary evidence to temperature values to capture long-term climate changes, because the documents often reflect temperatures relative to the contemporary authors' own perception of what constituted 'normal' conditions. By contrast, the simulated and reconstructed inter-annual variability is in rather good agreement.

Keywords :

past millennium, temperature reconstructions, documentary data, early instrumental temperature, climate simulations

1 Introduction

Global climate models are the most important tools to estimate future climate change caused by changes in the radiation balance of the Earth linked to increasing concentrations of greenhouse gases. There has been considerable progress in the understanding of the relevant long-timescale processes that give rise to low-frequency climate variability and of the mechanisms relevant for the modulation of the sensitivity of Earth's climate to external forcing. However, in spite of this progress, global climate models have some difficulties in simulating the present climate, specifically at regional scales, and uncertainties still remain concerning their ability to simulate climates different from

---

Universidad Complutense de Madrid,Spain
E-mail: fidelgr@fis.ucm.es

the present one (Bader et al., 2008). One test for climate models is to compare climate simulations, where the models are driven by external forcings and/or boundary conditions representing past periods, to climate reconstructions for the same periods based on natural proxies or on documentary records. Most previous comparative studies have been devoted to the Holocene Optimum, about 6,000 years BP (e.g. Wohlfahrt et al., 2008; Wanner et al., 2008), or to the last glacial maximum, about 21,000 years BP (e.g. Ramstein et al., 2007). Such comparisons are certainly useful for improving the understanding of climate dynamics under conditions that may be substantially different from those of today, in particular for the Last Glacial Maximum. They are also of importance for an assessment of the simulated climate sensitivity to changes in external forcings. However, the centennial-scale climate variability and climate sensitivity in those periods could have been different from the present.

In this paper we aim at comparing climate simulations and local-to-regional climate reconstructions for the last few centuries, i.e. in a period when the external forcing and climate itself was much closer to the present state. Over this period the number of comparative studies between simulations and reconstructions is rather limited (Jones and Mann, 2004; Jones et al., 2008), but nevertheless several studies have focused on regional scales (e.g. Casty et al., 2005; Goosse et al., 2005, 2006; Raible et al., 2006; Wilson et al., 2006; Luterbacher et al., 2006, 2008; Gouirand et al., 2007). Here, the comparison will not only focus on assessing the skill of the models in simulating the long-term trends, which are to a large extent directly related to the long-term variation in the external forcing in these centuries, but will also examine the amplitude of multi-decadal variations around the long-term trends. The amplitude of these variations, which are more related to internal variability, can indicate how strongly regional climates can deviate - and for how long - from a long-term global warming or cooling trends. These variations thus provide some information of how regional climates may potentially deviate over the coming decades from the anthropogenically-forced global warming trend. We also compare the amplitude of interannual temperature variations, which are even more strongly dominated by internal variability. Additionally, model simulations may also be useful for evaluation of proxy-based climate reconstructions, as these may suffer from biases in of low- as well as high-frequency variations which may be illuminated when a set of climate reconstructions from nearby regions are compared with simulated data for corresponding areas.

In comparing the output of global simulations with reconstructions, one should not expect complete agreement in an interannual basis, not even in the unlikely case that the global climate model, the external forcing with which it was driven, and the climate reconstructions were perfect (Yoshimori et al., 2005). The evolution of the unforced variability - the variability not related to the external forcing but caused by the internal non-linear climate processes - in the simulations and in the observed past climate will not necessarily agree in spatio-temporal details. For instance, since inter-annual climate variations are essentially due to the internal variability, a very low correlation should be expected between climate simulations and reconstructions at inter-annual timescales, with the possible exception of sudden temperature variations caused by short-lived intense volcanic eruptions (Fischer et al., 2007). The timescale at which we should expect agreement between climate simulations and reconstructions depends on the spatial scale and on the climate variable considered. This timescale is dictated primarily by the characteristic time required for the random internal variations to be filtered out, and thus on the turbulent character of the variable under consideration. This characteristic time is shorter for variables that are more directly driven by the external radiative forcing, such as for instance temperature, and it should be larger for variables that are more strongly affected by the turbulent atmospheric dynamics, such as precipitation, wind or sea-level-pressure (Bader et al., 2008).

The time scale and spatial scale problem is compounded by the limited spatial resolution of global climate models that hinders a faithful representation of the relevant local processes, such as topography, coast lines, land-sea, land-use, etc. A potential solution is the use of regional climate models driven at their boundaries by the output of global simulations. A comparison between local or regional climate reconstructions with regional climate simulations would be more meaningful than with the output of global models, but unfortunately, the computer requirements of regional models are also large and century-long regional simulations are just beginning to be performed and very little model output to compare with exist.

Being aware of all the potential problems when comparing local climate reconstruction with the output of coarse-resolution global models, herein we undertake an initial joint assessment of model simulations with three long-term records derived from historical documentary evidence and long instrumental temperature observations. These records should represent different aspects of European climate in the past centuries. The millennial simulations were performed with the global climate model ECHO-G (see e.g. Zorita et al., 2005). One of the temperature reconstructions is additionally compared with output from the regional climate model RCA3 (Kjellström et al., 2005) designed for the Scandinavian region. This regional model was driven by one of the global simulations with ECHO-G, and should represent more realistically the coastlines, topography, and ocean heat capacity of the actual region (Moberg et al., 2006). However, the regional ocean sub-model used in the regional simulation is very simplified and does not represent heat advection by Baltic Sea currents. Despite those limitations in the global and regional model comparisons, we expect to find robust features among all time series, as well as identifying sources of disagreement that could indicate further ways for improvement of reconstruction and modeling approaches.

The main reason why we restrict this study to a few European sites is that this work is part of a European project (MILLENNIUM) specifically aimed at studying climate in Europe over the last millennium (Gagen et al., 2006). Moreover, the fact that this paper is a part of a special issue devoted to documentary climate evidence, motivates why we restrict our model-proxy data comparison to only documentary data and long instrumental records. Documentary evidence has previously been used for developing a number of climate reconstructions for different parts of Europe (for more details see e.g. Brázdil et al., 2005), but has also been used together with natural proxy data and long instrumental records to derive climate reconstructions at the European continental scale (e.g. Luterbacher et al., 2004; Xoplaki et al., 2005; Pauling et al., 2006).

Traditionally, the first step in deriving a documentary based temperature reconstruction involves the compilation of a time series of temperature indices onto an ordinal index scale (Brázdil et al., 2005). Such an index series can then be calibrated in terms of past temperatures based on statistical analysis of the temperature indices and observed temperatures in an overlapping period (Dobrovolný et al., 2009). In addition, some types of documentary data (e.g. phenological data, starting dates of shipping seasons, etc.) directly reflect environmental conditions that are closely related to temperature. This kind of documentary data (Chuine et al., 2004; Rutishauser et al., 2008 Leijonhufvud et al., 2009) allow a direct statistical calibration to temperatures using approaches similar to the standard reconstructions methods applied to other natural proxies, such as tree ring data . In our study, one temperature reconstruction of each kind is included - based on temperature indices for Central Europe and the based on continuous data about shipping time for Stockholm - which allows a comparison of the two approaches. The third temperature series used here, the Central England temperature series (Manley, 1974), is comprised almost exclusively of instrumental observations, but also includes information from documentary sources in its earliest part (before 1720). Details of the reconstructed/instrumental temperature series and the models are given in the following section.

2 Simulated and reconstructed temperatures

The global climate model used here, ECHO-G, is an atmosphere-ocean coupled model that has been widely used, for instance, for estimations of future climate change by the Intergovernmental Panel on Climate Change (Min et al., 2005; Meehl et al., 2007). The atmosphere is represented by the atmosphere model ECHAM4, with a horizontal resolution of 3.75x3.75 degrees and 19 levels covering the troposphere and the lower stratosphere. The ocean model is HOPE-G, with a horizontal resolution of about 2.8x2.8 degrees, with an increasingly finer resolutions towards the equator in the tropical regions. The ocean model contains 20 levels irregularly spaced at various ocean depths, with higher vertical resolution in the upper ocean layers. A flux correction is applied to the coupling of the atmosphere and ocean models to avoid climate drift. This flux correction is constant in time and has zero global average (Legutke and Voss, 1999).

The model ECHO-G was driven by estimations of past external forcing in the past millennium taken from existing literature as described in Zorita et al. (2005). The solar irradiance was derived from proxy-based estimations of solar radiative forcing taking into account the geometry of the Earth and its albedo. The changes in the solar irradiance between the 1960-1990 mean and the mean in the Late Maunder Minimum (1680-1710) are, in these simulations, -0.3%. The short-wave volcanic forcing was derived from estimations based on acidity of ice layers in Greenland and Antarctica ice cores. This forcing was translated to an effective change of the solar constant in the model. This change takes place in the model along the whole year when an eruption occurs and is, for all eruptions, maximum in northern summer. Estimates of atmospheric concentrations of carbon dioxide and methane were obtained from measurements of air bubbles in ice cores. More details about the simulations of the past millennium with ECHO-G are given by Zorita et al. (2005) and in González-Rouco et al. (2006). A discussion of the uncertainties in the external forcings and the difficulties of comparing the ECHO-G data with local proxy data is given by Gouirand et al. (2007).

Two simulations (denoted here ERIK1 and ERIK2) were carried out with ECHO-G and with the same external forcing over the past millennium, with only the initial conditions being different in year 1000. Although two simulations are clearly not enough to assess the full range of internal model variability, the computer costs of performing a large ensemble of simulations over the whole millennium is very large. Two simulations should just deliver a rough idea about the magnitude of internal model variability. The initial conditions used in 1000 in the ERIK2 simulation were colder than in ERIK1. The simulated global and northern hemisphere mean temperature in both simulations show higher temperatures at the beginning and at the end of the millennium with colder temperatures in the central centuries, roughly between 1400 and 1800 (Zorita et al., 2005). From roughly the year 1300, both simulations show similar multi-centennial variations. Temperature output from the ERIK1 and ERIK2 simulations for the period 1500-1990 are used for comparison with the reconstructed/instrumental temperature series.

For the case of Stockholm, we also include in the analysis data from one of the first multi-centennial climate simulations performed with a regional climate model. Such data are available for the Scandinavian region for the period 1550-1929 . The regional model is the atmospheric RCA3 (Kjellström et al., 2005). This model originates in the numerical weather prediction model HIRLAM and was used in the present study with a horizontal resolution of $1^o x 1^o$ and 24 levels i the atmosphere. The regional model was driven at the boundaries of the domain by the global model ECHO-G (the ERIK1 simulation), and forced by the same external forcings as ECHO-G, with the exception of CH_4. The influence of this forcing on temperature is, however, communicated indirectly to the regional model by the lateral boundary conditions. To represent the Baltic sea and the lakes in the domain of simulation, the model RCA3 was coupled to the lake model FLAKE (http://nwpi.krc.karelia.ru/flake/). FLAKE takes as input radiative and heat fluxes from the atmosphere as well as solid precipitation. Calculations of sea (lake) surface temperatures and ice/snow conditions are calculated locally for each grid box. The Baltic Sea flow dynamics are, however, not represented in this model. A fully three-dimensional model of the Baltic Sea flow would have substantially increased computer time requirements, and this was considered too demanding (Moberg et al., 2006). In case of both the global and the regional model, the temperatures simulated in the respective single grid-cell nearest to Stockholm was considered for this study.

The Stockholm temperature reconstruction is a combination of temperature evidence from documentary data and long instrumental data. The instrumental data are available from 1756 onwards and the documentary data from 1502 to 1892. The instrumental record is the homogenized series updated from Moberg et al. (2002). The documentary data consist of several sub-series of estimates of the start of the sailing season after each winter, which are normalized in an evolutive approach (to preserve low-frequency variations) and combined into a dimensionless average time series. The underlying assumption to use this series as a proxy for air temperature is that the Baltic Sea as far north as Stockholm, including the archipelago outside the city, usually is ice-covered during the winter months, and the start of the sailing season indicates when the water became ice-free after each winter. The start of the sailing season (i.e. the approximate date ice break-up) is found to be strongly correlated with the mean temperature from January to April. This holds very well up to the late nineteenth century, when steel-hulled ships became more common and ice became less of a problem for shipping. Because of the character of the proxy data and the way they were processed, the Stockholm temperature reconstruction is developed as a January-to-April mean temperature reconstruction, and is thus not available for any other parts of the year.

Following this rationale, in this study, two slightly different versions of the reconstructions are considered: in one, the reconstruction consists of the calibrated proxy data series in the period 1502-1892, after having adjusted its mean and variance to agree with the instrumental data over the full overlapping period (1756-1892), and instrumental data from 1893 onward. The variance scaling is applied to avoid loss of variance caused by a regression calibration (Leijonhufvud et al., 2009). The second version, covering the period 1500-1892, is the result of a calibration of the proxy series to instrumental data by ordinary least squares in the period 1756-1892. After 1892, the reconstruction is spliced to the instrumental record from Stockholm.

The second temperature reconstruction, for Central Europe (henceforth CeurT), is a combination of monthly temperatures reconstructed from documentary temperature index data for the period 1500-1854 and long instrumental temperature measurements after that date. Altogether 11 instrumental station records from Germany, Switzerland, Austria and the Czech Republic, have been used to derive the CeurT areal average. This instrumental average series starts in 1760. All station series were homogenized and ten of them originate from the HISTALP database (Auer et al., 2007). The homogenization includes, among other aspects, adjustment for urban warming trends and an adjustment for the early instrumental warm-bias problem (Frank et al., 2007; Böhm et al., 2009). The temperature index data were derived from historical documentary evidence from the same countries, except Austria. First, a monthly index series was derived for each country, and then an average was formed to obtain a regional series of dimensionless monthly temperature proxy data for the period 1500-1854. None of the three national series, however, is complete over the entire period, wherefore the number of contributing series vary over time. Therefore, a variance adjustment was applied to minimize the problem with artificial changes in variance due to the changing number of individual series. Next, the variance adjusted monthly series were calibrated (1771-1816) and verified (1760-1770 and 1817-1854) against the homogenized instrumental temperatures (for details see Dobrovolný et al., 2009. Here, we use two versions of the CeurT reconstruction. In both cases, the proxy data are used for the period 1500-1854 and instrumental data from 1855 onwards. The only difference is that in one version, the proxy data are calibrated by means of linear regression (1771-1816) and in the other case the same data are adjusted to have the same mean and variance as the instrumental data over 1760-1854. By using both variants (in the case of both CeurT and Stockholm) we are able to compare the result of regression versus variance scaling for the calibration with the simulated temperatures.

The third temperature record is the well known Central England monthly temperature record (CET), starting in 1659. This series is exclusively based on early instrumental and instrumental observations from a set of stations in the Central England region back to 1723, but before this year it is based on a combination of various short early instrumental observations for sites more close to London, temperature observations from De Bilt in the Netherlands (1707-1722) and various historical documentary evidence (Manley 1953, 1974, Parker et al., 1992, Parker and Horton 2005). Manley (1974) clearly pointed out that data from before around 1720 must be considered as less reliable. As discussed by Jones (1999), Manley provided only very sketchy information about how the temperature estimates before 1723 were derived. It appears that he used the available information together with a good understanding of the meteorological and climatological conditions of the area, to derive an 'educated best guess' temperature estimate for each month. For the first decade, these estimates are given only with a precision of whole degrees. Hence, the very earliest part of the CET series has some similarities to an ordinal temperature index scale being calibrated to temperatures. The monthly temperatures in the following five decades are mostly given at a precision of 0.5 degrees, which is closer to the later standard precision of 0.1 degrees but nevertheless less precise. Despite this uncertainty in the early part of the CET record, comparison with temperature estimates derived from many boreholes on the British Isles show a remarkably good agreement of the long-term evolution of annual mean temperatures over the last three-and-a-half centuries (Jones, 1999).

3 Stockholm winter-spring temperature

Figure 1 shows the evolution of the reconstructed (two versions) and simulated (three simulations) January-to-April temperatures in Stockholm over the past 500 years. We consider deviations from the 1829-1929 mean (the last 101 years common to all simulated series). The data are shown as 31-year running means. The model temperatures correspond to the grid-cell that includes the geographical coordinates of Stockholm.

It is well known that the realism of climate simulations may be weak at local scales, because, for example, of the poor representation of coastlines, land-sea interactions and topographical features. This is quite clear in the case of the global model, where the land connection between the Baltic Sea and the North Sea is two grid-cells wide, about 700 km. This may cause an unrealistically large advection of heat from the Atlantic Ocean into the Baltic Sea region and bias the simulation of mean temperatures along the Baltic Sea coast. Therefore only the temperature deviations from a reference state are considered. For the regional model, the land/sea mask is more realistic, but still too coarse to resolve local features such as lakes and the archipelago near Stockholm.

Both reconstruction versions are very close after the 31-year mean smoothing. There are some agreements and also clear differences between the simulations and the reconstructions. Reconstructed temperatures start with a declining trend in the early part of the record, reaching a minimum between 1550-1650, followed by a prolonged long-term warming until present, punctuated by a roughly 80-year long warming episode during 1680-1760, peaking in the 1730s. The initial cooling, the minimum temperature around 1600, and the subsequent long-term warming up to the present (here neglecting the 1730s warmth) is broadly reproduced in both simulations. The latter point is interesting, since the present uncertainties in the magnitude of the changes of the solar forcing in the last centuries (see e.g. discussion in Gouirand et al., 2007) could easily have caused a mismatch between the simulated and observed long-term trends in the past few centuries.

A large discrepancy, however, is evident between simulations and reconstructions during around 1680-1760. The warming episode lasting for about 80 years in the reconstruction is absent in both simulations. This mismatch cannot be reconciled by invoking the estimated uncertainties in the reconstructed temperature or by internal model variations. Figure 1 illustrates this by showing the amplitude of the uncertainty in the reconstructions at multi-decadal timescales (the 31-year means). These uncertainties have been derived from the root-mean-square error (rmse) between the reconstructed record (regression-based version) and observed temperatures in a validation period (Leijonhufvud et al, 2009). The inter-annual errors in the validation period are serially uncorrelated, so that the approximate 2-sigma uncertainty in the 31-year running means is taken as 2x rmse divided by $\sqrt{31}$. The uncertainties associated with the internal variability in the model are derived from the amplitude of the simulated temperature variations in the same grid-cell in a control simulation with the global model along 1000 years , where the external forcing has been kept constant at the present values (Zorita et al., 2003). This amplitude should include all sources of variability that are not connected to the external forcing, e.g. internal variations of the North Atlantic Oscillation modulating the advection of maritime air towards Scandinavia, or variations of the sea-surface temperatures in the North Atlantic caused by internal ocean dynamics. The magnitude of the model uncertainty is taken as twice the standard deviation of the smoothed (31-year running mean) simulated Stockholm temperature in the control run.

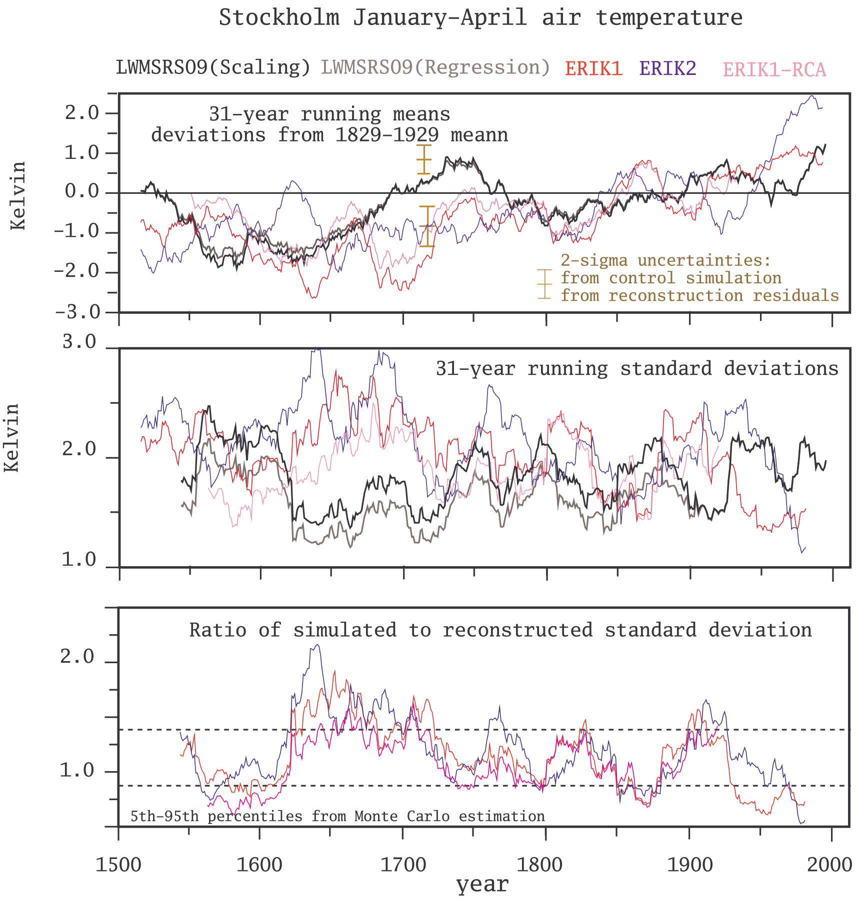

Fig. 1 January to April mean temperature in Stockholm reconstructed from documentary records (Leijonhufvud et al, 2009; LWMSR09) (two versions: variance-scaling and regression calibration) and simulated in two runs with the global climate model ECHO-G (ERIK1, ERIK2) and one run with the regional climate model RCA (ERIK1 - RCA). Upper panel: deviations from the 1829-1929 mean smoothed with a 31-year running mean; middle panel: 31-year running standard deviations; lower panel: ratio of the simulated standard deviations to the reconstructed (variance-scaling) standard deviation

The evolution of the temperature simulated by the regional model follows in general the temperature from the driving global simulation ERIK1 rather closely, although there are some differences between the global and regional model output. These differences are, however, clearly smaller than those between the output from the two global models.

It is unlikely that variations in the external forcing as prescribed in the model can explain the early 80-yr warming episode in the reconstruction, as this is absent in both global simulations. Also, the simulated internal climate variability seems unable to produce this reconstructed warming episode. This may be interpreted as an indication that the multi-decadal internal variability in the ECHO-G model is too small, perhaps due to too weak dynamics of the North Atlantic Ocean circulation. It may alternatively indicate that the reconstructed Stockholm temperatures are too warm around 1730. There is, however, independent support from the instrumental temperature measurements at the nearby city of Uppsala that the 1730s was an unusually mild period (Leijonhufvud et al., 2008; Moberg et al., 2005). Moreover, the Central England temperatures and also the De Bilt temperature in the Netherlands show a notable warmth in the same period (Jones and Briffa, 2006) (see also next section). The mismatch between proxy and model data for Stockholm during the decades around 1730 is therefore best explained as an incapability of the model to reproduce this warming feature (see also the conclusion section).

The Stockholm inter-annual temperature variability, as expressed by running 31-year standard deviations, show large inter-decadal changes. The inter-annual variability of the variance-scaled series is, by construction, larger than the regression-based reconstruction. The reconstruction standard deviation largely agrees with the simulations in some periods, but there are also periods when the simulated variability is larger in both simulations, in particular between 1620 and 1700 (Figure 1). The ratio of the modeled to reconstructed (variance-scaled) standard deviations is shown also in Figure 1. The significance of this ratio is assessed by constructing 1000 pairs of synthetic timeseries, both with standard deviation unity and the same autocorrelation as the simulated and reconstructed temperatures, respectively. Only in the decades around 1620-1700 the ratio of simulated to reconstructed variability is significantly (5% level) larger than it could be explained by chance.

The inter-annual variability in the regional model does, however, not significantly differ from the variance-scaled reconstruction in the 1620-1700 period. Rather, the regional model variability is essentially consistent with the Stockholm temperature reconstruction everywhere except before 1620, when the ratio between the simulated and reconstructed variability is above the 95^{th} percentile derived from the Monte Carlo simulations. This happens about 10% of the entire time period, one may conclude that the simulated regional inter-annual temperature variability is not significantly different from the reconstruction Notably, the variability in the regional model is almost consistently smaller than in the corresponding (ERIK1) global simulation, and also closer to the reconstruction than the global simulation. This indicates that the regional model provides a more realistic simulation of the local climate in this area.

All records display no clear trend in the inter-annual temperature variability along their corresponding period and the differences between the global model records are also large for much of the time. From the large differences between the simulated inter-annual variability in the two global simulations, it can be concluded that the external forcing can only account for a small amount of inter-annual variability, which thus must be dominated by internal variability.

4 Central European Temperature

In this section we compare the model simulations with ECHO-G with the reconstruction of the Central European air temperature based on documentary data (Dobrovolný et al., 2009). Although the reconstruction is available for all 12 calendar months, a more robust comparison is achieved on seasonal and annual basis. Figure 2 shows the annual and seasonal reconstructed Central European temperatures and the corresponding temperatures simulated by the model averaged over eight grid-cells that roughly

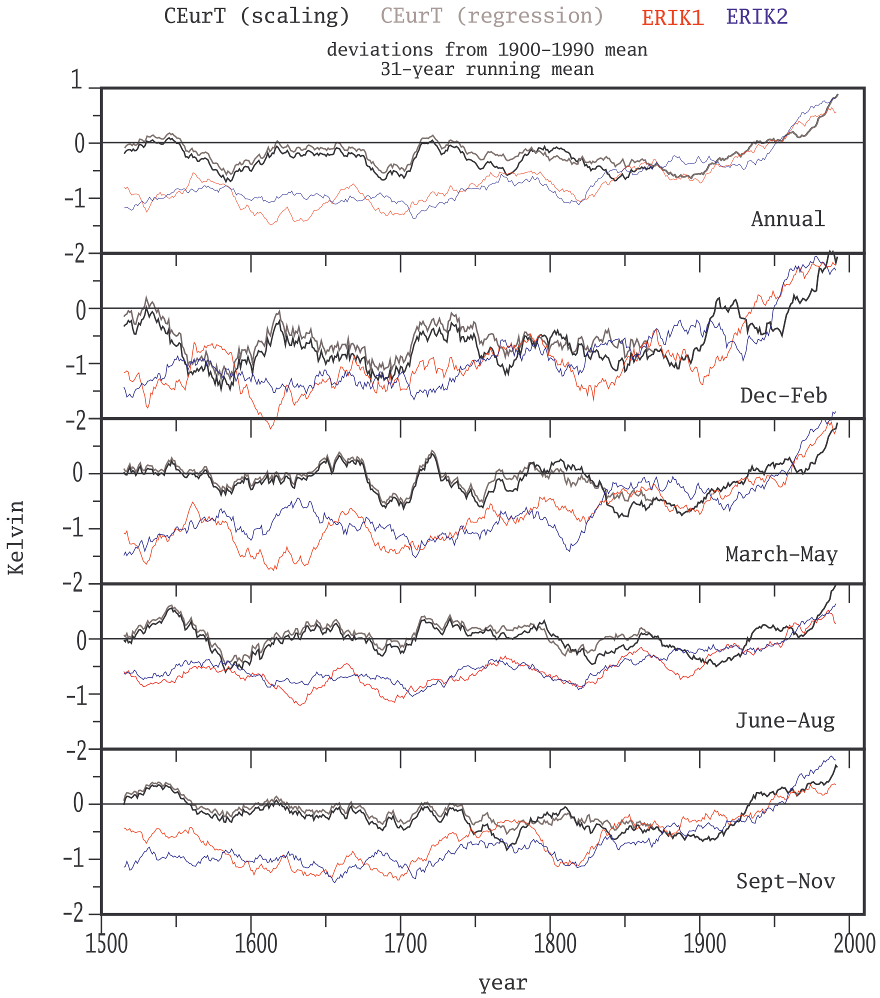

Fig. 2 Annual and seasonal means of the Central European Temperature reconstructions (Dobrovolný et al., 2009; CEurT scaling, CEurT regression) and of the corresponding grid-cells in two simulations with the global climate model ECHO-G (ERIK1, ERIK2). The series represent deviations from the 1900-1990 mean and are smoothed with a 31-year running mean.

correspond to the Central European area. Anomalies from the 1900-1990 mean are considered here. Because of the coarse resolution of the global model, the model topography in this area is quite flat (even the Alps are represented by elevations of just 600 meters). Therefore, the results are not critical to the exact choice of model grid-cells to represent the Central European area. The reconstruction is shown with the proxy data being calibrated with both regression and variance matching methos.

Overall, there are much larger differences between the simulated and reconstructed Central European temperatures than seen for Stockholm. There is good agreement between model and the reconstruction back to about 1850, but the reconstructions mostly show notably warmer temperatures than the model before 1850. This happens for all seasonal records and therefore for the annual means as well, although the differences between model and reconstructions are smaller for the winter season.

The seasonal reconstructions show multi-decadal variability with an amplitude similar to that displayed by the model simulations. The positions of the maxima and minima, however, do not tend to agree with the simulations. This can be attributed to internal regional variability, since the timing of the maxima and minima does not always agree in the two model simulations. A more striking difference between the reconstruction and the simulations is that the long-term warming seen over the last five centuries in all seasons in the simulations is much less pronounced in the reconstruction. There is even a slight cooling trend seen in the reconstruction over the first centuries after 1500, and this is not seen in the simulations.

The inter-annual variability in the reconstructions and simulations (31-yr running standard deviations; Figure 3) shows both similarities and differences depending on the time period and season. In winter and summer, the reconstructions show a large inter-annual variability at the beginning of the record, which diminishes along the half-millennium. In contrast the spring and autumn records do not show any discernible long-term trends in their inter-annual variability (Dobrovolný et al., 2009). The spring reconstruction nevertheless display somewhat smaller inter-annual variability in the twentieth century compared to the previous four centuries. The simulated records show a fairly constant inter-annual variability in all seasons, which stands in contrast to the instrumental data for winter and summer but agrees reasonably well in spring and autumn. Thus, at the beginning of the past half-millennium the reconstructions display larger inter-annual variability than the simulations during winter and especially so during summer. The reconstructed pre-instrumental variability could possibly be too high in these seasons (both in the regression and the variance matched version). If so, the reason for this is at present not known. Since the regression residuals for the summer and also annual reconstruction are significantly serially positively correlated according to a Durbin-Watson test (see Dobrovolný et al., 2009), the inter-annual variance of the reconstruction in summer could be incorrect. This may also affect the annual data. This may explain some of the differences in reconstructed and simulated inter-annual temperature variance in summer, but it does not explain the differences in winter. One possibility is that the real climate had on average more high-frequency temperature variability in summer and winter during the sixteenth to eighteenth centuries, perhaps caused by another circulation mode, and that this is correctly picked up by the reconstruction but not by the simulation. Previous analyses of reconstructed atmospheric circulation features over Europe (500 hPa geopotential height fields) and those simulated with ECHO-G over the last 350 years indicate quite substantial differences in some periods (Casty et al., 2005), which however could not be related to a response to the external forcing, either in the simulation or in the reconstructions (Casty et al., 2007). These differences could therefore, be caused by internal variability. In any case, it appears that further studies are needed to understand the differences between simulated and reconstructed temperature variability in the pre-instrumental period.

5 Central England Temperature

All proxy-based reconstructions of past temperatures may contain uncertainties of statistical nature, which can be partially quantified, but there are also likely sources of unknown error which have not been taken into account. To shed more light on

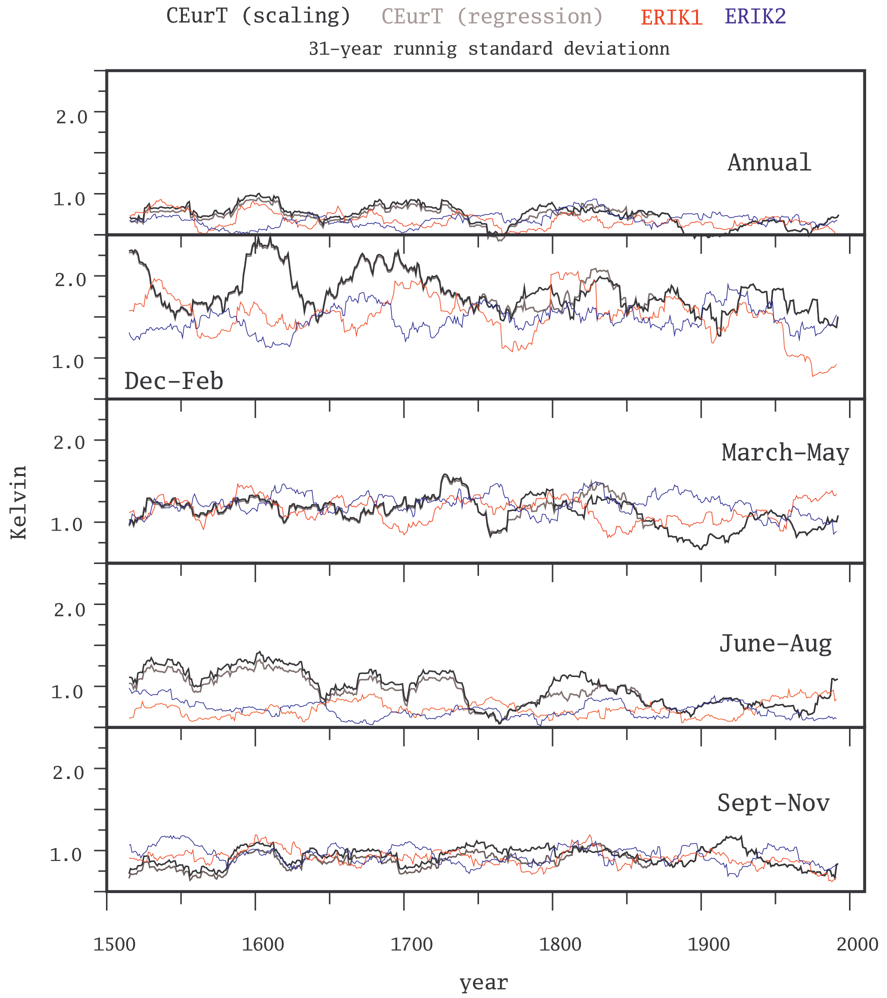

Fig. 3 Running inter-annual standard deviation of the annual and seasonal temperature derived from the reconstructed Central European temperature and from output of two simulations with the global climate model ECHO-G in the corresponding grid-cells.

whether or not the agreements and discrepancies found between the Stockholm and Central European temperature reconstructions and the model simulations are robust, we carry out a similar comparison between the simulation and the long instrumental Central England Temperature (CET) record, which starts in 1659. It should be borne in mind, though, that this 350-year long temperature record can hardly be considered homogeneous over the whole period (recall our brief discussion in section 2). The results of this comparison are conveyed by Figure 4 and 6. To represent the 'model CET' four model grid-cells located over Great Britain have been selected. Deviations from the 1900-1990 mean are considered.

Mean annual temperatures in the simulation agree quite well with CET in their long-term trend. Again, this supports the idea that the choice of external forcing com-

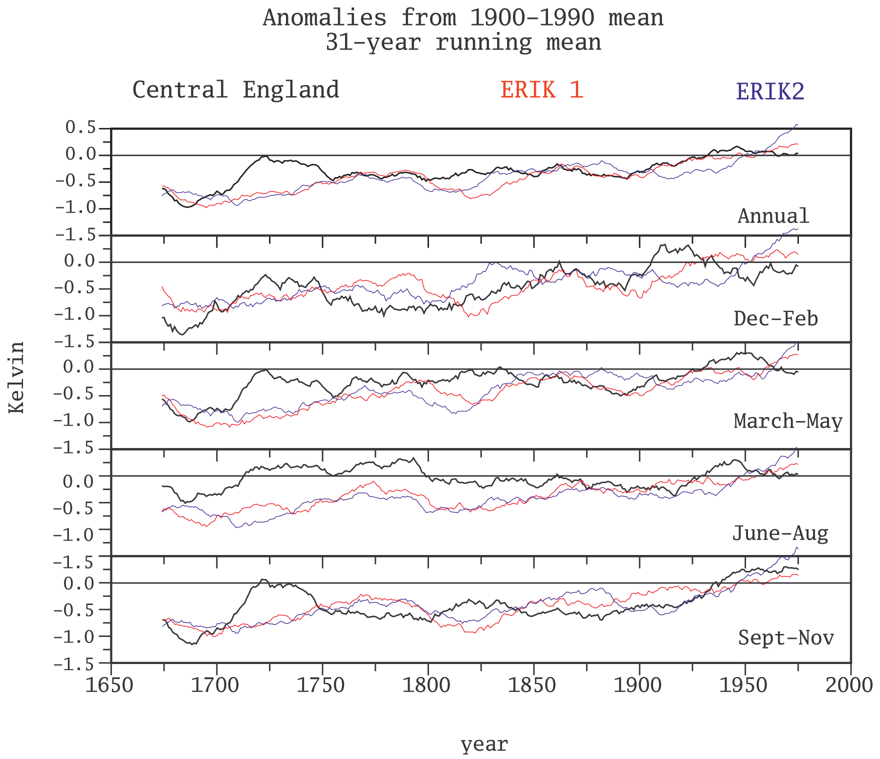

Fig. 4 Annual and seasonal means of the Central England Temperature and of the corresponding grid-cells in two simulations with the global climate model ECHO-G (ERIK1, ERIK2). The timeseries represent deviations from the 1900-1990 mean and are smoothed with a 31-year mean.

bined with the model climate sensitivity, yield a realistic picture of annual temperature evolution over the past recent centuries. However, a multi-decadal disagreement is noted around the 1710s to 1730s, similar to the one as noted earlier for the Stockholm January-April reconstruction. In CET, this mismatch is particularly clearly seen in the annual, March-May and September-November data but also in December-February data. The CET and Stockholm records have been derived independently, and therefore, one has to conclude that the relatively warm decades around the 1730s are very likely real phenomena (see also Luterbacher et al., 2004 and Xoplaki et al., 2005 (for European mean temperature); Jones and Briffa, 2006). This agreement between the CET and Stockholm series could indicate that its origin may lie in the North Atlantic and that it could have been a more widespread phenomenon. It is, however, not as easily seen in the CEurT series, although the reconstructed winters for that region during the 1710s to 1740s display a rather 'gently' warm level.

The long-term evolution of the CET summer temperatures over the whole period differs from the model simulations. The trend in the observed summer temperatures is very small, and summer temperatures in the 18th century seem to be higher than recent temperatures until 1990. This is a result which is not observed for the other seasons, where the observed and simulated trends are comparable. The practically constant

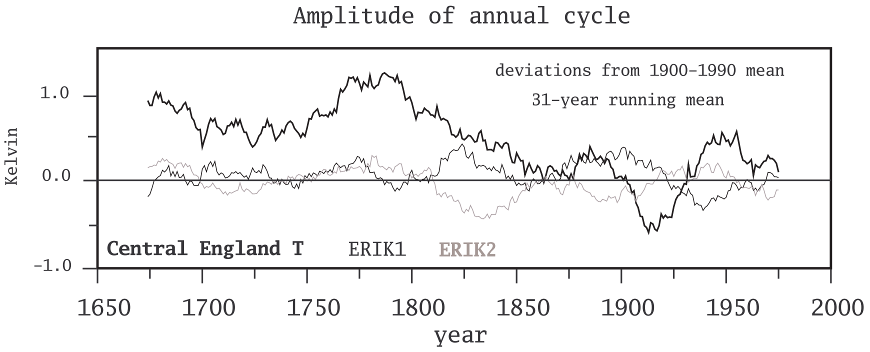

Fig. 5 Evolution of the amplitude of the annual cycle (June-August mean minus December-February mean) in the Central England temperature record and in two simulations with the climate model ECHO-G. The timeseries represent deviations from the 1900-1990 mean and are smoothed with a 31-year running mean.

observed summer temperatures together with the increasing winter temperatures along the past 350 years lead to a clear long-term decreasing amplitude of the observed annual cycle, which is absent in the simulations (Figure 5), and which has been reported also for other long instrumental temperature records in the Northern Hemisphere (Jones et al., 2003). However, the constantly warmer instrumental summer temperatures before the late 19th century in the CET data century suggests the presence of a warm bias in the early English observational data in this season, similarly to what has been concluded for Central European stations (Frank et al., 2007) and also proposed for Stockholm and Uppsala summer temperatures (Moberg et al., 2003). The argument for the presence of such a bias is that the thermometers were less well protected against radiation effects before the introduction of 'modern' screens, which typically occurred in the mid- to late-nineteenth century. This effect is largest in summer. This issue is discussed in more detail in (Böhm et al., 2009) for the Great Alpine Region. Böhm et al. (2009) estimate a correction for the summer temperature bias for this region in the range of 0.3-0.4 K (see their figure 13). A similar correction for the CET would bring the CET summer series closer to the model results, but still, simulations and CET would not fully agree in summer. Furthermore, a purported summer correction for England would probably be smaller than for Central European stations, as the effect of summer radiation was also presumably smaller. Nevertheless, an analysis of the possible presence of biased summer temperatures at the stations used to derive the CET record, in the period before the improvement of thermometer screens, seems worthwhile.

The inter-annual variability of the observed and simulated records (Figure 6) are notably different for the winter temperatures, when the observations systematically display a larger variability than the simulations. There is also an overall somewhat larger variability in the instrumental series than the simulated annual mean temperatures before around 1900, and particularly before 1800. None of the seasonal records of inter-annual variability, however, shows any clear trends, perhaps with the exception of a slight downward trend in the observed annual CET since 1800. Moreover, the inter-annual variability in the two simulations does not evolve in phase, possibly reflecting that the influence of the external forcing on the inter-annual variability is small.

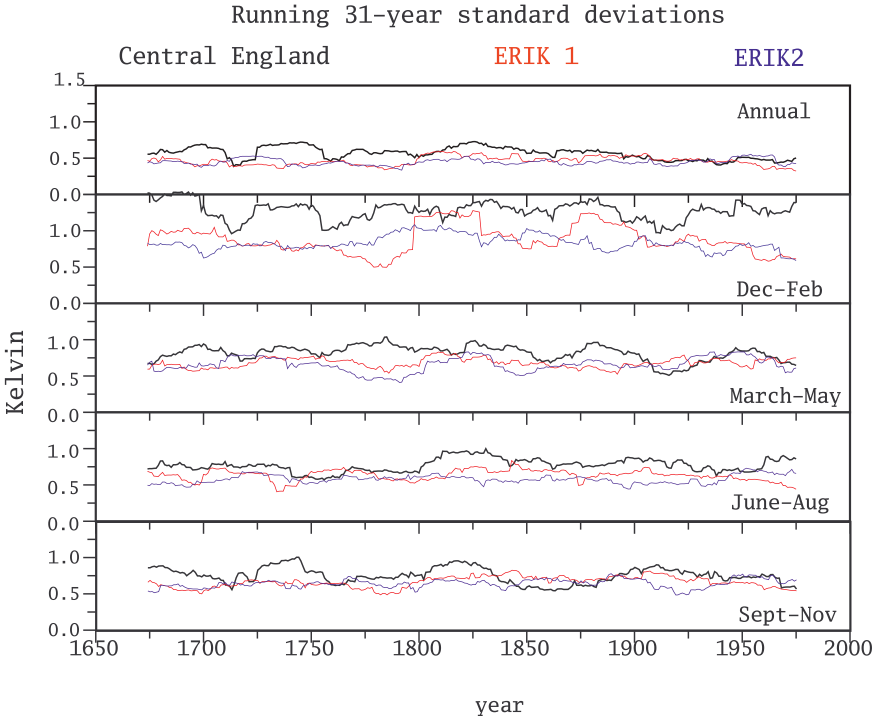

Fig. 6 Running inter-annual standard deviation of the annual and seasonal temperature derived from the Central England temperature records and from output of two simulations with the global climate model ECHO-G in the corresponding grid-cells.

6 Discussion and conclusions

The comparisons of reconstructed/instrumental and simulated temperatures for three European locations over the last 350 to 500 years display both encouraging agreements - especially considering the global model's coarse resolution - but also disagreements. We discuss in the following the main differences, which can help us to identify deficiencies in the models and reconstructions methods.

One important difference is the multi-decadal warm episode peaking around the 1730s that is displayed by the Stockholm winter-early spring temperature reconstruction and also by the Central England (CET) record. In CET, it is present not only in the winter season, but also in spring and autumn and hence is reflected also in the annual mean temperatures. The coarse resolution of the climate model is an unlikely cause of these differences compared to the reconstructions. The origin of the warm episode may more likely lie in the wintertime North Atlantic circulation, atmospheric or oceanic, since in summer, the season where the short-wave radiation forcing at the surface is most powerful, the deviation at this time are not very marked. A North Atlantic origin is supported by the fact that reconstruction of the Central European Temperatures does not clearly show the decades around 1730 as a notably mild period, but rather as a 'gently' mild one.

A tentative explanation for the mild decades around 1730 could be related to a prolonged positive phase of the North Atlantic Oscillation (Luterbacher et al., 1999), since some studies indicate that the atmospheric circulation is the most important factor for the relatively mild temperatures along the Western European coast (Seager et al., 2002). Increased solar irradiance at the end of the 17th century and through the first half of the 18th century might have induced a shift of the NAO toward a high NAO index. Similar findings were suggested for European winters (Luterbacher et al., 2004; Xoplaki et al. 2005). Thus, the solar irradiance changes might be a major trigger to explain the increasing trend in European winter and spring temperature at decadal time scale. However, reconstructions of the past NAO index do not completely agree (Schmutz et al., 2000). For instance, the NAO reconstruction by Cook et al. (2002) does not show such anomalous decades as the Luterbacher et al. (1999) does. From a modeling point of view, simulations for the Late Maunder Minimum (Shindell et al., 2003) indicate that increased solar irradiance at the end of the Maunder Minimum might have indeed induced a shift towards positive AO/NAO during November-April, leading to continental warming. However, it is known that climate models are not able to replicate the observed low-frequency variations of the atmospheric circulation in the North Atlantic in the 20th century (Osborn, 2004), and even future trends simulated by different models under much stronger external forcing according to future emissions scenarios do not agree in magnitude (Miller et al., 2007). It therefore could be possible that natural, multi-decadal excursions of the heat transport by the atmosphere are also not well replicated by models.

It has to be noted, however, that our conclusions about internal variability are based on a control simulation and only two forced simulations with one single model. A better estimation would require a much larger ensemble of forced simulations over the whole period. For instance, the difference in Stockholm temperature between ERIK1 and ERIK 2 around 1630 (Fig 1) is larger than our estimation of internal variability based on a control simulation, and indeed the NAO indices in ERIK1 and ERIK2 show the largest differences of the whole millennium around 1630 (Zorita et al., manuscript submitted).

Other clear differences are the long-term temperature trends of the reconstructed Central European temperature record, which are much smaller in magnitude than those simulated by the climate model. Although a model deficiency cannot be ruled out, the fact that the long-term trends of the Stockholm January-April temperature and the seasonal Central England Temperatures (except in summer) are much more in agreement with the model results, may indicate that the reconstructed Central European record underestimates the multi-centennial variations. In contrast to the Stockholm temperature record, which is derived from physically-anchored documentary information related to the ice break-up dates in the port of Stockholm, the Central European record is composed by more varied documentary records, some of which could suffer from a variant of the so called 'segment-length-curse' (Cook et al., 1995). As mentioned by Dobrovolný et al. (2008, 2009), the documentary sources contain descriptions of weather recorded according to the contemporary authors' own perceptions of what constituted 'normal' conditions. These perceptions were framed by the period in which the authors were living. Therefore, the index series derived from documentary evidence expresses primarily deviations from the respective normal conditions. Slow changes in these normal conditions can thus likely not be fully captured by the reconstructed indices. This drawback can possibly be partly removed if other proxies are used that are more directly anchored in the prevailing environmental conditions, such as series of phenophases, freezing of rivers and lakes, start of agricultural harvests, etc. Certainly, more carefully analyses and comparisons with other proxy-based records are necessary to ascertain whether or not this is relevant for this particular record.

The absence of a clear long-term summer temperature warming trend in the Central European Temperature reconstruction, which agrees with the Central England record in this respect, is in contrast with the summer temperature trends in the climate simulations, and also with the reconstructed summer temperature trends based on tree-ring data from the Alpine region (Büntgen et al., 2005; 2006). This mismatch can have significant ramifications in the discussion about the amplitude of past changes in solar irradiance but also about the uncertainties in the long-term instrumental records. One possible explanation for the disagreement between the simulated and reconstructed summer temperature trends is that the long-term variations in the external forcing may have been overestimated in the climate simulations. More recent estimations of the changes in solar irradiance between the Late Maunder Minimum and present are of the order of 0.17% (Wang et al., 2005), or even less than 0.1% (Foukal et al., 2004), instead of the 0.3% in the ECHO-G simulations. Such a mismatch would likely be more evident for the summer temperatures. On the other hand, such small amplitude changes in solar irradiance as less than 0.1% would hardly be compatible with the general agreement seen between the simulated temperatures in the other seasons and the CET or the Stockholm series. The climate sensitivity of the model ECHO-G (2.5 K caused by a doubling of atmospheric CO_2 concentrations) lies well within the range of the IPCC climate models (Meehl et al., 2007), so that an anomalously high sensitivity of the model to external perturbations cannot be invoked.

As regards the CET record, the notably warm temperatures in summer before 1800 compared to the twentieth century, which are also warmer than the simulated summer temperatures before the late nineteenth century, may perhaps indicate that the early observations suffer from a positive bias before the introduction of modern-type thermometer screens, as it has been found in early instrumental records from the Great Alpine Region (Frank et al., 2007; Böhm et al., 2009) and also suggested for Stockholm (Moberg et al., 2003). This points to the importance of further investigating the possible early warm instrumental temperature bias, not only in the CET, but potentially in all temperature records extending back to before the introduction of modern-type thermometer screens, i.e. typically before the mid-to late-19th centuries. The absence of a long-term warming trend in the reconstructed Central European summer temperatures analyzed here can, however, not be easily explained by any warm bias in the early instrumental observations as these were adjusted for this problem (Böhm et al., 2009).

In summary, the comparison between modeled and reconstructed records points to particular lines of research for all teams involved. Particularly important seems the inability of the climate model to simulate large multi-decadal regional temperature excursions in the past , as it could point to possible future regional deviations from the global warming trend. This should be more accurately ascertained with ensemble simulations. The physical origins of these excursions need also to be identified. In the end, the analysis should yield a consistent picture of model and observational results, thus further constraining the possible range of models and identifying possible biases in climate reconstruction methods.

Acknowledgements This study is part of the European project MILLENNIUM (contract no 017008-2). A. Moberg, L. Leijonhufvud and J. Söderberg acknowledge the Swedish Research Council. L- Leijonhufvud expresses her gratitude to FORMAS. F. González-Rouco was supported by the Spanish 'Ministerio de Educación y Ciencia' through grant REN-2002-04584-CLI. and J. Lüterbacher by the Swiss National Science Foundation (NCCR Climate).

References

1. Auer, I., Böhm, R,, Jurkovic, A., Lipa, W., Orlik, A., Potzmann, R., Schöner, W., Ungersböck, M., Matulla, C., Briffa, K., Jones, P.D., Efthymiadis, D., Brunetti, M., Nanni, T., Maugeri, M., Mercalli, L., Mestre, O., Moisselin, J.-M., Begert, M., Müller-Westermeier, G., Kveton, V., Bochnicek, O., Stastny, P., Lapin, M., Szalai, S., Szentimrey, T., Cegnar, T., Dolinar, M., Gajic-Capka, M., Zaninovic, K., Majstorovic, Z., Nieplova, E. : 2007, ' HISTALP - Historical instrumental climatological surface time series of the greater Alpine region 1760-2003', *Int. J. Climatol.* **27**, 17-46.
2. Böhm, R., Jones, P.D., Hiebl, J., Frank, D., Brunetti, M., Maugeri, M.: 2009, ' The early instrumental warm-bias: a solution for long central European temperature series 1760-2007', *Climatic Change*, this issue.
3. Bader, D.C, Covey, C., Gutowski, W.J., Held, I.M. , Kunkel, K.E., Miller, R.L., Tokmakian,R.T and Zhang, M.H.: 2008, ' Climate models: an assessment of strengths and limitations'. Synthesis and Assessment products 3.1 Report by the U.S. Climate change Science Program and the subcommittee on global change research, 133 pp, Washington DC.
4. Brázdil, R., Pfister, C., Wanner, H., von Storch, H., Luterbacher, J. :2005 ' Historical climatology in Europe - the state of the art.', *Climatic Change*, **70**, 363-430.
5. Büntgen, U., Frank, D.C., Nievergelt, D., Esper, J. : 2006, ' Summer temperature variations in the European Alps: AD 755-2004', *J. Clim.* **19**, 5606-5623 .
6. Casty, C., Handorf, D., Raible, C.C., González-Rouco, J.F., Weisheimer, A., Xoplaki, E., Luterbacher, J., Dethloff, K,, Wanner, H.:2005, 'Recurrent climate winter regimes in reconstructed and modelled 500 hPa geopotential height fields over the North Atlantic/European sector 1659-1990'. *Climate Dynamics* **24**, 809-822.
7. Casty, C., Raible C.C., Stocker, T.F., Wanner, H. and Luterbacher, J.: 2007 'A European pattern climatology 1766-2000', *Climate Dynamics* **29**, 791-805.
bibitemRefJ Chuine, I., Yiou, P., Viovy, N., Seguin, B., Daux, V. and Le Roy Ladurie, E.: 2004: 'Grape ripening as a past climate indicator', *Nature* **432**, 289-290.
8. Cook, E.R., Briffa, K.R., Meko, D.M., Graybill, D.A., Funkhouser, G.:1995 'The 'segment length curse' in long tree-ring chronology development for palaeoclimatic studies'. *The Holocene* **5**, 229-237.
9. Cook, E., D'Arrigo, R. and Mann, M.E.: 2002, 'A well verified multiproxy reconstruction of the winter North Atlantic Oscillation index since A.D. 1400'. *J. Clim.* **15**, 1754-1764.
10. Dobrovolný, P., Br?zdil, R., Val??ek, H., Kotyza, O., Macková, J., Halíčková, M.: 2008, ' A standard paleoclimatological approach to temperature reconstruction in historical climatology: an example from the Czech Republic, AD 1718-2007', *Int. J. Climatol.*, published online.
11. Dobrovolný, P., Moberg, A., Brázdil, R., Pfister, C., Glaser, R., Wilson, R.J.S., van Engelen, A., Limanowka, D., Kiss, A., Halíčková, M., Macková, J., Riemann, D., Luterbacher, J. and Böhm, R.:2009, 'Monthly and seasonal temperature reconstructions for Central Europe derived from documentary evidence and instrumental records since AD 1500'. *Climatic Change* (this issue).
12. Fischer, E.M., Luterbacher, J., Zorita, E., Tett, S.F.B. , Casty, C. and W. Wanner: 2007, 'European climate response to tropical volcanic eruptions over the past half millennium'. *Geophys. Res. Lett.* **34**, L05707, doi:10.1029/2006GL027992.
13. Foukal, P., North, G. and Wigley, T.: 2004, 'A stellar view on solar variations and climate'. *Science* **306**, 68-69.
14. Frank, D., Büntgen, U. , Böhm, R. , Maugeri, M. and Esper, J.: 2007, 'Warmer early instrumental measurements versus colder reconstructed temperatures: shooting at a moving target', *Quarter. Sci. Rev.* **26** 3298-3310.
15. Gagen M., McCarrol D., and Hicks, S.: 2006, 'The Millennium project: European climate of the last millennium'. *PAGES News* **14**,: 4-4.
16. Goosse, H., Renssen, H., Timmermann, A. and Bradley, R.S.: 2005 'Internal and forced variability during the last millennium: a model-data intercomparison using ensemble simulations'. *Quat. Sci. Rev* **24** 1345-1360.
17. Goosse, H., Arzel, O., Luterbacher, J., Mann, M.E., Renssen, H., Riedwyl, N., Timmermann, A., Xoplaki, E., and Wanner, H.: 2006, 'The origin of the European 'Medieval Warm Period'. *Clim. Past* **2**, 99-113.
18. Gouirand, I., Moberg, A. and Zorita, E.: 2007, 'Climate variability in Scandinavia for the past millennium simulated by an atmosphere-ocean general circulation model', *Tellus* **59A**, 30-49.

19. Jones, P.D.: 1999, 'Classics in physical geography revisited. Manley, G. 1974: Central England temperatures: monthly means 1659-1973' *Progr. Phys. Geogr.* **23**, 425-428.
20. Jones, P.D. and Briffa, K.R.: 2006, 'Unusual climate in northwestern Europe during the period 1730 to 1745 based on instrumental and documentary data'. *Climatic Change* **79**, 361-379.
21. Jones, P.D., Briffa, K.R. and Osborn, T.J.:2003, 'Changes in the Northern Hemisphere annual cycle: implications for paleoclimatology'. *J. Geophys. Res.* **108**, 4588, doi:10.1029/2003JD003695.
22. Jones, P.D. et al., 2008: 'High-resolution paleoclimatology of the last millennium: a review of current status and future prospects'. *The Holocene*, in press.
23. Jones, P.D. and Mann, M.E.: 2004, 'Climate over past millennia', *Rev. Geophys.* **42**, RG2002.
24. Kjellström, E., Bärring, L., Gollvik, S., Hansson, U., Jones, C., Samuelsson, P., Rummukainen, M., Ullerstig, A., Willén, U., Wyser, K.: 2005, 'A 140-year simulation of European climate with the new version of the Rossby Centre regional atmospheric climate model (RCA3)'. Reports Meteorology and Climatology 108, SMHI, SE-60176 Norrköping, Sweden.
25. Legutke, S. and Voss, R.:1999, 'The Hamburg atmosphere-ocean coupled model circulation model ECHO-G'. Technical Report No. 18. German Climate Computing Centre (DKRZ), Hamburg, 62 pp (1999).
26. Leijonhufvud, L., Wilson, R. and Moberg, A..: 2008, 'Documentary data provide evidence of Stockholm average winter to spring temperatures in the eighteenth and nineteenth centuries'. *Holocene* **18**, 333-343, doi: 10.1177/0959683608086770.
27. Leijonhufvud, L., Wilson, R., Moberg, A., Söderberg, J., Retsö, D., Söderlind, U.: 2009, 'Five centuries of winter/spring temperatures in Stockholm reconstructed from documentary evidence and instrumental observations', *Climatic Change*, this issue .
28. Luterbacher, J., and 48 coauthors: 2006, 'Mediterranean climate variability over the last centuries: A review, in: The Mediterranean Climate: an overview of the main characteristics and issues', Eds. Lionello, P., Malanotte-Rizzoli, P., and Boscolo, R., Elsevier, Amsterdam, the Netherlands, 27-148. .
29. Luterbacher, J., Dietrich, D., Xoplaki, E., Grosjean, M. and Wanner:, H. 2004, 'European seasonal and annual temperature variability, trends and extremes since 1500', *Science* **303**, 1499-1503.
30. Luterbacher, J., Schmutz, C., Gyalistras, D., Xoplaki, E. and Wanner, H.:1999, 'Reconstructions of monthly NAO and EU indices back to A.D. 1675' *Geophys. Res. Lett.* **26**, 2745-2748.
31. Luterbacher, J., Xoplaki, E., Küttel, M., Zorita, E., González-Rouco, F.J., Jones, P.D., Stössel, M., Rutishauser, T., Wanner, H., Wibig, J., and R. Przybylak, 2009: 'Climate change in Poland in the past centuries and its relationship to European climate: evidence from reconstructions and coupled climate models', in review.
32. Manley, G.: 1953, 'The mean temperature of Central England, 1698 to 1952', *Q.J.R. Meteorol. Soc.* **79**, 242-261.
33. Manley, G.:1974, 'Central England Temperatures: monthly means 1659 to 1973', *Q.J.R. Meteorol. Soc.* **100**, 389-405.
34. Meehl, G.A., Stocker, T.F. , Collins, W.D., Friedlingstein, P. , Gaye, A.T. , Gregory, J.M., Kitoh, A., Knutti, R., Murphy, J.M., Noda, A., Raper, S.C.B., Watterson, I.G. , Weaver, A.J. and Zhao, Z.-C. : 2007: 'Global Climate Projections. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change' [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
35. Meier, N., Rutishauser, T., Pfister, C., Wanner, H. and Luterbacher J.: 2007, 'Grape harvest dates as a proxy for Swiss April to August temperature reconstructions back to AD 1480', *Geophys. Res. Lett.* **34**, L20705, doi:10.1029/2007GL031381.
36. Miller, R.L., Schmidt, G.A. and Shindell, D.T.: 2006, 'Forced annular variations in the 20th century Intergovernmental Panel on Climate Change Fourth Assessment Report models'. *J. Geophys. Res.* **111**, D181101, doi:10.1029/2005JD006323.
37. Min, S.K., Legtuke, S. , Hense, A. and Kwon, W.-T.: 2005,'Internal variability in a 1000-year control simulation with the coupled climate model ECHO-G. I. Near surface temperature,precipitation, and mean sea-level pressure', *Tellus* **57A**, 605-621.
38. Moberg, A., Bergström, H., Ruiz Krigsman, J., Svanered, O.: 2002, 'Daily air temperature and pressure series for Stockholm (1756-1998)', *Climatic Change* **53**, 171-212 .

39. Moberg, A., Alexandersson, H., Bergström, H. and Jones, P.D. :2003: 'Were Southern Swedish summer temperatures before 1860 as warm as measured?', *Int. J. Climatol.* **23**, 1495-1521.
40. Moberg, A., Tuomenvirta, H. and Nordli, Ø.:2005, 'Recent climatic trends. In: Physical Geography of Fennoscandia'. (Ed: Seppälä M). Oxford Regional Environments Series, Oxford University Press, Oxford: 113-133.
41. Moberg, A., Gouirand, I., Schoning, K., Wohlfarth, B., Kjellström, E., Rummukainen, M., de Jong, R., Linderholm, H., Zorita, E..:2006, 'Climate in Sweden during the past millennium - Evidence from proxy data, instrumental data and model simulations'. SKB Technical Report TR-06-35, Swedish Nuclear Fuel and Waste Management Co., Stockholm. 87 pp.
42. Osborn, T.J.: 2004, 'Simulating the winter North Atlantic Oscillation: the roles of internal variability and greenhouse gas forcing', *Climate Dynamics* **22**, 605-623.
43. Parker, D.E. and Horton, E.B.: 2005, 'Uncertainties in Central England Temperature 1878-2003 and some improvements to the maximum and minimum series', *Int. J. Climatol,* **25**, 1173-1188.
44. Parker, D.E., Legg, T.P. and Folland, C.K.: 1992, 'A new daily Central England Temperature Series, 1772-1991', *Int. J. Climatol.* **12**, 317-342.
45. Pauling, A., Luterbacher, J., Casty, C.and Wanner, H.:2006, 'Five hundred years of gridded high-resolution precipitation reconstructions over Europe and the connection to large-scale circulation', *Climate Dynamics***26**, 387-405.
46. Raible, C.C., Casty, C., Luterbacher, J., Pauling, A., Esper, J., Frank, D.C., Büntgen, U., Roesch, A.C., Tschuck, P., Wild, M., Vidale, P.-L., Schär, C. and Wanner, H.:2006, 'Climate variability - observations, reconstructions and model simulations for the Atlantic-European and Alpine region from 1500-2100 AD', *Climatic Change* **79**, 9-29.
47. Rutishauser, T., Luterbacher, J., Defila, C., Frank, D., and Wanner, H., 2008: 'Swiss Spring Plant Phenology 2007: Extremes, a multi-century perspective and changes in temperature sensitivity', *Geophys. Res. Lett.* **35**, L05703. doi:10.1029/2007GL032545.
48. Schmutz, C., Luterbacher, J. , Gyalistras, ,D. Xoplaki, E. and Wanner, H.:2000, 'Can we trust proxy-based NAO index reconstructions?', *Geophys. Res. Lett.* **27**, 1135-1138.
49. Shindell, D. T., Schmidt, G.A., Mann M.E., Rind, D., and Wapple A., 2001: 'Solar forcing of regional climate change during the Maunder Minimum', *Science* **294**, 2149-2152.
50. Seager, R., Battisti, D. S. ,Yin, J. , Gordon, N., Naik, N. H., Clement, A. C. and Cane, M. A.: 2002, 'Is the Gulf Stream responsible for Europe's mild winters?' ,*Q.J.R. Meteor. Soc.* **128**, 2563-2586.
51. Wohlfahrt, J., Harrison S.P., Braconnot P., Hewitt C.D., Kitoh A., Mikolajewicz U., Otto-Bliesner B.L.and Weber S.L.: 2008, 'Evaluation of coupled ocean-atmosphere simulations of the mid-Holocene using palaeovegetation data from the northern hemisphere extratropics', *Clim.Dyn.* **31**, 871-890, DOI - 10.1007/s00382-008-0415-5.
52. Wagner, S. and Zorita, E. :2005, 'The influence of volcanic, solar and CO_2 forcing on the temperatures in the Dalton Minimum (1790-1830): a model study', *Climate Dynamics* **25**, 205-218.
53. Wang, Y.-M., Lean J. L. and Sheeley, N. R.: 2005, 'Modeling the Sun's magnetic field and irradiance since 1713', *Astrophys. J.* **625**, 522-538.
54. Wanner, H., Beer, J., Bütikofer, J., Crowley, T. J., Cubasch, U., Flückiger, J., Goosse, H., Grosjean, M., Joos, F., Kaplan, J. O., Küttel, M., Müller, S. A., Prentice, I. C., Solomina, O., Stocker, T. F., Tarasovk, P., Wagner, M., and Widmann, M.: 2008, .'Mid- to Late Holocene climate change: an overview', *Quat.. Sci. Rev.* **27**, 1791-1828.
55. Wilson, R., Tudhope, A., Borhan, P.,Briffa, K., Osborn, T. and Tett. S.:2006, 'Two-hundred-fifty years of reconstructed and modeled tropical temperatures', *J. Geophys. Res.* **111**, doi:10.1029/2005JCC003188.
56. Xoplaki, E., Luterbacher, J., Paeth, H., Dietrich, D., Steiner, N. and Grosjean, M.:2005, 'European spring and autumn temperature variability and change of extremes over the last millennium', *Geophys. Res. Lett.* **32**, L15713, doi:10.1029/2005GL023424.
57. Yoshimori, M., Stocker, T. F., Raible, C. C. and Renold, M.: 2005, 'Externally forced and internal variability in ensemble climate simulations of the Maunder Minimum', *J. Climate* **18**, 4253-4270.
58. Zorita, E., González-Rouco, F. and Legutke, S.: 2003, 'Testing the Mann et al. (1998) Approach to Paleoclimate Reconstructions in the Context of a 1000-Yr Control Simulation with the ECHO-G Coupled Climate M odel' , *J. Climate* **16**, 1378-1390.
59. Zorita, E., González-Rouco, F., von Storch, H., Montavez, J.P and Valero, F.: 2005, 'Natural and anthropogenic modes of surface temperature variations in the last thousand years', *Geophys. Res. Lett* **32**, L08707.