UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS FÍSICAS
Departamento de Física de la Tierra, Astronomía y Astrofísica II

TESIS DOCTORAL

Forced and internal variability in temperature simulations and reconstructions of the common era

Variabilidad interna y forzada en simulaciones y reconstrucciones de temperatura de los últimos dos mil años

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Laura Fernández Donado

Directores
Jesús Fidel González Rouco
Elena García Bustamante

Madrid, 2016

Forced and internal variability in temperature simulations and reconstructions of the Common Era

Variabilidad interna y forzada en simulaciones y reconstrucciones de temperatura de los últimos dos mil años

Memoria que presenta

Laura Fernández Donado

para optar al grado de

Doctor en Ciencias Físicas

Directores:

Dr. J. Fidel González Rouco
Dra. Elena García Bustamante

Departamento de Física de la Tierra, Astronomía y Astrofísica II
Facultad de Ciencias Físicas
Universidad Complutense de Madrid

Uncertainty is always with us and can never be fully eliminated from our lives, either individually or collectively as a society. Our understanding of the past and our anticipation of the future will always be obscured by uncertainty.
(...)
Has science been debilitated by uncertainty? To the contrary, the successes of science, and indeed there are many, arise from the ways that scientists have learned to make use of uncertainty in their quests for knowledge. Far from being an impediment that stalls science, uncertainty is a stimulus that propels science forward. Science thrives on uncertainty.

Henry N. Pollack, Uncertain Science... Uncertain World
Cambridge University Press, 2005

Contents

Acknowledgements

I would like to thank the labour of the two external reviewers Prof. Jürg Luterbacher and Prof. Christoph C. Raible in revising this manuscript. Thank you for your time and for helping us to improve the present piece of work.

This PhD has been funded by the FPU grant AP2009-4061 from the Ministerio de Educación, Cultura y Deporte and has been also supported by the following research projects: MOVAC (200800050084028), SPEQ-TRES (CGL2011-29672-C02-01), AVAVIP-plus (CGL2011-29677-C02-02) and ILModelS (CGL2014-59644-R). A short stay of 3 months at the WSL-Birmensdorf (Switzerland) was funded by a ESF-MedCLIVAR grant.

List of acronyms

AOGCM	Atmosphere-Ocean General Circulation Model
AR4	Fourth Assessment Report of the IPCC
AR5	Fifth Assessment Report of the IPCC
CE	Common Era
CFR	Climate Field Reconstruction
EBM	Energy Balance Model
ECS	Equilibrium Climate Sensitivity
EMIC	Earth system Model of Intermediate Complexity
ESM	Earth Syste Model
FR	Forward Regression
GHG	Greenhouse Gases
GL	Global
IPCC	Intergovernmental Panel on Climate Change
Last2k	Last 2 Millennia
LIA	Little Ice Age
LM	Last Millennium
LMTCR	Last Millennium Transient Climate Response
LULC	Land Use - Land Cover
MCA	Medieval Climate Anomaly
NH	Northern Hemisphere
PMIP3	Paleo Model Intercomparison Project Phase 3
SH	Southern Hemisphere
STSI	TSI variations of larger amplitude
ssTSI	TSI variations of weaker amplitude

TCR	Transient Climate Response
TEF	Total External Forcing
TSI	Total Solar Irradiance
VM	Variance Matching

Summary

Understanding climate variability and change, including recent anthropogenic warming, poses questions that cannot be answered based solely upon instrumental records. The Common Era (CE), and specifically the last millennium (LM), are immediate temporal intervals with comparable external radiative forcings as those of present day. Exploring the climate system blended response to the forcing conditions and to the mechanisms imposed by its internal dynamics during the CE, and particularly the LM, has the potential to expand our understanding of climate variability from inter-annual and decadal to multi-centennial timescales. This provides a wider context for current warming that might help constraining the uncertainty embedded in the future climate response to a sustained anthropogenic pressure. Proxy-based climate reconstructions, paleoclimate model simulations and estimations of external radiative forcing stand as fundamental tools that allow gaining insights about past climate variations, their amplitude and causes.

Aim of the study

This Thesis addresses the relative roles of internal variability and external natural or anthropogenic induced changes in temperature reconstructions and model simulations of the CE. To this end, available continental, hemispherical and global temperature reconstructions, the ensemble of simulations including both Paleoclimate Modelling Intercomparison Project Phase III / Coupled Model Intercomparison Project Phase 5 (PMIP3/CMIP5) and non-PMIP3 model experiments, as well as the external forcing configurations applied are herein exhaustively compiled, analysed and intercompared, thereby providing a comprehensive overview of the current state of knowledge of the temperature evolution for this period.

The relatively short ranges of external forcing variability within the CE/LM, compared to longer timescales like glacial/interglacial changes, for instance, represent a challenge in as much as the consistency between simulations and reconstructions can be affected by the large uncertainties in their respective responses to the external forcings. One of the core questions within this work relates therefore the extent to which a straight response to the external forcing can be identified during the period under study and whether this signal is common to simulated and reconstructed temperature.

The comparison of simulations and reconstructions involves thus a suitable frame to identify robust features of past climate variations and also weaknessess that might affect the model simulations, the reconstructions or both. Within such a frame this work discuses the impact that a range of generally accepted methodological approaches might have on inferences about the consistency between models and the estimations from reconstructions, overall their influence on uncertainty.

Also this work is aimed at responding to questions related to the influence of the spatial scales, from global to continental, on the temperature variations of the CE. The internal dynamics of the climate system, associated feedbacks and mechanisms gain a more significant role in as much as the regional scales reach prominence. Evidencing to what extent the role of the externally forced response in simulations and reconstructions suffers from the increasing presence of the internally driven temperature variations is one key concern of this Thesis.

Main results

This study provides a state-of-the-art comprehensive collection of mean average temperature series from model simulations and reconstructions during the CE. The thorough analysis performed herein represents an exhaustive update since AR5 by characterizing the state of knowledge of the CE temperature evolution and its relationship to the external forcing.

At hemispheric and global scales, simulations and reconstructions broadly agree on the major temperature changes above multidecadal timescales and suggest, despite the important influence of the internal variability, an overall linear response to external forcing. The rate of temperature response to LM changes in total external forcing (TEF) is quantified as a metric of the transient climate response during the LM (LMTCR) and its values from the model and reconstructed ensembles are compared to other estimates of the sensitivity of climate. The LMTCR evidences that a certain amount of response to the forcing can be considered linear, in constrast to the other estimates, that include different feedbacks and delayed responses to the forcing.

The climate sensitivity estimated from reconstructed temperatures suggests that there is also a linear response to the external forcings in the reconstructions. The ratios of temperature response to the external driving factors between reconstructions and simulations are found to be overall consistent. Nonetheless, the uncertainties involved in both, model simulations and proxy-based reconstructions condition to a great extent the coherence between reconstructed and simulated temperature. In particular, this work provides evidences of how the uncertainties associated to methodological issues in the reconstructions significantly burdens the traditionally accepted consistency between reconstructed and simulated hemispheric temperature of the CE, specially for the acknowledged warm medieval episode. The most relevant intervals of the LM are analysed to provide hints about potential explanations for their agreement and discrepancies based on the relative influences of the external versus internal variability.

At regional/continental scales, where it is well-known that the internal variability plays a more significant role, we focus on the assessment of PMIP3/CMIP5 experiments and temperature reconstructions developed within the PAGES 2k project and their responses to forcing. Temperature response among the different PAGES2k regions is more homogeneous in the simulated than in the reconstructed climate. Agreement between simulations and reconstructions is higher for Northern Hemisphere regions whilst models disagree more with the reconstructions in the Southern Hemisphere.

The consistency revealed between reconstructed and simulated temperature responses as well as their discrepancies, evidenced and quantified throughout this study, ultimately assesses how this results change our understanding of the relative roles of external forcing and internal variability during the CE and what implications the latter has for projections of future climate.

Publications related to this Thesis in which the author has participated

- Fernández-Donado, L., J. F. González-Rouco, C. C. Raible, C. M. Ammann, D. Barriopedro, E. García-Bustamante, J. H. Jungclaus, S. J. Lorenz, J. Luterbacher, S. J. Phipps, J. Servonnat, D. Swingedouw, S. F. B. Tett, S. Wagner, P. Yiou and E. Zorita, 2013: Large-scale temperature response to external forcing in simulations and reconstructions of the last millennium. *Climate of the Past*, **9**, 393-421. DOI 10.5194/cp-9-393-2013.
- Fernández-Donado, L., J. F. González-Rouco, E. García-Bustamante, J. E. Smerdon, S. J. Phipps, J. Luterbacher and C. C. Raible, 2015: Northern Hemisphere temperature reconstructions of the Common Era: Ensemble un-

certainties and their influence on model-data comparisons. *Geophysical Research Letters*, **Under review.**

- González-Rouco,J. F., L. Fernández-Donado, C. C. Raible, , D. Barriopedro, J. Luterbacher, J. H. Jungclaus, D. Swingedouw, J. Servonnat, E. Zorita, S. Wagner and C. M. Ammann, 2011:Medieval Climate Anomaly to Little Ice Age transition as simulated by current climate models. *Medieval Climate Anomaly. Pages News.*, **19**, 7-8.
- PAGES2k-PMIP3 group, 2015: Continental-scale temperature variability in PMIP3 simulations and PAGES 2k regional temperature reconstructions over the past millennium. *Climate of the past*, **Under review**.
- EuroMed2k Consortium, 2015: European summer temperatures since Roman times. *Environmental Research Letters*, **Under review**.
- Luterbacher, J., R. Neukom, J. F. González-Rouco, L. Fernández-Donado, C. C. Raible and E. Zorita, 2011: Reconstructed and simulated Medieval Climate Anomaly in southern SouthAmerica. *Medieval Climate Anomaly. Pages News.*, **19**, 20-21.
- Raible, C. C., F. Lehner, J. F. González-Rouco and L. Fernández-Donado, 2014:Changing correlation structures of the Northern Hemisphere atmospheric circulation from 1000 to 2100 AD. *Climate of the Past*, **10**, 537-550.
- Seim, A., K. Treydte, V. Trouet, D. Frank, P. Fonti, W. Tegel, M. Panayotov, L. Fernández-Donado, P. Krusic, U. Büntgen, 2014: Climate sensitivity of Mediterranean pine growth reveals distinct eastwest dipole. *International J. Climatology*,**35**, 2503-2513.

Resumen

La variabilidad del clima presente, incluyendo el calentamiento antropogénico de las últimas décadas, plantea diversas preguntas a las que no se puede dar respuesta sólo en base a los registros instrumentales. Los últimos dos mil años, y específicamente el último milenio (del inglés LM), son los períodos pasados más cercanos al presente en los que los forzamientos externos al sistema climático son en buena medida comparables a los actuales, a excepción, claro está, del forzamiento de caracter antropogénico post-industrial. Examinar la respuesta del sistema climtico a los factores externos que se mezcla con la influencia de la dinmica interna del clima permite expandir nuestro conocimiento sobre la variabilidad climática desde escalas interanuales y decadales a seculares. A su vez esto posibilita emplazar la tendencia de calentamiento actual en un contexto climático más amplio y, de este modo, contribuir a acotar la incertidumbre asociada a la respuesta futura del sistema a un forzamiento prolongado de origen antropogénico. Las reconstrucciones paleoclimáticas basadas en medidas indirectas (registros *proxy*), las paleo-simulaciones con modelos climáticos y las estimaciones del forzamiento externo radiativo son herramientas clave para una mejor caracterización de las variaciones del clima en el pasado, la amplitud de dichas variaciones así como sus posibles causas.

Objetivos centrales

Esta Tesis está enfocada a evaluar la influencia relativa de la variabilidad interna frente a la respuesta forzada en simulaciones y reconstrucciones de temperatura de últimos dos mil años. Con este objetivo fundamental, este trabajo presenta la más actualiza colección disponible de reconstrucciones de temperatura a escala continental, hemisférica y global, un amplio conjunto de simulaciones climáticas

de modelos de alta complejidad, incluyendo tanto experimentos del *Paleoclimate Modelling Intercomparison Project Phase III / Coupled Model Intercomparison Project Phase 5 (PMIP3/CMIP5)* como otros experimentos desarrollados al margen de estos proyectos coordinados, así como las configuraciones de forzamiento externo aplicadas a cada simulación. Estas fuentes de informacin han sido exhaustivamente analizadas e intercomparadas, ofreciendo una visión general del estado actual del conocimiento sobre la evolución de la temperatura durante el periodo de estudio y nos permiten explorar las respuestas del sistema climático tanto a la variabilidad interna propia del sistema como a los forzamientos externos.

Comparar las simulaciones y las reconstrucciones de temperatura del último milenio representa un reto dado que los rangos de variabilidad del forzamiento externo en este periodo son relativamente pequeños, comparados con los forzamientos característicos de escalas temporales más largas, por ejemplo los cambios durante periodos glaciales e inter-glaciales. Por tanto, el acuerdo entre las variaciones del clima que sugieren las simulaciones y las reconstrucciones puede verse afectado por la incertidumbre que afecta a la respuesta forzada tanto en las simulaciones como en las reconstrucciones. Una de las cuestiones fundamentales a que esta Tesis pretende dar respuesta consiste en entender en qué medida se puede identificar una respuesta directa al forzamiento externo que sea común a las simulaciones y a las reconstrucciones en el periodo de estudio.

La comparación entre simulaciones y reconstrucciones garantiza un marco adecuado que permite identificar qué aspectos son consistentes en los cambios del clima sugeridos por ambas fuentes de información, de igual modo que facilitan el poder identificar las limitaciones que pueden afectar a simulaciones, a reconstrucciones o a ambas. Es en ese marco comparativo en el que esta Tesis presenta una discusión acerca de la influencia que ejerce un conjunto de estrategias metodológicas comúnmente aceptadas sobre el grado de consistencia entre simulaciones y reconstrucciones del último milenio que se ha asumido hasta la fecha. Es por tanto un objetivo fundamental de este trabajo establecer las premisas para un profundo debate acerca de las incertidumbres que afectan a simulaciones y reconstrucciones y cómo ello produce un impacto en la comparación entre ambas fuentes.

Adicionalmente, este trabajo está orientado a responder preguntas relacionadas con la influencia de las distintas escalas espaciales, desde lo global hasta escalas espaciales continentales, en las variaciones de temperatura de los dos últimos milenios. El papel que desempeña la variabilidad interna del sistema climático y los mecanismos físicos asociados a ella se ve amplificado a medida que cobran importancia las escalas espaciales más regionales. Un aspecto fundamental en este trabajo es evidenciar hasta qué punto la respuesta en temperatura al forzamiento externo se ve afectada por un aumento progresivo de la presencia

de variabilidad de origen interno en los cambios de temperatura de los últimos dos mil años.

Resultados más relevantes de esta Tesis

Este estudio proporciona una compilación estado del arte de simulaciones y reconstrucciones de los últimos dos mil años. Los análisis exhaustivos que se han llevado a cabo en este trabajo representan un actualizaci'on completa y detallada del último informe del IPCC (AR5).

A escala global y hemisférica, simulaciones y reconstrucciones muestran un amplio acuerdo en los cambios de temperatura evidenciados y sugieren, a pesar de la influencia de la variabilidad interna, una respuesta lineal al forzamiento externo en escalas multidecadales y más largas. Se ha estimado el ratio entre la respuesta en temperatura durante el LM y cambios en el forzamiento externo total (del inglés TEF), lo que proporciona una métrica de la respuesta climática transitoria durante el LM (del inglés LMTCR). Los valores de LMTCR obtenidos en los conjuntos de simulaciones y reconstrucciones son comparados con otras estimaciones de la sensibilidad climática y la respuesta transitoria. El LMTCR muestra que una buena porción de la respuesta de la temperatura al forzamiento en las simulaciones es lineal frente al tipo de respuesta que proporcionan otras estimaciones de la sensibilidad climática que incluye respuestas no linales o con retardo al forzamiento.

La sensibilidad climática que muestran las reconstrucciones de temperatura sugiere también una cierta linealidad en su respuesta al forzamiento externo del sistema. Se observa que existe un buen grado de acuerdo en las respuestas respectivas si se compara la temperatura simulada con la reconstruída, si bien las simulaciones tienden a sobrestimar la respuesta al forzamiento en comparación con las reconstrucciones. La incertidumbre que afecta a ambas condiciona sin embargo el grado de acuerdo que puede esperarse entre la temperatura según los modelos climáticos y la reconstruída. En particular, en esta Tesis se proporcionan evidencias acerca de cómo las incertidumbres relacionadas con el dominio espacial que representan las reconstrucciones menoscaban en buena medida el grado de acuerdo admitido hasta la fecha entre la temperatura simulada y reconstruída, especialmente durante el periodo cálido medieval. Los periodos más notables del LM se han explorado con el fin de proporcionar argumentos que sustenten tanto el acuerdo como las discrepancias entre simulaciones y reconstrucciones.

A escalas regionales/continentales se analiza la consistencia entre simulaciones y reconstrucciones PAGES2k, así como su relación con el forzamiento externo en escalas interanuales y decadales. Los resultados obtenidos varían para cada

región, observándose una mayor consistencia entre simulaciones y reconstrucciones en regiones del Hemisferio Norte y un mayor desacuerdo en las regiones del Hemisferio Sur. Las simulaciones muestran una mayor homogeneidad espacial que las reconstrucciones.

El grado de consistencia que se evidencia en este trabajo entre la temperatura simulada y reconstruída, así como las discrepancias constatadas y cuantificadas este estudio, tienen un efecto en nuestra comprensioón y nuestro conocimiento sobre la importancia relativa del forzamiento externo del sistema climático frente a su variabilidad interna durante los últimos dos mil años, lo que en último término, tiene implicaciones para las proyecciones de cambio climático en condiciones de forzamiento externo prolongado en el futuro.

Publicaciones relacionadas con la Tesis en las que ha participado la autora

- Fernández-Donado, L., J. F. González-Rouco, C. C. Raible, C. M. Ammann, D. Barriopedro, E. García-Bustamante, J. H. Jungclaus, S. J. Lorenz, J. Luterbacher, S. J. Phipps, J. Servonnat, D. Swingedouw, S. F. B. Tett, S. Wagner, P. Yiou and E. Zorita, 2013: Large-scale temperature response to external forcing in simulations and reconstructions of the last millennium. *Climate of the Past*, **9**, 393-421. DOI 10.5194/cp-9-393-2013.
- Fernández-Donado, L., J. F. González-Rouco, E. García-Bustamante, J. E. Smerdon, S. J. Phipps, J. Luterbacher and C. C. Raible, 2015: Northern Hemisphere temperature reconstructions of the Common Era: Ensemble uncertainties and their influence on model-data comparisons. *Geophysical Research Letters*, **En revisión**
- González-Rouco,J. F., L. Fernández-Donado, C. C. Raible, , D. Barriopedro, J. Luterbacher, J. H. Jungclaus, D. Swingedouw, J. Servonnat, E. Zorita, S. Wagner and C. M. Ammann, 2011:Medieval Climate Anomaly to Little Ice Age transition as simulated by current climate models. *Medieval Climate Anomaly. Pages News.*, **19**, 7-8.
- PAGES2k-PMIP3 group, 2015: Continental-scale temperature variability in PMIP3 simulations and PAGES 2k regional temperature reconstructions over the past millennium. *Climate of the past*, **En revisión**.
- EuroMed2k Consortium, 2015: European summer temperatures since Roman times. *Environmental Research Letters*, **En revisión**.
- Luterbacher, J., R. Neukom, J. F. González-Rouco, L. Fernández-Donado, C. C. Raible and E. Zorita, 2011: Reconstructed and simulated Medieval Climate Anomaly in southern SouthAmerica. *Medieval Climate Anomaly. Pages News.*, **19**, 20-21.

- Raible, C. C., F. Lehner, J. F. González-Rouco and L. Fernández-Donado, 2014:Changing correlation structures of the Northern Hemisphere atmospheric circulation from 1000 to 2100 AD. *Climate of the Past*, **10**, 537-550.
- Seim, A., K. Treydte, V. Trouet, D. Frank, P. Fonti, W. Tegel, M. Panayotov, L. Fernández-Donado, P. Krusic, U. Büntgen, 2014: Climate sensitivity of Mediterranean pine growth reveals distinct eastwest dipole. *International J. Climatology*,**35**, 2503-2513.

1

Introduction

The level of understanding of the present climate state relies to a large extent on the analysis of instrumental data (e.g. Myhre et al., 2013; Hartmann et al., 2013; Vaughan et al., 2013) and of experiments with Atmosphere-Ocean General Circulation Models (AOGCMs, e.g. Randall et al., 2007; Flato et al., 2013; Collins et al., 2013). Current climate conditions can be viewed as the result of different processes interacting at a wide range of timescales, many of which are longer than the length of the instrumental period (Peixoto and Oort, 1984; Houghton, 2005; Masson-Delmotte et al., 2013). Such sources of variability may be related to internal dynamics and feedbacks in the system (Delworth and Zeng, 2012), or may be a response to changes in natural or anthropogenic external forcings (Ottera et al., 2010; Myhre et al., 2013). The limited time span of the instrumental period (e.g. Brohan et al., 2006; Lawrimore et al., 2011) hampers the study of mechanisms operating at long temporal scales and the characterization of the levels of internal and forced variability of the climate system. The use of AOGCMs and the analysis of indirect (proxy) sources of climate information can help to expand the knowledge gained from instrumental records alone to multicentennial and longer timescales (Jones et al., 2009; Masson-Delmotte et al., 2013).

In contrast to other paleoclimatic periods such as the Last Glacial Maximum (21000 years before present), the climate conditions of the Late Holocene are, to a great extent, comparable to present climate. Therefore the recent warming trends can be placed in a wider context that considers climate variations in the preceding period (Jansen et al., 2007; Masson-Delmotte et al., 2013). The availability of high resolution proxy data for the last thousand years in comparison to earlier periods allows the development of reconstructions of the temporal evolution, and sometimes even the spatial distribution, of some important climate parameters as well as the past external forcing conditions. Reconstructions of external forcings

have been used, in turn, as boundary conditions for AOGCM simulations, many of which span at least the last millennium. Indeed, it is specifically the last millennium (LM) the period that can be considered as an optimal time interval to frame the assessment of current climate variability, as it includes the largest amount of proxy reconstructions and simulations. Nevertheless, recent advances in simulating and reconstructing past climate (Masson-Delmotte et al., 2013; PAGES 2k Consortium, 2013; PAGES2k-PMIP3 group, 2015) have allowed to extend this analysis back in time to different periods within the Common Era (CE; from year 0 onwards), even extending beyond the last two millennia (Last2k) in the case of some reconstructions(e.g. Moberg et al., 2005; Ljungqvist, 2010; Christiansen and Ljungqvist, 2012b).

Reconstructions and simulations are the main tools for improving our knowledge of the evolution of past climate. Nevertheless, they are subject to their own strengths and weaknesses since they are both affected by various sources of uncertainty. These issues will be discussed in the following sections.

1.1 Climate reconstructions and simulations of the Common Era

Climate reconstructions are based on documentary observations as well as on geological and biological data that offer indirect (proxy) information of past climate variability (Jones et al., 2009). These data are used in statistical models that are calibrated with instrumental data and they can therefore provide an estimation of the past climate evolution of a particular variable of interest (North et al., 2006). Although most of the studies have focused on the reconstruction of past temperature and precipitation, a considerable number of them have also explored atmospheric circulation patterns and indices (e.g. Luterbacher et al., 2002; Trouet et al., 2009; Ortega et al., 2015). Climate reconstructions have targeted different spatial scales, from local, regional and continental (e.g. Linán et al., 2012; PAGES 2k Consortium, 2013) to hemispheric and global (e.g. Briffa et al., 1998; Mann et al., 2008).

In turn, numerical simulations of the CE climatic states have used models of varying complexity, from Energy Balance Models (EBM; e.g. Crowley, 2000; Hegerl et al., 2006) and Earth system Models of Intermediate Complexity (EMIC; e.g. Bauer et al., 2004; Goosse et al., 2005) to comprehensive AOGCMs (e.g. Ammann et al., 2007; González-Rouco et al., 2003b; Servonnat et al., 2010; Swingedouw et al., 2010) or Earth System Models (ESM; e.g. Jungclaus et al., 2010; Lehner et al., 2015), which include a more realistic representation of some system components, such as a dynamic vegetation or a carbon cycle component. Such simulations contribute to the understanding of the climate of the CE and may

have implications on the estimations of future climate change. For instance they can contribute to a better understanding of the climate response to natural and anthropogenic external forcings and the mechanisms involved (e.g. Zorita et al., 2005; Goosse et al., 2006; Ammann et al., 2007); to a validation of the response to external forcing in simulations through comparison with climate reconstructions (e.g. Crowley, 2000; Bauer et al., 2003; Hegerl et al., 2011), they can serve to narrow the ranges of estimates of climate sensitivity (Hegerl et al., 2006); or they can be used as a pseudo-reality to validate the statistical methodologies applied in proxy-based climate reconstructions (see Smerdon, 2012).

The present work will focus on the analysis of temperature variability during the CE and for this purpose, it will make use of these two paleoclimatic sources of information at global, hemispheric and continental scales.

The availability of AOGCM/ESM simulations and reconstructions (Masson-Delmotte et al., 2013) is a result of a community effort over the last two decades. The recent development of both sources of paleoclimate information is linked in turn, among other factors, to the evolution of the computational knowledge and the growing availability of estimates of changes in past external forcing and climate reconstrucions throughout the advancements from the proxy collections and the methodologies applied. Attending to these aspects, the following section presents an overview of the evolution of our understanding of the temperature evolution during the CE. This knowledge has only lately been extended to the first millennium (0 - 1000 CE) and since the late 1980s most research has been focused on the LM. The next section briefs the main steps of this evolution mostly sketched by the milestones of Intergovernmental Panel on Climate Change (IPCC) reports.

1.2 An overview of 25 years analysing the climate of the Last Millennium

The importance of improving our knowledge about LM climate was already acknowledged in the IPCC First Assessment Report by Folland et al. (1990), in which a schematic plot of the global LM temperature evolution was provided. Figure 1.1a reproduces the temperature series shown in Folland et al. (1990). Although the illustration did not include a temperature scale, that schematic representation already evidenced a warming period at the beginning of the millennium, the so-called Medieval Climate Anomaly (MCA, previously known as Medieval Warm Period; Lamb, 1965) followed by a cooler period, the so-called Little Ice Age (LIA; Diaz et al., 2011). This schematic plot was derived from the study of Lamb (1965) over Central England (see discussion in Jones et al.,

2009) as a quantitative estimation of the large-scale LM temperature variability was missing at that time. Nevertheless, various regional reconstructions (e.g Briffa et al., 1990) were available at that time as well as a couple of Northern Hemisphere (NH) temperature reconstruction exercises reaching back the 16-17th centuries (Groveman and Landsberg, 1979; Jacoby and D'Arrigo, 1989). Groveman and Landsberg (1979) presented a NH temperature reconstruction since 1579 AD based mostly on early instrumental data. In contrast, Jacoby and D'Arrigo (1989) provided the first NH temperature reconstruction back to 1671 based entirely on non-instrumental data and they suggested despite the modest techniques available, the importance of preserving low-frequency variability. Both reconstructions already evidenced that the 20th century was anomalously warm in the context of the reconstructed inmediate past.

Supporting the latter result, Bradley and Jones (1993) provided a new NH temperature reconstruction spanning a longer period, since 1400 to 1960 CE. This series (Figure 1.1b) was the first paleoclimatic evidence included in the Second Assessment Report (AR2; Nicholls et al., 1995) and it provided a quantitative estimation of the LM NH temperature evolution. AR2 prompted the need for an evolution of this field that would ideally span the whole LM and helped to improve the knowledge about key periods like the MCA at hemispherical and global scales.

It was by the time of the Third Assessment Report (AR3; Folland et al., 2001) when the first 1000 yr long NH reconstructions were introduced (Mann et al., 1999; Jones et al., 1998). Interestingly, the number of reconstructions by that time enabled (Figure 1.1d; Folland et al., 2001) the first intercomparison of NH temperature reconstructions in AR3 (Mann et al., 1999; Jones et al., 1998; Briffa, 2000). All reconstructions (Jones et al., 1998; Mann et al., 1999; Briffa, 2000) agreed on pointing out the 1990s as the warmest decade of the LM. Indeed, according to the existing evidences, past temperature changes were characterized by moderate low frequency variability. A large-scale signature of the MCA was not evident at hemispheric scale and the transition from the relatively warmer medieval to the cooler preindustrial times was essentially linear (Figure 1.1d).

The Mann et al. (1999) reconstruction (Figure 1.1c) showed, for the first time, an explicit estimation of quantitative uncertainty and was based on a remarkable amount of individual proxies: more than 100 after 1400 CE and 12 reaching the beginning of the LM. This series (Mann et al., 1999) fostered considerable discussion within the community (see Jansen et al., 2007, for a review) that raised critical issues about the method applied and the uncertainties involved. The influence of various subjective methodological decisions, during the calibration and reconstruction intervals, such as the normalization of the data or the choice of the centering period were under debate (e.g. McIntyre and McKitrick, 2003, 2005;

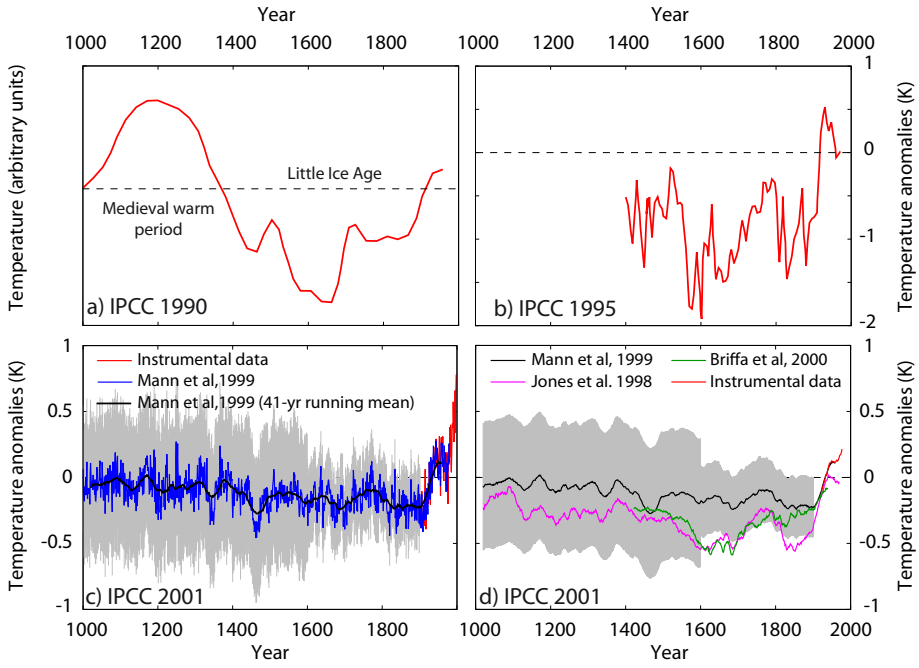

Fig. 1.1: Representation of the evolution of knowledge about the LM temperature variation as represented in the first three IPCC assessment reports: a) Schematic representation included in Folland et al. (1990); b) Bradley and Jones (1993) temperature reconstruction as shown in Nicholls et al. (1995); c) Representation in Folland et al. (2001) of the Mann et al. (1999) reconstruction together with its uncertainty; d) NH temperature reconstructions (Jones et al., 1998; Mann et al., 1999; Briffa, 2000) intercomparison exercise included in Folland et al. (2001) in which the gray band corresponds to the uncertainty of Mann et al. (1999) reconstruction. Panels b-d show temperature anomalies wrt 1961-1990 CE. All figures are adaptations from the original figures in the IPCC reports.

Von Storch and Zorita, 2005; Buerger and Cubasch, 2005; Buerger et al., 2006). In particular, the decision about whether detrending or not the data during the calibration period was an issue that raised considerable debate (Von Storch et al., 2004; Wahl et al., 2006; Von Storch et al., 2006b,a; Rahmstorf et al., 2006). Trends in proxy records during the industrial period can include relevant temperature information (Ammann and Wahl, 2007). However, they also can be induced by other (non-temperature) climate and non-climatic influences (Gagen et al., 2011;

Jones et al., 2009), thereby with a potential of biasing the calibration to undesired signals. In spite of this, experiments generating ensembles of reconstructions with various methods and pseudoproxies contamined with noise showed that reconstruction had a poorer quality when calibration was done with detrended data (Christiansen et al., 2009). Most of this discussion was developed under the concern that methods used so far would underestimate the amplitude of past low frequency variability (e.g. North et al., 2006; Jansen et al., 2007; Masson-Delmotte et al., 2013). Other methodological and technical proxy-related issues extended the discussion to different reconstructions showing various relatively high ranges of low frequency variability like a tree-ring maximum density reconstruction (Esper et al., 2002; Mann and Hughes, 2002), a multi-proxy reconstruction blending the information from different proxy types to reconstruct different spectral bands (Moberg et al., 2005; Mann et al., 2005) or some borehole reconstructions that specifically target long term trends (Mann and Schmidt, 2003; González-Rouco et al., 2003a, 2006; Mann and Jones, 2003; Pollack and Smerdon, 2004). The use of AOGCM simulations as a surrogate reality in which pseudoproxy records are created and reconstruction methods can be replicated and tested (see Smerdon, 2012, for a review) has provided improved understanding of reconstruction methods and their potential biases. An important finding from these studies is that many published reconstructions may indeed understimate the amplitude of low frequency variability (Christiansen et al., 2009; Lee et al., 2008; Smerdon et al., 2010). This effect has been shown to be diminished in some reconstructions if temporal smoothing techniques were applied increasing in that way the correlation between the proxy series and the temperature data (Lee et al., 2008), or by using different variants of compositing plus scaling (Christiansen and Ljungqvist, 2011; Christiansen, 2011; Hegerl et al., 2007b). The degree of amplitude attenuation in real world reconstructions is uncertain, although it will increase for cases of weaker correlation between instrumental temperatures and proxies, i.e. a low signal to noise ratio (Christiansen et al., 2009; Lee et al., 2008; Smerdon et al., 2011), if errors in the proxy or instrumental data are not taken into account (Ammann et al., 2010; Hegerl et al., 2007b) or, as discussed above, if the data are detrended in the calibration phase (e.g. Christiansen et al., 2009; Lee et al., 2008).

The Fourth Assessment Report of the IPCC (AR4; Jansen et al., 2007), made a qualitative leap forward towards gaining knowledge about the LM NH temperature. On the one hand, a total of 12 NH temperature reconstructions were considered for the LM. In addition to the three reconstructions already cited in Folland et al. (2001), 9 new reconstructions (e.g. Esper et al., 2002; Mann and Jones, 2003; Moberg et al., 2005; D'Arrigo et al., 2006; Hegerl et al., 2006) were included; some of them spanning the last 1200 years. Figure 1.2a shows

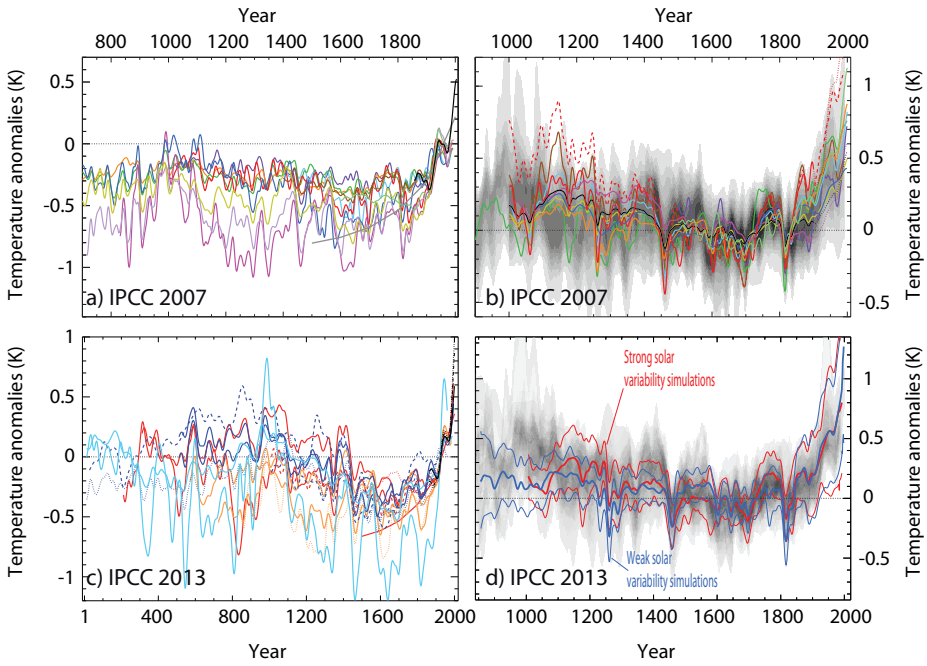

Fig. 1.2: LM reconstructions and simulations as represented in the two last IPCC assessment reports: a) NH temperature reconstructions included in Jansen et al. (2007); b) comparison between model simulations and reconstructions as shown in Jansen et al. (2007); simulations (lines) are shown over the uncertainty band built from the reconstructions ensemble; c) NH temperature reconstructions included in Masson-Delmotte et al. (2013); d) model-data comparison exercise as shown in Masson-Delmotte et al. (2013); two ensembles of simulations (lines), attending to two different levels of variability of solar forcing applied, are shown over the uncertainty band built from the reconstructions ensemble. Panels a) and c) show temperature anomalies wrt 1961-1990 CE, while b) and d) refer to 1500-1899 and 1500-1850 CE, respectively. All figures are adaptations from the original figures in the IPCC reports.

the complete ensemble of reconstructions included in AR4 and illustrates the consensus on evidencing a large-scale warmer time interval during the MCA and recent decades but also a relatively large dispersion of records in depicting the amplitude of past temperature changes (Frank et al., 2010a). This dispersion of the series illustrates the existing uncertainty among the reconstructions but also allowed an estimation of the most probable evolution of the last millennium NH temperature (grey shading in Figure 1.2b).

On the other hand, AR4 also presented, for the first time, simulated evidences to describe the LM temperature evolution and to compare with the available reconstructions. Most of these experiments came from EBMs or EMICs, while the AOGCM experiments spanning the LM were a minority (González-Rouco et al., 2003b, 2006; Ammann et al., 2007). The scarcity of these kind of experiments coming from comprehensive AOGCMs with high temporal and spatial resolution was largely related to the extensive computational requirements (González-Rouco et al., 2009; Braconnot et al., 2012). Based on all these reconstructed and simulated evidences, AR4 presented a first model-data comparison of the LM temperature evolution (Figure 1.2b). Simulations are plotted over the uncertainty band of the reconstructions, illustrating a significant agreement between both sources, i.e. simulations are embedded within the ensemble spread of the reconstructions with largest discrepancies occurring during the MCA.

Since AR4 new progress have been made within the reconstructions and the simulations of the CE, as evidenced in the Fifth Asessment Report (AR5; Masson-Delmotte et al., 2013). A total of 15 reconstructions were included in AR5 as an update of AR4. For the first time the assessment was extended to the Last2k, as illustrated in Figure 1.2c. This collection of series (Figure 1.2c), as in the case of AR4 (Figure 1.2a), agreed on showing the acknowledged warm and cold intervals of the LM. Nevertheless it also evidenced large differences in the low frequency temperature variability mong the different reconstructions, which is used as an estimation of the level of associated uncertainty.

AR5 introduced the relevance of considering proxy-based reconstructions at the regional scale. Regional reconstructions allow to better understand the intrinsic spatial heterogeneity of large-scale climate, the interaction of the large-scale circulation with the finer resolution details of orography and land surface properties as well as the mechanisms involved in its variability. Within this context, important advances have also taken place during last years, highlighting the continental-scale reconstructions developed by the coordinated effort of the PAGES 2k Network (PAGES 2k Consortium, 2013; McKay and Kaufman, 2014). This joint effort has produced so far a total of eight new continental-scale temperature reconstructions based on a global data set of proxy records developed during last years (Kaufman et al., 2014). Figure 1.3 illustrates the reconstructed

regional temperature changes for four of the PAGES2k continental areas as an example. All continents except North America show warming trends in the industrial period and some of them (Arctic, Europe, North America) warmer temperatures during medieval times than in the so-called LIA centuries before 1850 CE. The ranges of multidecadal to multicentennial variability vary largely for different continents, with Arctic or Europe showing much higher variability than North America, for example. This evidence, as well as model-data comparison studies at continental scales (PAGES2k-PMIP3 group, 2015) has revealed new insights about internally and externally driven variability.

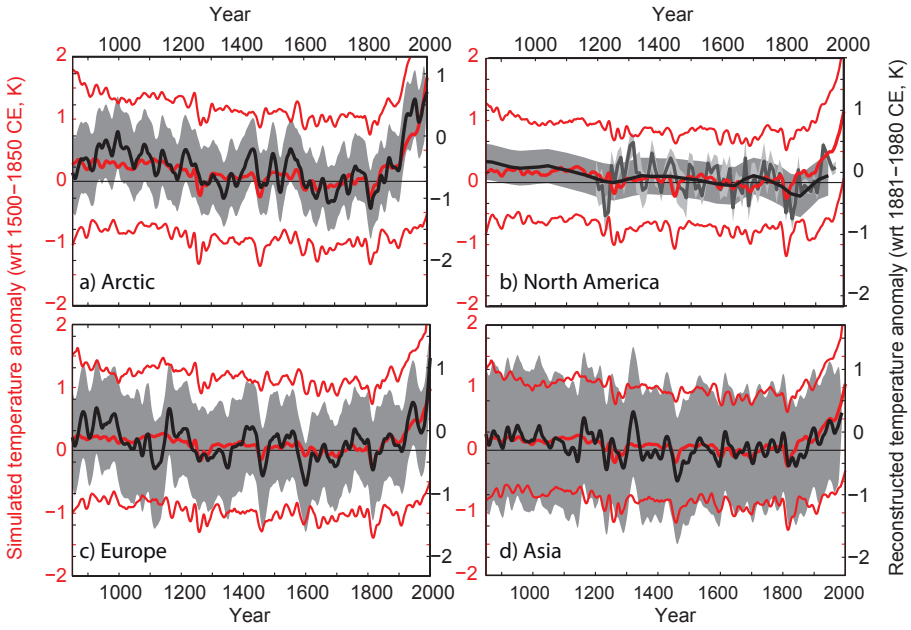

Fig. 1.3: LM temperature reconstructions and simulations as shown in Masson-Delmotte et al. (2013) for the regions: a) Arctic; b) North America; c) Europe; and d) Asia. Reconstructions (black lines) are shown over a shaded envelope depicting their uncertainties; their anomalies are calculated with reference to 1881-1980 CE. Temperature anomalies from PMIP3/CMIP5 simulations are shown in red (thick line stands for the multi-model mean; thin lines represent the 90 % multi-model-range) and they are calculated with reference to the period 1500 - 1850 CE. All panels are adaptations from the original figures in the IPCC technical summary (Stocker et al., 2013).

In turn, during last years, the impressive increase in computing power has favoured the development of new AOGCM/ESM paleosimulations. Thus, the number of high complexity model simulations has increased considerably from AR4 to AR5, where a total of 18 LM simulations were included, some of them extending to the last 1200 years. Figure 1.2d shows the model data comparison presented in AR5, in which, instead of the individual simulations, two ensemble averages of two simulations subgroups were shown over the uncertainty band provided by the reconstructions. The two subensembles of simulations were grouped according to the high or low level of solar forcing variability applied as in Fernández-Donado et al. (2013). The ensemble applying a solar forcing with higher variability evidenced a slightly warmer MCA; nevertheless, the temporal evolution of both ensemble means over the rest of the LM was quite similar. The simulations were embedded in the reconstructions spread, showing, despite some discrepancies during the MCA, a general qualitative agreement between both simulated and reconstructed evidences. In addition to these conclusions, that can be considered as an update of AR4, AR5 provided a basic assessment of the temperature response to external forcing (volcanic and solar) in simulations and reconstructions over the LM. Simulations evidenced a larger response to volcanic forcing than reconstructions at interannual scales, although these responses became similar in magnitude at multidecadal timescales with a larger time interval needed from the reconstructions to recover pre-eruption temperature values. The response to the solar forcing analysed during solar minima intervals was also very similar in reconstructions and simulations, nevertheless these responses were influenced also by the effect of volcanic forcing. Furthermore, the temperature changes associated to the MCA-LIA transition obtained by simulations and reconstructions were intercompared in AR5. Reconstructions showed a larger spread of temperature changes than simulations, and it was highlighted that at hemispheric scales the temperature changes for the MCA-LIA transition were consistent with the changes in external forcing (Fernández-Donado et al., 2013). However, when considering the spatial distribution of these temperature differences, the spatial patterns obtained by the available reconstructions (Mann et al., 2009; Ljungqvist et al., 2012), were not reproduced by the simulations (González-Rouco et al., 2011; Fernández-Donado et al., 2013; Masson-Delmotte et al., 2013). The discrepancies were larger over the equatorial Pacific Ocean, where the reconstructions indicated a mean state during the MCA compatible with stable "La Niña" conditions (Mann et al., 2009). The fact that this pattern is not reproduced by any simulation is suggestive of several possibilities. One option hints deficiencies in the reconstructions during this period potentially related to problems in the climate field reconstruction methods (CFRs; Smerdon et al., 2010, 2015). The other posibility is that AOGCMs cannot reproduce the

large scale structure of the spatial pattern because either they do not have a correct representation of the mechanisms of response to changes in external forcings (Meehl et al., 2009; Swingedouw et al., 2010), or the spatial variability during the MCA is largely influenced by internal variability (Goosse et al., 2012b). The influence of the internal variability is also highly noticed at regional scales when considering the multi-model-data comparison exercises based on the continental PAGES2k reconstructions, as the ones shown in Figure 1.3. Note that the model ensemble reproduces many of the details of the temperature evolution in the four different regions indicating a clearly detectable signature of external forcing on both simulations and reconstructions (Chapter 7; PAGES2k-PMIP3 group, 2015).

In approximately two decades, the state of knowledge about the evolution of the NH last millennium temperature has evolved from Figure 1.1a to Figures 1.2c,d and 1.3. The consistency between simulations and reconstructions, both for the NH and at continental scales, is remarkable considering the large amount of available evidences and that they are independent sources of information. The model-data comparison exercises, like those proposed in AR5 and others discussed below represent the main tool to improve our understanding about changes in climate variability during the CE and the mechanisms involved. Both simulations and reconstructions are subject to different sources of uncertainty that influence model-data comparison exercises. Understanding, and if possible, narrowing this uncertainty becomes a challenge and helps moving forward in our knowledge of climate variability and change during the CE.

1.3 Uncertainty in simulations and reconstructions of the Common Era

This work focuses on assessing the evolution of temperature during the CE and will thus make extensive use of a diversity of reconstructions and models to explore the plausible range of temperature variability. For this purpose, an overview of the main limitations and uncertainties regarding both paleoclimatic tools becomes relevant.

Uncertainty in reconstructions is associated with how proxy archives respond to local/regional past temperatures (Jones et al., 2009) and with the skill of the methods used to infer past hemispheric-scale temperature changes from information at local/regional scales (Smerdon, 2012). Proxy data have different temporal resolutions and spatial coverages, are often biased to certain seasons, and show sensitivity to different climate parameters, as well as environmental factors not necessarily related to climate (Jones et al., 2001, 2009). When used in multiproxy approaches that integrate local or regional information from different sources and

areas, all these factors may contribute to larger uncertainties and noise. In addition, the integration of local and/or regional proxy information into large scale, hemispheric or global reconstructions, is performed with a variety of methodological approaches: from simple compositing and scaling of local/regional series (e.g. Hegerl et al., 2007a; Mann et al., 2008; Ljungqvist, 2010), and regression based approaches of various levels of complexity (e.g. Guillot et al., 2015; Tingley and Li, 2012; Tingley et al., 2012; Luterbacher et al., 2004; Mann et al., 2009) to Bayesian methods (Tingley and Huybers, 2010; Li et al., 2010; Werner et al., 2013). Most of these methods are prone to various degrees of variance loss that can affect the amplitude of low frequency variability (Masson-Delmotte et al., 2013). This loss can arise, among other factors, from variance underestimation implicit in regression methods, contribution of low frequency nonclimatic noise from proxies that perturbs the calibration process, low climate signal to noise ratios in proxies or spatial underrepresentation (e.g. Buerger and Cubasch, 2005; Buerger et al., 2006; Juckes et al., 2007; Christiansen et al., 2009; Smerdon, 2012). This loss of variance can be reduced if the reconstruction method is refined, by, for example, attributing part or all of the temperature-proxy differences to errors in proxy data and thus increasing the correlation between proxies and temperature (e.g. Hegerl et al., 2007b). The latter can be achieved if, for instance, a site-by-site calibration is applied (Christiansen and Ljungqvist, 2012b). Nevertheless, this strategy might also inflate the variability, leading to an upper bound for low-frequency variability (Moberg, 2013). Additionally, the variance can also be inflated when not considering the existence of nonlinear relationships between proxies and the climate variable (Emile-Geay and Tingley, 2014).

Moreover, the somewhat subjective methodological decisions in the process of implementing these reconstruction methods may also have an impact, specially on the reconstruction ensemble spread and on the assessments involving comparisons of climate reconstructions and simulations. Related to, for instance, the selection of the reconstruction method and the choice of the configuration for the statistical model, the selection of regression parameter values, the distribution of proxies and the decisions made during the training process. The latter includes the choice of the instrumental dataset, the calibration and verification periods or the spatial domain represented by the target variable (e.g. Buerger et al., 2006; Frank et al., 2010b; Smerdon, 2012).

In turn, despite AOGCM and ESM simulations provide the most comprehensive and exhaustive representation of the climate system, they also contain weaknesses that arise from the fact that models, independently on their level of complexity, are a simplified representation of reality (Oreskes et al., 1994; von Storch, 2010). Models are also subject to uncertainties that stem from a relatively coarse temporal and spatial resolution and a limited knowledge of the climate

system and its interactions with external forcings. This leads to model codes being in continuous improvement introducing new developments as they become available and computing resources enable higher complexity experiments (Flato et al., 2013; Oreskes, 2003). These sources of uncertainty also contribute to model spread in projections of future climate change (Collins et al., 2013; Flato et al., 2013). The slightly different approaches to simulate atmosphere/ocean dynamics and different schemes for unresolved physical processes such as cloud feedbacks in the atmosphere (e.g. Soden and Held, 2006) are major contributors to model uncertainty. Additionally, the limitations in the representation of the various climate subsystems like ice sheets, permafrost, land surface processes, convection parametrizations, etc, all contribute to each model having different climate sensitivities (Flato et al., 2013). Further, uncertainties in the estimation of the evolution of forcing factors during the last millennia add to the uncertainty of the model response.

Forcing factors such as changes in orbital parameters are well known (e.g. Berger and Loutre, 1991; Laskar et al., 2004), and this is arguably also the case of changes in the concentrations of greenhouse gases (e.g. Joos and Spahni, 2008). Other forcings such as solar variability, the effect of volcanic eruptions or land use are comparatively more uncertain (Fernández-Donado et al., 2013; Schmidt et al., 2011, 2012). Thus, different simulations of the climate of the last millennium have used different forcing specifications as new and improved estimates became available. Also depending on the competence of models to account for given forcings in their code, the implementation of external forcings has been model-dependent (Jones and Mann, 2004).

An attempt to reduce the uncertainty related to the implementation of external forcings in simulations of LM climate has lead to a new generation of AOGCM simulations within the frame from the Coupled Model Intercomparison Project Phase 5-Paleoclimate Modelling Intercomparison Project Phase 3 (CMIP5-PMIP3, Taylor et al., 2012). In contrast to previous experiments, this coordinated effort focuses on developing simulations with some common criteria on the use of established external forcing configurations (Taylor et al., 2012). In the case of the LM, an specific selection of various external forcing reconstructions has been provided in Schmidt et al. (2011, 2012).

1.4 Comparison of simulations and reconstructions

In spite of the uncertainties existing in both sources of information, assessing the consistency between climate reconstructions and simulations seems pertinent given that AOGCMs/ESMs are the main tools for producing projections of future climate change (Collins et al., 2013; Kirtman et al., 2013). The comparison

between both approaches enables the climate model evaluation prior to the instrumental era (Cane et al., 2006; Braconnot et al., 2012). These exercises are also important because knowledge about past external forcings and the temperature response of the system is informative about the relative role of internal versus forced variability (Schurer et al., 2013; Fernández-Donado et al., 2013) and ultimately about the Earth system energy balance (Crowley, 2000; Trenberth et al., 2009) and its climate sensitivity (Hegerl et al., 2006).

Assessments of consistency between climate reconstructions and simulations are not only burdened by the various sources of uncertainty discussed above, but also by the fact that both approaches target conceptually different representations of reality. While climate reconstructions aim to capture the precise evolution of a climate variable in the past, simulations provide a time evolution that is consistent with the physical equations and with the imposed initial and boundary conditions. In fact, different simulations performed with the same climate model generate different climate evolutions (e.g. González-Rouco et al., 2009) when started from different initial conditions (Lorenz, 1963). Therefore, an ensemble of simulations using identical boundary specifications (external forcing) and performed either with different models or with the same model but starting from different initial conditions would only be comparable in those aspects that relate to the forced model response. Accordingly, climate simulations and reconstructions will only be correlated when the system evidences a response to the external forcing and to the extent that the current estimations of external forcings used to drive the model experiments are reliable.

An additional important point in model-data comparison relates to the spatial aggregation of proxy and AOGCM information. This rationale relates to the fact that AOGCMs show highest skill on large scales (von Storch, 1995, 2004) while proxy based reconstructions often target local and regional scales. Strategies to circumvent this problem may be derived based on upscaling (e.g. Jones and Widmann, 2003), downscaling (e.g. Wagner et al., 2007) or forward modelling of proxy variables (e.g. Evans et al., 2006; Ohlwein and Wahl, 2012; Baker et al., 2012; González-Rouco et al., 2009). Alternatively, considering continental, hemispheric and global scales, where the influence of internal variability is reduced by spatial averaging, constitutes a sound basis for the comparison of simulations and reconstructions and optimizes the detection of the external forcing signal (e.g. Schurer et al., 2013).

The most comprehensive model-data comparison assesments before this Thesis can be found in AR4 and AR5, as mentioned in Section 1.2. Additional to those assessments, an important amount of works including some level of model-data comparison have been published, most of them focused either on assessing the consistency of a new numerical experiment in the context of the available

reconstructions (e.g. Jungclaus et al., 2010; Servonnat et al., 2010; Lehner et al., 2015), or focused on comparing new reconstructions with some model evidences (e.g. Mann et al., 2009; Moberg et al., 2005). However, analyses of the multi-model consistency with the ensemble of available reconstructions are more limited. This type of exercise targets a better understanding of the role of the external forcing and the internally induced variability in the real climate system. A few examples can be cited in this context. Hind et al. (2012) consider a statistical framework developed by Sundberg et al. (2012) that allows to rank the simulations according to the "goodness-of-fit" of a forced simulation to the proxy/instrumental series, using the unforced signal (i.e. the control run simulation) as the reference. With this method they test whether different levels of amplitude of the solar forcing signal are detectable in a pseudoreconstruction built upon pseudoproxies from the EC5MP (Jungclaus et al., 2010) From their results, they extend their conclusions to future model-data comparison exercises, and they suggest the need of improving the signal to noise ratio in the proxies (i.e. their quality), more than the amount of proxy series to better distinguish the forced multi-decadal temperature signal. In turn, Bothe et al. (2013a) analyzed, under the paradigm of a statistically indistinguishable ensemble, the model-data consistency for the EC5MP and some regional and large-scale reconstructions. Their study revealed a limited consistency, found only over some regions and some periods within the LM such as the Central Europe during the last 500 years. Interestingly, if this analysis is extended to an ensemble of multimodel simulations (Bothe et al., 2013b), the level of low consistency found is similar to the case of single-model ensemble simulations. This lack of consistency stresses the necessity, as in Hind et al. (2012), to improve simulations and reconstructions for more accurate model-data comparisons in future exercises. Other analyses are directly focused on detecting the forced signal in the reconstructions, considered as the representation of the real climate system. For example, Schurer et al. (2013) analysed through a detection/attribution exercise, the role of the internally induced and the externally forced response of temperature in a multi-model simulation ensemble as well as in some NH reconstructions. A large influence of the external forcing in simulations and reconstructions is found since 1400 CE, while the internal variability seems to play a major role before. The alternative hypothesis deals with the idea of limitations in simulations, in reconstructions or in both. Fernández-Donado et al. (2013) offered also an analysis at hemispheric scales in which the linear response to changes in external forcing in simulations was explored and compared to the equivalent signal in reconstructions. Based on the linear relation between forcing and response supported by multi-model simulations and reconstructions, a simple metric to assess the amplitude of the response to the forcing over the LM was developed to assess the model-data consistency. Recently,

the multi-model-data comparisons have also been extended to continental scales. Thus, PAGES2k-PMIP3 group (2015) and EuroMed2k Consortium (2015) offer relevant insights on the larger influence of internal variability when considering continental and regional spatial scales. Specifically, the availability of a gridded european temperature reconstruction (EuroMed2k Consortium, 2015) allows to identify the subregions with a potential larger influence of the internal variability.

So far, the most recent and exhaustive assessment of the model-data consistency involving global, hemispheric and continental scales can be found in Masson-Delmotte et al. (2013). This Thesis builds on the analysis on the evolution of temperature during the CE and the discussions therein.

1.5 Main objectives and structure of the thesis

Estimating and understanding the past variations of the temperature from global to continental scales greatly benefits drawing inferences about the expected evolution of the temperature in the future. The temperature variability at multi-decadal and longer timescales is the result of the external forcing imposed to the climate system combined with the comparatively less well-known role of the internal variability. Disentangling the relative contributions of the externally forced and internally induced temperature variability is the basis on which simulations and reconstructions of last millennia can be compared and allows constraining the uncertainty of the climate in the future.

Therefore, the main goal of this Thesis is to provide an analysis of the role of the externally forced temperature response relative to the signature of the internal variability embedded in the evolution of the large-scale temperature during the CE. Such analysis is grounded on the evidences yielded by a comprehensive set of general circulation model simulations and proxy-based climate reconstructions.

The simulations, their external forcings drivers and the reconstructions are subject to uncertainty. The spread of temperature estimations throughout the CE can be largely considered a measure of the impact of the uncertainties affecting simulations and reconstructions. Therefore, understanding the sources of the respective uncertainties is pivotal if simulated and reconstructed evidences are to be compared. Examining the sources of uncertainty that prevent from straight forward comparisons between reconstructed and simulated trajectories of temperature during the last millennia is hence a major endeavor of this work.

The central scope of this study can be segregated in the following specific objectives:

This Thesis provides a compilation of AOGCM/ESM LM simulations and global, hemispheric and continental temperature reconstructions that constitute the most

comprehensive assemblage of simulated and proxy-reconstructed evidences up to date. The model simulations pool collected herein embraces most of the experiments published within the CMIP5/PMIP3 community for the LM as well as the available simulations that has been developed aside this coordinated effort. Indeed, the set of simulations presented in this manuscript represents the state-of-the-art compilation of model experiments driven by estimations of changes in natural and anthropogenic forcings. In addition, the assortment of global and hemispheric temperature reconstructions, as well as the ensemble of continental-scale reconstructions from the Pages2k project, integrates an extensive compendium of temperature evidences based on proxy reconstructions during the CE. The vast collection of reconstructions also encompasses a variety of methodological approaches and proxy and geographical diversity available up to date. Together, the two sets of temperature data used in this Thesis thoroughly account for the-state-of-the-art knowledge on temperature variations during the last two millennia based on simulations and reconstructions. Detailed information about the datasets that will serve as the basis of the analyses in this Thesis is provided Chapter 2.

A detailed description and discussion of all existing possibilities for the external forcing configurations applied to the AOGCM/ESM is provided in this Thesis. Such precise survey of current forcing conditions and a discussion about their historical usage is unique in the scientific literature. The uncertainty related to each type of forcing, natural or anthropogenic, is explored taking into account all alternative boundary conditions that can be implemented within models. In addition, an estimate of the total external forcing for each simulation that includes all individual forcing contributors is calculated herein. The quantification of the response to external forcing from simulations and also reconstructions in the following chapters lies upon this estimation of the total external forcing. Chapter 3 gathers this discussion about the external forcings of the climate system.

Once the simulations, the reconstructions and the external forcing configurations considered in this work have been introduced in Chapters 2 and 3, subsequent chapters are devoted to analyse the temperature response of both simulations and reconstructions and to link them with the external forcing.

The global and hemispheric temperature response of model simulations to the external forcing is examined and quantitave estimates of such a response are provided at multidecadal and above timescales. A calculation of the rate of temperature variations relative to unit changes in external forcing implies thus a measure of the climate sensitivity to changes in boundary conditions of the cli-

mate system and allows evaluating to what extent the simulated temperature response is linearly related to the total variation of the external forcing during the CE. Moreover, the previous quantification permits a comparison to other estimates of climate sensitivity that include conversely delayed responses in the system or other feedbacks (e.g. ECS). The study of the relation between simulations response and forcing conditions is convened in Chapter 4 of this book.

The evaluation provided of the strength of the connection between the reconstructed temperature and the known external forcing determines a very useful framework to assess the consistency between model simulations and reconstructions of the CE. Indeed, a comparison of both is solely possible in the basis of common responses to the external forcing, since the internal variability contributes to dampen the forcing signal to noise ratio. To what extent the global and hemispheric temperature reconstructions respond the the forcing signal at the different timescales is largely connected to several methodological issues and subjective decisions in the process of reconstructing temperature or other climate variables. The methodological variability evidences indeed an influence on the spread of the reconstructions ensemble and therefore also on model-reconstruction comparisons. Thus, Chapter 5 provides complementary assessement to the previous Chapter 4 by presenting an analysis of the reconstructed temperature response to the external forcing at hemispherical and global scales.

This work provides a thorough evaluation of the coherence of simulated and reconstructed temperature variations during the last millennium. Robust features between the two independent sources of paleoclimatic evidences contribute to reinforce the state of knowledge about last millennium temperature variations, while weaknesses arising from the comparison help identifying the sources of uncertainty that will require further efforts from the community in the future. Several approaches are worth to be explored in the present comparison as, for instance the estimates of temperature response to external forcing changes in simulations and reconstructions that are inter-compared at this stage. The latter establishes a suitable metric for a quantitative assessment of consistency between reconstructions and simulations. Additionally, the temperature changes and amplitude as well as their spatial distribution during the well-known MCA and LIA periods are examined in order to gain some insight into the role of the external forcing and the relative contribution of internal variability that could give raise to significant anomalies of temperature during these periods of the LM. This model-data comparison analysis is reported in Chapter 6.

The question whether to what extent the signal to noise ratio decreases as the

contribution of internal variability increases at more regional spatial scales is addressed in this work. Consequently, the model-data consistency explored in the previous chapter at global and hemispheric scales is extended herein to continental scales. The analysis proposed at this stage illustrates a larger influence of the internal variability as a result of an increased effect of regional climate dynamics and other non-climatic features at smaller spatial scales. The influence of the external forcing is analysed at different regions of both hemispheres. Additionally, the different levels of spatial homogeneity evidenced by simulated and reconstructed temperature are examined. The analysis of consistency between simulations and reconstructions at the continental scale is provided in Chapter 7.

A final account of the main conclusions, including a discussion about the open questions is developed in Chapter 8.

2

Data: simulations and reconstructions of the Common Era*

This chapter describes the database of AOGCM/ESM climate model simulations and reconstructions that will be used in this work. The collection of simulations and reconstructions considered corresponds to the most updated suite of published experiments, thus updating Masson-Delmotte et al. (2013) and Fernández-Donado et al. (2013). Section 2.1 focuses on describing the ensemble of AOGCMs and simulations while Section 2.2 presents the collection of temperature reconstructions considered. This chapter is complemented with Chapter 3 where a detailed description of the various external forcing factors and configurations used as boundary conditions in AOGCMs/ESMs are analysed.

2.1 Models and simulations

The present work considers a total of 37 forced climate simulations of the LM produced with 15 different AOGCMs/ESMs. Table 2.1 shows the list of models and experiments considered, indicating the name of the model and the acronym, whether it varies from the original name, used herein. A general description of each simulation, including horizontal resolution, number of atmospheric and oceanic levels, the set of external forcings considered, the exact period of simulation and the original references is also provided.

* The main contents of this chapter are included in:
- Fernández-Donado, L., J. F. González-Rouco, C. C. Raible, C. M. Ammann, D. Barriopedro, E. García-Bustamante, J. H. Jungclaus, S. J. Lorenz, J. Luterbacher, S. J. Phipps, J. Servonnat, D. Swingedouw, S. F. B. Tett, S. Wagner, P. Yiou and E. Zorita, 2013: Large-scale temperature response to external forcing in simulations and reconstructions of the last millennium. *Climate of the Past*, **9**, 393-421. DOI 10.5194/cp-9-393-2013.

Table 2.1: Description of the climate models and experiments considered in this Thesis: institutional model acronym and short acronym used herein, if different from the official one (column 1); resolution and number of levels in their atmospheric (column 2) and oceanic (column 3) components; set of external forcing factors considered in the experiment configuration (column 4); period of simulation (column 5); and the original reference describing the experiments (column 6). Legend for external forcing factors considered in column 4: (S) solar forcing using stronger changes in amplitude (i.e. larger than 0.20 % TSI change since LMM to present); (ss), solar focing using weaker changes in amplitude (i.e. lower than 0.1 % TSI change since LMM to present; (V) volcanic activity; (G) greenhouse gases; (A) anthropogenic aerosols; (L) land-use changes; and (O) orbital variations. Experiments are classified as PMIP3 if they follow the community protocols for the external forcing configurations (Schmidt et al., 2011) and nonPMIP3 otherwise.

Model name (short acronym)	**Atmosphere [Resolution/vertical levels]**	**Ocean [Resolution/vertical levels]**	**Forcings**	**(n° runs) length CE**	**Reference**
			PMIP3		
BCC-csm1-1 (BCC)	T42/26	1x1/40	ssVGAO	(1) 850-2005	–
CCSM4	0.9x1.25/27	1x1/60	ssVGALO	(1) 850-2004	Landrum et al. (2013)
CSIRO-MK3L-1-2 (CSIRO)	R21/18	2.8x1.6/21	ssVGO	(1) 851-2000	Phipps et al. (2013)
GISS-E2-R (GISS)	2x2.5/40	1x1.25/32	ssVGALO ssVGALO	(1) 850-2004 (2) 850-2004	Schmidt et al. (2014)
HadCM3	3.75x2.5/19	1.25x1.25/20	ssVGALO	(1) 800-2000	Schurer et al. (2013)
IPSL-CM5A-LR (IPSLCM5)	1.9x3.75/39	2x2/31	ssVGO	(1) 850-2005	Dufresne et al. (2013)
MPI-ESM-P (MPI)	T63/47	1.5x1.5/40	ssVGALO	(1) 850-2005	Jungclaus et al. (2014)
*CESM(CAM5)	2x2/30	1x1/60	ssVGALO	(10) 850-2005	Otto-Bliesner et al. (2015)
			nonPMIP3		
CCSM3	T31/18	3.6x2.8/25	SVG	(1) 1000-2100 (4) 1500-2100	Hofer et al. (2011)
CNRM-CM3.3 (CNRM)	T42/31	2x2/31	SVGAL	(1) 1001-1999	Swingedouw et al. (2010)
CSIRO-MK3L-1-2 (CSIRO-pre)	R21/18	2.8x1.6/21	ssGO ssVGO	(3) 1-2000 (3) 501-2000	Phipps et al. (2012)
CSM1.4	T31/18	3.6x1.8/25	SVGA	(1) 850-1999	Ammann et al. (2007)
ECHAM5/MPIOM (EC5MP)	T31/19	3x3/40	ssVGALO SVGALO	E1 (5) 800-2000 E2 (3) 800-2000	Jungclaus et al. (2010)
ECHO-G	T30/19	2.8x2.8/20	SVG SGO	(2) 1000-1990 (1) 8000 BP - 1998	González-Rouco et al. (2006) Wagner et al. (2007)
HadCM3 (HadCM3-pre)	3.75x2.5/19	1.25x1.25/20	SVGALO	(1) 1492-1999	Tett et al. (2007)
IPSLCM4	3.75x2.5/19	2x2/31	SGAO	(1) 1001-2000	Servonnat et al. (2010)
FGOALS-gl (FGOALS)	5x4/26	1x1/30	SVG	(1) 1000-1999	Zhou et al. (2011)
CESM1.0 (CESM)	0.9x1.25/27	1x1/60	SVGAL	(1) 850-2099	Lehner et al. (2015)

All the model experiments considered herein, except the CESM simulation, that was published recently (Lehner et al., 2015), are included within the most recent IPCC assessment report (Masson-Delmotte et al., 2013). These simulations constitute the actual playground for analysing the CE, and more specificaly the LM, climate. The reader is addressed to the original references in Table 2.1 for a more in depth description of the models and configurations applied. Nine out of the fifteen models are effectively different AOGCMs, whereas the ECHO-G/EC5MP, the CCSM1.4/CCSM3 and the IPSLCM4 are earlier versions of the MPI, CCSM4 and IPSLCM5 models, respectively, and the CESM is a version of the CCSM4 but including a carbon cycle module (Lehner et al., 2015; Hurrell et al., 2013). Recently an ensemble of 30 experiments developed with the CESM(CAM5) with various forcing configurations (10 out of the 30 considering a full forcing scenario) has been published (Otto-Bliesner et al., 2015). These simulations are not included in this work. However, previous works from Lehner et al. (2015) and Landrum et al. (2013) stand as representative experiments with this model. Indeed the new experiments are developed with a similar model configuration to the one used by Lehner et al. (2015), only differing in the atmospheric model by applying the most recent version CAM5 instead of CAM4 but with a coarser spatial resolution ($\sim 2°$ instead of $\sim 1°$ to reduce the computational demand).

Table 2.1 presents the different horizontal and vertical resolutions of the atmospheric and oceanic subsystems from each model, illustrating as well he range of spatial resolutions included. CSIRO presents the coarsest atmospheric resolution with 64 longitudes and 56 latitudes (2.8°x1.6°) in contrast to the finest, CCSM4, with 288 and 192 (0.9°x1.25°), respectively. The various models could also be characterized attending to the different sensitivity evidenced to changes in external forcing through their corresponding values of the Equilibrium Climate Sensitivity (ECS) and Transient Climate Response (TCR), indicated in Table 2.2. ECS assesses the temperature change, after reaching the equilibrium, due to a sudden doubling of CO_2 above pre-industrial levels (Schneider et al., 1980), while the TCR focuses on the temperature change at the time of atmospheric CO_2 doubling under an idealized scenario of CO_2 increase by 1% each year (Knutti et al., 2005). Models considered within this work present a range of temperature increase from 2 to 4.1 (°C) in the case of ECS, and from 1.2 to 2.3 (°C) for the TCR. These values will be further discussed in Chapters 4 and 6.

The total of the 37 simulations considered herein have been developed since 2003 (González-Rouco et al., 2003b) and constitute the complete suite of currently published AOGCM simulations for the CE, thereby updating the assessment by Masson-Delmotte et al. (2013). Since these simulations were produced

with different AOGCMs, the total ensemble considered is therefore heterogeneous in terms of forcing configurations and initial conditions. Further and most importantly, different external forcing boundary conditions were used, depending on the institutions that carried out the simulations and successive updates of the forcing estimates that progressively became available (see Chapter 3). Indeed, the total ensemble of simulations includes experiments developed following the CMIP5-PMIP3 protocol (Taylor et al. 2012; Schmidt et al. 2011, 2012; PMIP3 experiments hereafter) but also a large amount of runs developed aside this project, the so-called nonPMIP3 experiments hereafter (see Table 2.1). The main difference between these two sub-ensembles of simulations is the existence of a coordinated protocol for the external forcing configurations in the case of PMIP3 (Schmidt et al., 2011, 2012) experiments. However, the nonPMIP3 group is characterized by a wide diversity of forcing boundary conditions (Chapter 3). This heterogeneity in the simulations and forcings will allow exploring a larger spectrum of plausible scenarios for the last millenium and, in some instances, to assess the degree to which the agreement between simulated and reconstructed responses can improve or deteriorate by considering different specifications of a given forcing.

The PMIP3 simulations are part of the *past1000 tier* (Taylor et al., 2012) in which the time span is defined from 850 to 1850 CE. Herein the corresponding industrial extensions of each run are included, reaching thus, at least, the year 2000 CE. The nonPMIP3 runs show different time spans, most of them covering the complete LM, i.e. 1000 - 2000 CE. Despite this existing variety of time intervals (Table 2.1), most of the analysis carried out within the present manuscript focuses on the last 12 centuries (800 - 2000 CE). Thus, the shorter simulations (the 550 yr HadCM3 and CCSM3 runs) are used only in some parts of this study while the longer runs (the 8 kyr ECHO-G and the 2 kyr CSIRO simulations) are only considered within the time interval of the PMIP3 experiments, i.e. after 800 CE.

In addition, in some parts of the work only the simulations including a minimum set of external forcing factors such as solar, volcanic and greenhouse gases concentrations are considered, as it will be stated. Note also that the series considered through the manuscript are usually 31-yr low-pass filter outputs since the focus of the present work is on timescales above multidecadal variability where the response to the external forcing becomes more detectable.

2.2 Reconstructions

This section provides information about the ensemble of temperature reconstructions that will be used through the text.

Table 2.2: Equilibrium Climate Sensitivity (ECS) and Transient Climate Response (TCR) estimates for the AOGCMs/ESMs listed in Table 2.1. The equivalent increase in temperature degrees for a doubling of CO_2 is shown between square brackets. Values were extracted from references in Table 2.1, Solomon et al. (2007) and Flato et al. (2013). The conversion to $^{\circ}C/2 \times CO_2$ was done following Houghton et al. (2001). ECS value for CESM is not available.

Model	**ECS** $^{\circ}C/Wm^{-2}$ [$^{\circ}C/(2 \times CO_2)$]	**TCR**
BCC	0.76 [2.82]	0.46 [1.71]
CCSM4	0.86 [3.20]	0.46 [1.82]
CSIRO	1.08 [4.00]	0.40 [1.50]
GISS	0.57 [2.11]	0.40 [1.50]
HadCM3	0.80 [2.97]	0.54 [2.00]
IPSLCM5	1.11 [4.12]	0.54 [2.00]
MPI	0.95 [3.52]	0.54 [2.00]
CCSM3	0.62 [2.30]	0.38 [1.41]
CNRM	0.59 [2.19]	0.43 [1.59]
CSM1.4	0.54 [2.00]	0.39 [1.45]
EC5MP	0.92 [3.41]	0.59 [1.19]
ECHO-G	0.86 [3.19]	0.46 [1.71]
IPSLCM4	1.02 [3.78]	0.57 [2.11]
FGOALS	0.99 [3.70]	0.38 [1.41]
CESM	0.86 [3.20]	0.46 [1.72]
CESM(CAM5)	1.10 [4.10]	0.62 [2.30]

The analyses developed in Chapters 5 and 6 focus on global and hemispheric scales, thus considering temperature reconstructions at these spatial scales. These series are described in Section 2.2.1. In turn, Section 2.2.2 briefly describes the existing continental-scale temperature reconstructions that will be examined in Chapter 7.

2.2.1 Large scale reconstructions

The set of hemispheric and global reconstructions considered herein is presented in Table 2.3. This suite of reconstructions represents the state-of-the-art of available large-scale temperature reconstructions in the CE, in a similar manner that

was done in Masson-Delmotte et al. (2013) and presented in Figure 1.2c, but including some additional series. The latter provides a representation of the different existing multi proxy sources and the various reconstruction methods applied. The present ensemble includes a total of 16 reconstructions for the NH, 6 for the SH and 4 for the global (GL) scale.

If the original reconstruction presented minimal methodological variants or alternative but very similar instrumental dataset, only the most recent version was selected. This follows the same convention as Masson-Delmotte et al. (2013) and Fernández-Donado et al. (2013). Some general characteristics of the reconstructions selected are summarized in Table 2.3, such as the time span (column 2), the proxy-type (column 3) and the target region (column 4), although for further details the reader is addressed to the original reference of each reconstruction (column 1). All reconstructions in Table 2.3 have a minimum length of four centuries and in some cases start well before 800 CE, even spanning the whole CE (Moberg et al., 2005; Ljungqvist, 2010; Christiansen and Ljungqvist, 2012b). The time resolutions are annual for all reconstructions except for Ljungqvist (2010), that has decadal resolution. Even if records have annual resolution, this does notnecessarily imply that the real temporal resolution is annual; for instance this is the case of the GL and SH borehole data (Huang et al., 2000) that provide information only on multicentennial trends. This also occurs with some other reconstructions that show low variance on interannual timescales, despite the data was originally provided at annual resolution (Briffa et al., 2001; Hegerl et al., 2007b; Loehle, 2007; Mann et al., 2009; Leclercq and Oerlemans, 2012; Christiansen and Ljungqvist, 2011). Additionally, some of the reconstructions are tree-ring based and thus potentially biased to the growing season (Jones et al., 2009). Nevertheless, all the reconstructions included herein purport to represent annual variations except for Briffa et al. (2001), that is presented as a summer temperature reconstruction. Within each large-scale domain (i.e. GL, NH, SH), different regions are distinguished, according to the spatial target of the series, depending whether they include or not oceanic areas or they just focus on extratropical regions. Thus within the GL and SH, reconstructions targeting only the continental regions, i.e. GL-LO and SH-LO, will be distinguished from those including also the oceans (i.e. the complete spatial domain, GL-C and SH-C). In turn within the NH, apart from the complete (NH-C) and the land only (NH-LO) domains, the extratropical land only target will be considered (NH-LOE).

One of the GL reconstructions listed in Table 2.3, the so-called "hybrid", has been built within this study as a simple average from the NH series from Frank et al. (2010b) and the SH Neukom et al. (2014) reconstruction. Frank et al. (2010b) did not really present a single new NH reconstruction, but an ensemble based on nine different reconstructions and their respective variants when mod-

Table 2.3: Temperature reconstructions used in this work. The original publication (column 1), the time span (column 2), the proxy-type (column 3) and the target region (column 4) are specified for each record. A variety of regional targets are identified for each large-scale region, corresponding to the complete region (GL-C, NH-C and SH-C, land and ocean regions over all latitudes), continental or land only regions (GL-LO, NH-LO and SH-LO), and extratropical continental regions (NH-LOE). (*) This GL reconstruction is calculated in this study as an average from the NH Frank et al. (2010b) and SH Neukom et al. (2014) reconstructions.

Reconstruction	**Period (CE)**	**Proxy type**	**Target region**
NH:			
Ammann and Wahl (2007)	1000-1980	Multiproxy	NH-C
Briffa et al. (2001)	1402-1980	Tree rings	NH-LOE
Christiansen and Ljungqvist (2012b)	1-1973	Multiproxy	NH-C
D'Arrigo et al. (2006)	713-1995	Tree rings	NH-LOE
Frank et al. (2007)	831-1992	Tree rings	NH-LOE
Hegerl et al. (2007b)	1251-1960	Tree rings	NH-LOE
Huang (2004)	1500-1980	Boreholes	NH-LO
Juckes et al. (2007)	1000-2000	Multiproxy	NH-C
Leclercq and Oerlemans (2012)	1600-2000	Glaciar length fluctuations	NH-LO
Ljungqvist (2010)	1-1999	Multiproxy	NH-C
Loehle and McCulloch (2008)	16-1935	Multiproxy (non tree-rings)	NHC
Mann et al. (2008) – CPS	1500-1995	Multiproxy	NH-C
Mann et al. (2008) – EIV	299-2005	Multiproxy	NH-C
Mann et al. (2009) – CFR	500-2006	Multiproxy	NH-C
Moberg et al. (2005)	1-1979	Multiproxy	NH-C
Shi et al. (2013)	1000-1998	Multiproxy	NH-LO
SH:			
Huang et al. (2000)	1500-2000	Boreholes	SH-LO
Jones et al. (1998)	1000-1991	Multiproxy	SH-C
Leclercq and Oerlemans (2012)	1600-2000	Glaciar length fluctuations	SH-LO
Mann et al. (2008) – CPS	1000-1995	Multiproxy	SH-C
Mann et al. (2008) – EIV	399-2005	Multiproxy	SH-C
Neukom et al. (2014)	1000-2000	Multiproxy	SH-C
GL:			
Hybrid reconstruction*	1000-1995	Multiproxy	GL-C
Huang et al. (2000)	1500-2000	Boreholes	GL-LO
Leclercq and Oerlemans (2012)	1600-2000	Glaciar length fluctuations	GL-LO
Mann et al. (2008) – EIV	499-2005	Multiproxy	GL-C

ifying the calibration period or the spatial target. The mean of the Frank et al. (2010b) ensemble is presented in Neukom et al. (2014) as a representation of the NH reconstructed temperature evolution for comparison to the SH. This ensemble mean is used here, together with the new SH temperature reconstruction of Neukom et al. (2014) in order to produce a new GL "hybrid" reconstruction. This "hybrid" reconstruction presents new evidence for global scales, a domain that, indeed suffers from a lack of temperature reconstructions.

2.2.2 Continental scale reconstructions

The suite of continental scale temperature reconstructions considered for the present work is presented in Table 2.4. These continental reconstructions have been developed within the PAGES "2k Network" project[†] with the aim of producing a global dataset of regional climate reconstructions that would span as much as possible over the CE. Some general remarks about these reconstructions are presented below, nevertheless the reader is guided to PAGES 2k Consortium (2013) and Kaufman et al. (2014) for an in-depth description of the series.

The set of reconstructions includes eight different PAGES2k series targeting seven continental-scale regions, i.e. Arctic, Antarctica, Asia, Australasia, Europe, North America (2 series) and South America, spanning at least the last 800 years (Table 2.4). The series reconstruct annual or warm-season temperature variability within each continental scale region (Table 2.4). All regions present annual resolution except North America, which has a tree-ring-based reconstruction with 10 year resolution and a pollen-based reconstruction with 30 year resolution (PAGES2k-PMIP3 group, 2015).

The selection of the proxies used for each continental reconstruction, as well as the choice of the methodology and regional calibration targets was based on expert criteria (PAGES 2k Consortium, 2013). Thus, as in the case of the large-scale reconstructions, different methods have been applied to develop these continental series. Nevertheless, alternative reconstructions for all regions based on exactly the same statistical procedures were also produced and were found to compare reasonably well with the original PAGES2k temperature reconstructions provided by each group (PAGES 2k Consortium, 2013).

2.3 Conclusions

The extensive ensemble of simulations and reconstructions that will be considered for this work has been presented in this chapter. AOGCM/ESM experiments

[†] www.pages-igbp.org/workinggroups/2k-network

Table 2.4: PAGES2k Continental scale temperature reconstructions used in this work. For each record the reconstructed area (column 1), the time span (column 2), the proxy-type (column 3), the temporal resolution (column 4) and the spatial target region (column 5) are specified. See PAGES 2k Consortium (2013) for further details. Specifically the reader is guided to McKay and Kaufman (2014) for further information about the Arctic reconstruction and to EuroMed2k Consortium (2015) for the Europe reconstruction.

Reconstruction	**Period**	**Proxy type**	**Resolution**	**Target region**
Arctic	1-2000	Multiproxy	Mean annual	Land and Ocean
Europe	1-2003	Multiproxy	Summer (JJA)	Land only
Asia	800-1989	Tree rings	Summer (JJA)	Land only -
North America	1204-1974	Tree rings	10-yr, mean annual	Land and Ocean
	480-1950	Pollen	30-yr, mean annual	Land and Ocean
South America	857-1995	Multiproxy	Summer (DJF)	Land only
Australasia	1001-2001	Multiproxy	Warm season (Sept-Feb)	Land and Ocean
Antarctica	167-2005	Ice cores	Mean annual	Land only

developed within the PMIP3 common effort are included together with those developed aside this project (nonPMIP3), conforming a total of 37 transient simulations. The suite of simulations is heterogeneous in terms of the climate model used, spatial resolution, time interval spanned or the external forcing configuration applied. Thus, a comprehensive collection of high complexity climate simulations is considered for the analys of the temperature variability over the CE.

In turn, the collection of hemispheric and global temperature reconstructions considered include a total of 16 series for the NH, 6 for the SH and 4 for the GL scales, representing the most updated suite of large-scale temperature reconstructions. A new GL temperature reconstruction has been developed in this Thesis based on two relatively recent hemispheric reconstructions. Apart from the new GL reconstruction developed herein, a new SH reconstruction (Neukom et al., 2014) that does not appear in AR5 has been also included. Another difference from the AR5 selection is related to the choie of the Huang (2004) borehole reconstruction instead of the analogous reconstruction of Pollack and Smerdon (2004) since the latter includes only low frequency information. Additionally, from the several versions of the Mann et al. (2008) reconstruction, only the land and ocean variants have been considered since they provide a better representation of the location of the proxies used. In addition, the PAGES2k continental scale recon-

structions has been introduced for the following temperature analysis developed at sub-hemispheric or continental scales.

3

The external forcing of the climate system*

LM climate simulations may embrace different sets of forcing factors depending on the model considered and also depending on the particular realization of the same model. Additionally, there are alternative estimates of the temporal evolution of these forcings used as boundary conditions to reproduce realistically the states of the climate system in the past. While all simulations incorporate solar and greenhouse gas (GHG) forcing and most of them consider volcanic forcing (except for the IPSLCM4, one ECHO-G and some CSIRO simulations, see Table 2.1), only some of them introduce anthropogenic aerosols (BCC, CCSM4, GISS, MPI, CSM1.4, EC5MP, CNRM, IPSLCM4 and HadCM3) and land use changes (CCSM4, GISS, MPI, EC5MP, CNRM and HadCM3). This variety of configurations allows to some extent to explore the uncertainty stemming from our lack of robust knowledge about the past evolution of some of the forcing factors.

This chapter illustrates and compares the main differences between the various existing forcing estimations that have been considered by the suite of simulations (Table 2.1). For more in-depth details of the natural and anthropogenic forcings applied the reader is addressed to the original references in Table 3.1. The text will be organized herein into natural (Section 3.1) and anthropogenic (Section 3.2) forcing factors. Section 3.1 will describe thus the various scenarios of solar

* The main contents of this chapter are included in:
- Fernández-Donado, L., J. F. González-Rouco, C. C. Raible, C. M. Ammann, D. Barriopedro, E. García-Bustamante, J. H. Jungclaus, S. J. Lorenz, J. Luterbacher, S. J. Phipps, J. Servonnat, D. Swingedouw, S. F. B. Tett, S. Wagner, P. Yiou and E. Zorita, 2013: Large-scale temperature response to external forcing in simulations and reconstructions of the last millennium. *Climate of the Past*, **9**, 393-421. DOI 10.5194/cp-9-393-2013.

variability, volcanic activity and orbital variations, while Section 3.2 will focus on the GHG, anthopogenic aerosols and land use changes. GHGs will be included in the group of anthropogenic forcings due to their gradient during the industrial period, even if most contributions to their global variability before 1850 CE may be of natural origin. The same exception applies to land use. Additionally, a total external forcing expressed in radiative forcing units has been obtained for each model by integrating all natural and anthropogenic contributions for the purpose of a better comparison among the various forcing configurations, and also with simulations and reconstructions (Section 3.3). Section 3.4 summarizes the main results of this chapter.

Note that most of the external forcing factors explored will be divided, for the sake of clarity, between PMIP3 and nonPMIP3 (Chapter 2), referring to those experiments following the Schmidt et al. (2011, 2012) external forcing criteria and those developed aside these common criteria, respectively (Fernández-Donado et al., 2013). In addition, it will be shown that the estimations of past Total Solar Irradiance (TSI) can be clustered in two groups, which assume a weak (ss) or strong (S) amplitude of variations, respectively. This classification is extended to the ensemble of simulations (Section 3.1) depending on the solar variability scenario they consider. An example of the latter are the eight EC5MP (nonPMIP3) simulations that group into two sub-ensembles (EC5MP-E2 and EC5MP-E1, see Table 2.1) produced following comparatively stronger and weaker changes in TSI through the LM, respectively (Jungclaus et al., 2010, see Section 3.1). Overall, it can be said that most of the nonPMIP3 experiments consider strong amplitude variations, while all the PMIP3 simulations stand for the low solar variability scenario (see Table 2.1).

3.1 Natural forcing

Solar irradiance changes can play a major role in forcing decadal to centennial climate variability through the CE (e.g. Crowley, 2000; Zorita et al., 2005). The amplitude of its variations is nowadays estimated to be much smaller (e.g. Lean et al., 2002; Foukal et al., 2004; Solanki and Krivova, 2004; Wang et al., 2005; Krivova et al., 2007; Gray et al., 2010) than previous published estimates (e.g. Hoyt and Schatten, 1993; Lean et al., 1995; Bard et al., 2000). Yet, a relatively recent reconstruction (Shapiro et al., 2011) still endorses large background variations in irradiance (see discussion in Schmidt et al., 2011, 2012; Feulner, 2011; Masson-Delmotte et al., 2013). Thus, the solar variability amplitude is still under debate in the community.

Table 3.1: Reconstructions of natural and anthropogenic forcings applied to each model in Table 2.1.

Model	Solar	Volcanic	CO_2	CH_4	N_2O	Aerosols	Land use	Orbital
				PMIP3				
BCC	VS10+W05	Gao08	McFM06	McFM06	Fluck02	Lam10	—	Berg78
CCSM4	VS10+W05	Gao08	McFM06	McFM06	Fluck02	Lam10	Pong08+H11	Berg78
CSIRO	Stein09+W05	Crow12	McFM06	McFM06	Fluck02	—	—	Berg78
GISS	Stein09+W05 VS10+W05	Crow12	McFM06	McFM06	Fluck02	Lam10	Pong08+H11	Berg78
HadCM3	Stein09+W05	Crow12	Johns03	Johns03	Johns03	Johns03	Pong08	Berg78
IPSLCM5	VS10+W05	Amma03	McFM06	McFM06	Fluck02	—	—	Berg78
MPI	VS10+W05	Crow12	McFM06	McFM06	Fluck02	Lam10	Pong08	Berg78
				nonPMIP3				
CCSM3	Bard00+Lean95	Amma03	Ether96	Blun95	Fluck02	—	—	—
CNRM	Bard00+Lean95	Amma03	McFM06	Blun95	Fluck02	Bouch02	R99	—
CSIRO-pre	Stein09	Gao08	McFM06	McFM06	McFM06	—	—	Berg78
CSM1.4	Bard00	Amma03	Ether96	Blun95	Fluck02	Joos01	—	—
EC5MP	E1-Kriv07 E2-Bard00	Crow08+12	Dg.+Marl03	McFM06	McFM06	Lefohn99	Pong08	Breta88
ECHO-G	Bard00+Lean95	Crow00	Ether96	Ether98	Batt96	—	—	—
HadCM3-pre	Bard00+Lean95	Crow03	Johns03	Johns03	Johns03	Johns03	W85+R99+G01	Berg78
IPSLCM4	Bard00+Lean95	—	McFM06	Blun95	Fluck02	Bouch02	—	Lask04
FGOALS	Bard00+Lean95	Crow00	Ether96	Blun95	Fluck02	—	—	—
CESM	[VS10+W05]*	Gao08	McFM06	McFM06	Fluck02	Lam10	Pong08+H11	Berg78

Key for labels:

Amma03 (Ammann et al., 2003) **Crow08** (Crowley et al., 2008) **H11** (Hurtt et al., 2011) **Marl03** (Marland et al., 2003)
Bard00 (Bard et al., 2000) **Crow12** (Crowley and Unterman, 2012) **Johns03** (Johns et al., 2003) **McFM06** (MacFarling Meure et al., 2006)
Batt96 (Battle et al., 1996) **Dg.** (Diagnosed by the model) **Joos01** (Joos et al., 2001) **Pong08** (Pongratz et al., 2008)
Berg78 (Berger, 1978) **Ether96** (Etheridge et al., 1996) **Kriv07** (Krivova et al., 2007) **R99** (Ramankutty and Foley, 1999)
Bouch02 (Boucher and Pham, 2002) **Ether98** (Etheridge et al., 1998) **Lam10** (Lamarque et al., 2010) **Stein09** (Steinhilber et al., 2009)
Blun95 (Blunier et al., 1995) **Fluck02** (Flückiger et al., 2002) **Lask04** (Laskar et al., 2004) **VS10** (Vieira and Solanki, 2010)
Breta88 (Bretagnon and Francou, 1988) **Gao08** (Gao et al., 2008) **Lean95** (Lean et al., 1995) **W05** (Wang et al., 2005)
Crow00 (Crowley, 2000) **G01** (Goldewijk, 2001) **Lefohn99** (Lefohn et al., 1999) **W85** (Wilson and Henderson-Sellers, 1985)
Crow03 (Crowley et al., 2003) [VS10+W05]* as VS10 and W05, but scaled to double the Maunder Minimum-Present Day amplitude.

Figure 3.1a,b presents the estimations of solar forcing evaluation used to drive the suite of simulations listed in Table 2.1 for the nonPMIP3 and PMIP3 experiments, respectively. Despite the suite of series proposed by Schmidt et al. (2011, 2012), the available PMIP3 experiments consider so far only two different estimations of solar variability forcing (Steinhilber et al., 2009; Vieira and Solanki, 2010). Figure 3.1b displays the two reconstructions; see Table 3.1 for the correspondence between these series and PMIP3 model simulations.

All scenarios agree in depicting higher irradiance values during the so called Medieval Maximum (ca. 1100-1250 CE) and between two minima in the 11th (Oort Minimum; 1010-1080 CE) and 13th (Wolf Minimum; ca. 1280-1350 CE) centuries. The lowest irradiance during the millennium is reconstructed during the 15th to 17th centuries in the Spörer (ca. 1460 to 1550 CE) and Maunder (ca. 1645 to 1715 CE) minima; the last interval with low reconstructed TSI being the Dalton minimum (ca. 1790 to 1820 CE). During the 19th and 20th centuries all models use irradiance values that are comparable in magnitude or higher than those of the Medieval Maximum. The various solar forcing scenarios group into two types, one involving TSI variations of comparatively larger amplitude through the LM (STSI hereafter; comprising the CCSM3, CNRM, CSM1.4, EC5MP-E2, ECHO-G, HadCM3-pre, IPSL4, CESM and FGOALS runs) and one involving changes of comparatively weaker amplitude (ssTSI hereafter; comprising the CSIRO-pre, EC5MP-E1 and the complete suite of PMIP3 runs). This is quantified in Table 3.2 where the percentage of TSI change between three key periods of the LM is provided: the interval of highest TSI during the Medieval Maximum (1130-1160 CE), the Late Maunder Minimum (LMM, 1680-1710 CE) and the late 20th century (1960-1990 CE). The STSI group clusters with an increase in TSI larger than 0.20 % from the LMM to present and a decrease of more than 0.13 % from the Medieval Maximum to the LMM. The EC5MP and ECHO-G show the largest percentage of change in the transition to the LMM-present as is also evidenced in Figure 3.1a. The ssTSI group, displays a maximum reduction of 0.06 % during the Medieval to LMM transition and an increase of less than 0.09 % from LMM to present.

Except for the CESM, that it will be commented later, the coherent evolution within the STSI solar forcing stems from the use of a single record of production rates of the cosmogenic isotope ^{10}Be in Antarctic ice cores from Bard et al. (2000). The CSM1.4 and the EC5MP-E2 ensemble use the original values provided by Bard et al. (2000) (note that series overlap in Figure 3.1a) and do not include estimations of the 11-yr solar cycle. In turn, the CCSM3, CNRM, ECHO-G, HadCM3-pre, FGOALS and IPSLCM4 simulations use a version of the Bard et al. (2000) record spliced by Crowley (2000) to a reconstruction of TSI (Lean et al., 1995) based on the sunspot record of solar activity since 1610 CE. Therefore, all these records include an estimate of the 11-yr solar cycle since this date. The slightly different forcings adopted by the various AOGCMs are due to different calibrations of the net radiative forcing data provided by Crowley (2000) to the original TSI values of Lean et al. (1995).

The TSI used by the CESM, although included within the STSI group due to the large percentage of TSI change between the key periods, presents a more similar evolution at high frequencies to the ssTSI group. The latter is based on

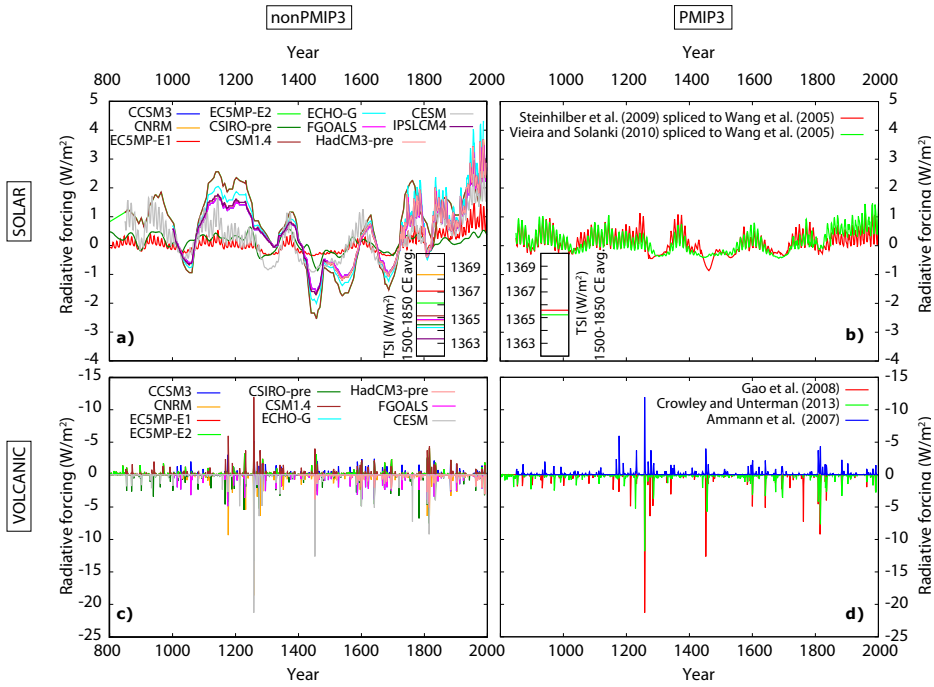

Fig. 3.1: Estimations of solar (top) and volcanic (bottom) external forcings used to drive the nonPMIP3 (left) and PMIP3 (right) simulations in Table 2.1. a) and b) TSI anomalies wrt 1500-1850 CE. The insets show the corresponding TSI averages for the reference period. c) and d) Volcanic forcing estimations in radiative forcing units. Note: the volcanic forcing is always negative. The orientation above or below the x axis is abitrary and used for clarity in the display.

the origin of the TSI forcing applied to the CESM which comes from one of the Schmidt et al. (2011) recommendations for the PMIP3 simulations (Vieira and Solanki, 2010) but scaled to double the Maunder Minimum-Present Day amplitude (Lehner et al., 2015). This forcing represents, thus, an intermediate state between the previous TSI forcing estimations and those of the PMIP3 generation.

Within the ssTSI group two clusters can be distinguished: some simulations following the PMIP3 guidelines (Schmidt et al., 2011) and the EC5MP-E1 and CSIRO-pre runs within the nonPMIP3 group. The EC5MP-E1 simulations use TSI sunspot-based reconstructions since 1610 (Krivova et al., 2007) spliced to

Table 3.2: Percentage of TSI change between the period with highest TSI values (1130-1160) during the Medieval Maximum and the LMM and also from LMM to late 20th century. Percentages are calculated with reference to the LMM average. The last two rows indicate the solar forcing reconstructions used by the PMIP3 experiments (see Table 3.1). Note: a 0.25 % change between the LMM and late 20th century is equivalent to a variation of the TSI between the two periods of $\sim$3.4 W/m^2.

Model	**Medieval Maximum-LMM (%)**	**LMM-late 20th (%)**
CCSM3/FGOALS	- 0.27	0.23
CESM	- 0.13	0.20
CNRM	- 0.18	0.25
CSIRO-pre	- 0.04	0.05
CSM1.4	- 0.17	0.24
EC5MP-E1	- 0.04	0.09
EC5MP-E2	- 0.27	0.24
ECHO-G	- 0.22	0.29
HadCM3-pre	–	0.25
IPSLCM4	- 0.18	0.25
Steinhilber	- 0.05	0.05
Vieira	- 0.06	0.09

records of ^{14}C isotope concentrations in tree-rings (Solanki et al., 2004; Usoskin et al., 2007; Krivova and Solanki, 2008). The reconstructed 11-yr cycle is extended before the 17th century by artificially superimposing the average 11-yr solar cycle between 1700 CE and present. The CSIRO-pre simulations use a ^{10}Be based reconstruction by Steinhilber et al. (2009) with no 11-yr cycle. None of the simulations consider stratospheric photochemistry nor ozone interactions (Shindell et al., 2001). Estimations of variability in solar wavelengths (Haigh et al., 2010) are neither included. In the case of PMIP3 simulations, all the experiments use the reconstruction from Wang et al. (2005) from 1850 CE based on sunspot data, spliced to two different alternatives, depending on the model, for the past variations. CSIRO, HadCM3 and one simulation from GISS consider, similar to the CSIRO-pre, the Steinhilber et al. (2009) reconstruction based on ^{10}Be, although including an 11-yr cycle. In turn, BCC, CCSM4, IPSLCM5, MPI and the other two simulations from GISS use the reconstruction of Vieira and Solanki (2010) based on records of ^{14}C isotope concentrations.

Figure 3.1a,b also shows the mean value of TSI for each reconstruction calculated within the reference interval 1500-1850 CE. These vary from ∼1362 to 1369 W/m^2. The range of average TSI values is bounded by the IPSLCM4 and CNRM models in the lower and upper limit, respectively, interestingly both of them using the same TSI reconstruction (Bard et al., 2000; Lean et al., 1995) and identical sequence of anomalies through the LM. For the case of the PMIP3 TSI forcing, the mean values in 3.1b, correspond to those of the original reconstructions, Steinhilber et al. (2009) and Vieira and Solanki (2010), although they could have been modified by each model when implementing these forcings. For the current study, however, the difference in TSI mean values between simulations is not expected to have an influence on the simulated climate evolution during the CE.

Volcanic forcing is included in all simulations except for the IPSLCM4, the 8000-yrs ECHO-G run and 3 simulations of the CSIRO-pre model (Table 2.1). The various estimations of volcanic forcing are shown in Figure 3.1c,d, in radiative forcing units, for the nonPMIP3 and PMIP3 simulations, respectively. CCSM3, CNRM and the PMIP3 experiments originally incorporate this forcing in aerosol optical depth values and their conversion into radiative forcing units has been done following Hansen et al. (2002), where a factor of -21W/m^2 is suggested for the conversion. This value is within the range of estimations made also by Wigley et al. (2005).

The reconstructions of stratospheric aerosols from volcanic eruptions are based on ice-core data from Antarctica and Greenland. Our knowledge of volcanic forcing over the CE has recently improved significantly according to Myhre et al. (2013), although there exists still some uncertainty in regard to the strengths of individual eruptions (Schmidt et al., 2011). Indeed, the derived time series of volcanic forcing tend to display consistent timing for major eruptions (Figure 3.1c,d), but they often present differences on the magnitudes of individual events, highlighting the largest values shown by the Gao et al. (2008) reconstruction.

ECHO-G and CSIRO-pre incorporate volcanic forcing following Crowley (2000) and Gao et al. (2008), respectively. HadCM3-pre uses also annual globally defined data updated from Crowley et al. (2003) and converted to monthly aerosol depths assumming a Pinatubo optical-depth time decay. CCSM3, CSM1.4 and CNRM incorporate latitudinal distributions of aerosols following Ammann et al. (2003), albeit with different parametrizations and scaling. EC5MP uses time series of aerosol optical depth at 0.55 μm and of effective radius (Crowley et al., 2008; Crowley and Unterman, 2012). Within the PMIP3 experiments, following Schmidt et al. (2011), the BCC and CCSM4 consider the latitudinally variable optical depth reconstruction from Gao et al. (2008), while CSIRO, GISS, HadCM3 and MPI use the recent Crowley and Unterman (2012) reconstruction,

as the EC5MP. In addition, IPSLCM5 is considered as PMIP3 (Masson-Delmotte et al., 2013) but using a volcanic forcing aside the recommendations of Schmidt et al. (2011). This run follows the CSM1.4 (Ammann et al., 2007) and uses the volcanic forcing estimations originally presented in Ammann et al. (2003). See original references in Tables 2.1 and 3.1 for details on aerosol load and forcing conversions and on parametrizations.

The implementation of volcanic forcing into AOGCMs was done by including either the net effect of stratospheric volcanic aerosols on the global radiation balance (ECHO-G, CSIRO) or by latitudinally resolved changes in optical depth in the stratosphere (EC5MP, HadCM3-pre and PMIP3 runs). These differences may have an impact on the climatic effects in the aftermath of volcanic eruptions on the atmospheric circulation, especially over the extratropical hemispheres during wintertime (e.g. Robock, 2000; Fischer et al., 2007). Although volcanic impacts are restricted to a few years, the temporal clustering of volcanic outbreaks may also have impacts beyond these time scales (see Chapter 4).

The radiative forcing associated to orbital variations is globally small during the CE but potentially important for seasonality changes at high latitudes (Kaufman et al., 2009). While all the PMIP3 runs include the Berger (1978) orbital estimations (see Table 3.1), only CSIRO-pre, IPSL, HadCM3-pre and one of the ECHO-G simulations include these changes following also Berger (1978), and Laskar et al. (2004) in the case of the IPSLCM4 model. EC5MP follows Bretagnon and Francou (1988) for orbital changes and additionally considers nutation.

3.2 Anthropogenic forcing

Estimates of the concentration changes of the main well mixed GHGs (CO_2, CH_4 and N_2O) are obtained from Antarctic ice cores (Forster et al., 2007; Joos and Spahni, 2008). The records used to produce the simulations in Table 2.1 were selected according to the availability of data at the time of production of model experiments. Indeed, in the case of the most recent simulations, they apply almost the same forcing estimations (see Table 3.1).

Figure 3.2 shows the CO_2 concentrations considered by each model simulation or group of simulations that apply the same estimation (e.g. PMIP3 experiments). The CO_2 concentrations in each model (Figure 3.2) were prescribed by different reconstructions as it will be detailed later, except for the EC5MP that calculates it interactively (see Jungclaus et al., 2010). Although the CESM contains also a carbon cycle module that calculates interactively the CO_2 concentrations, this is radiatively inactive (Lehner et al., 2015) and instead the prescribed concentrations from MacFarling Meure et al. (2006) are included (see Table 3.1).

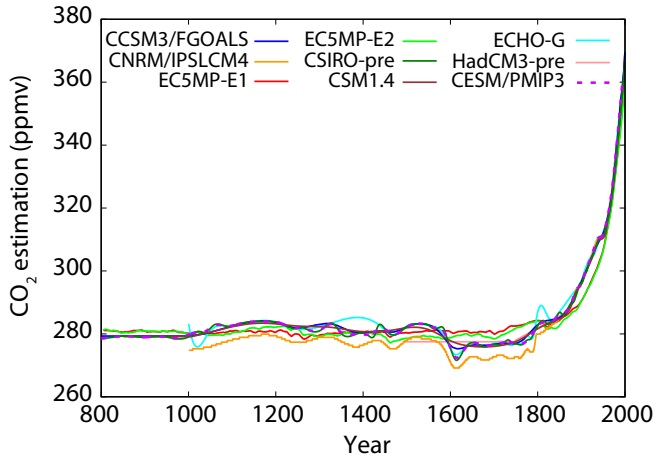

Fig. 3.2: Available CO_2 concentrations (ppmv) used to drive the simulations in Table 2.1; see also Table 3.1 for the original references. Values from the EC5MP are diagnosed online and correspond here to 31-yr filtered outputs of the ensemble averages.

Figure 3.3a,b shows an estimation of the GHG radiative forcing obtained from the concentrations of the three GHGs (CO_2,CH_4 and N_2O) in each model following Myhre et al. (1998) for the nonPMIP3 and PMIP3 experiments, respectively. This allows the comparison of the total effect of CO_2, CH_4 and N_2O between different simulations and also with other anthropogenic and natural forcings as well as with the single CO_2 forcing variations (Figure 3.2).

The GHG forcing estimation alternatives for the PMIP3 are only two, therefore they are referred by the original references instead of the names of the models (Figure 3.3b). Nevertheless, it is worth to mention that all the PMIP3 models, except the HadCM3, follow the GHG estimations suggested by Schmidt et al. (2011), i.e. estimations of MacFarling Meure et al. (2006) for CO_2 and CH_4 and Flückiger et al. (2002) for the N_2O for the preindustrial period tied to Hansen and Sato (2004) estimations after 1850 CE. The other alternative, followed by the HadCM3 (Schurer et al., 2013), although does not differ substantially from the proposed by Schmidt et al. (2011), evidences a slightly larger radiative forcing during the industrial period, based on the Johns et al. (2003) instead of the Hansen and Sato (2004) estimations.

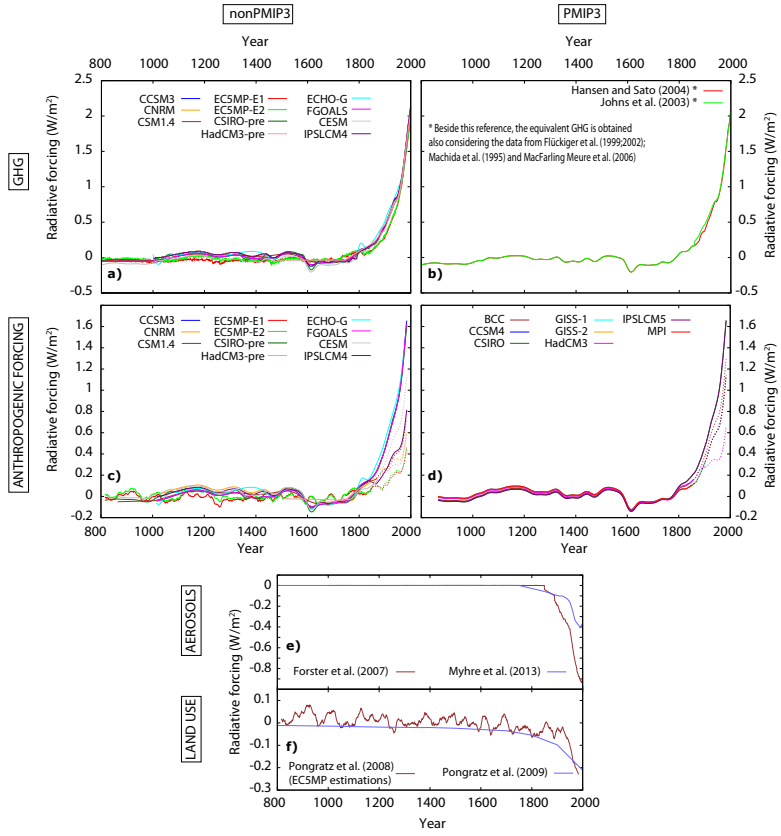

Fig. 3.3: Estimations of anthropogenic external forcings used to drive the simulations in Table 2.1 (see labels). The reference period for all panels showing anomalies is 1500-1850 CE. a, b) Equivalent CO_2 forcing (W/m^2) including well-mixed GHGs for the nonPMIP3 and PMIP3 simulations, respectively. In the case of PMIP3, only the original series are referred; see Table 3.1 for correspondence with the models. Note that the CNRM and IPSLCM4 models depict identical values. c, d) Estimations of total anthropogenic forcing (W/m^2) including the contributions of GHGs, aerosols and land use for the nonPMIP3 simulations for the PMIP3 ones, respectively. Dashed lines indicate that the forcing includes, in at least one of the contributors to the anthropogenic forcing, an approximation to the original forcing applied during that period (see text for details). e) Radiative forcing associated to the two anthropogenic aerosol estimations. f) Radiative forcing associated to the land use changes.

Regarding the nonPMIP3 experiments, the variability among the various GHG forcing estimations is larger. CESM follows, as the PMIP3 cases, the Schmidt et al. (2011) recommendations. HadCM3-pre used estimated changes in GHGs only in the industrial period; values after 1750 CE were taken from Johns et al. (2003) and constant pre-industrial values were assumed before this date. The ECHO-G model used LM reconstructions from Etheridge et al. (1996) for CO_2, and from Etheridge et al. (1998) for CH_4; Battle et al. (1996) estimates for N_2O were included after 1850 CE and assumed constant before. CSM1.4, CCSM3 and FGOALS used reconstructions from Etheridge et al. (1996) for CO_2, Blunier et al. (1995) for CH_4, and Flückiger et al. (2002) for N_2O. All simulations from the last four models use different spline interpolations to obtain annual concentration values. In addition to the different origin of the data the different interpolation approaches produce variability in the evolution of pre-industrial concentrations and forcings in Figure 3.3a. The CSIRO-pre runs used values derived from updated reconstructions provided by MacFarling Meure et al. (2006). CNRM and IPSLCM4 incorporate also estimations of MacFarling Meure et al. (2006) for CO_2. However, transient changes in CH_4 and N_2O concentrations are considered only after 1850 CE and taken from Blunier et al. (1995) and Flückiger et al. (2002), respectively. Before this date the concentrations are kept constant. These pre-1850 CE concentration values are higher than those suggested by MacFarling Meure et al. (2006). Thus the CO_2 concentrations in the CNRM and IPSLCM4 simulations were indeed lowered by about 5 ppmv to compensate for the relatively high CH_4 and N_2O levels. This can be appreciated in the lower CO_2 of CNRM/IPSLCM4 in Figure 3.2, while in Figure 3.3a the GHG forcing of CSIRO-pre, CNRM and IPSLCM4 co-vary in phase.

CO_2 and GHGs forcing evolve very similarly for all simulations that prescribed GHG concentration values in Figures 3.2 and 3.3a,b. Excluding arbitrary changes produced by spline interpolations, the multicentennial changes displayed by the various forcings are due to natural feedbacks from the ocean and terrestrial biosphere in response to variations in climate; additional effects of land cover change are also possible (Pongratz et al., 2010). The diagnosed ensemble averages of CO_2 concentrations simulated by EC5MP (Figure 3.2) are below the MacFarling Meure et al. (2006) observations in the 20th century. This discrepancy is arguably due to an underestimation of the emissions related to land-use change among other factors discussed in Jungclaus et al. (2010). The pre-industrial CO_2 concentration values show more variability in the E2 ensemble, albeit with changes of somewhat smaller magnitude than in the MacFarling Meure et al. (2006) reconstruction. The larger variability in the E2 (relative to the E1) ensemble may be related to its slightly larger temperature fluctuations (see Chapter 4), with higher values during the MCA and lower during the LIA.

The smaller number of members in E2 (3) relative to E1 (5) may also have contributed to this effect, with less chances of cancelling out deviations associated to internal variability among ensemble members. The observed minimum in the 17th and 18th centuries is not reproduced.

Land use and land cover changes are considered by some of the simulations that incorporated information from several datasets available through time (see Table 3.1). Land use changes as reconstructed by Ramankutty and Foley (1999) are used in the CNRM simulation since 1700 CE. HadCM3-pre uses land surface data from Wilson and Henderson-Sellers (1985) modified with crop history from Ramankutty and Foley (1999) and pasture change data from Goldewijk (2001). The EC5MP ensembles use a reconstruction of global agricultural areas and land cover from Pongratz et al. (2008). The latter is the one recommended for the PMIP3 simulations (Schmidt et al., 2011) together with new available reconstructions (Kaplan et al., 2011; Goldewijk et al., 2011).

Details regarding the inclusion of anthropogenic sulphate aerosols (CNRM, CSM1.4, HadCM3, IPSLCM4, CESM, BCC, CCSM4, GISS, MPI) and halocarbons in the simulations are not provided here and the reader is addressed to the original references in Table 3.1 for more information. Nevertheless, the estimations of radiative forcing associated to this and the land use forcings will be provided later on.

Indeed, Figure 3.3c,d shows the equivalent anthropogenic forcing integrating the available forcing data of GHGs, anthropogenic sulfate aerosols and land use changes for the nonPMIP3 and the PMIP3 models, respectively.

Aerosol forcing data were available from the CNRM and IPSLCM4 experiments in which it was prescribed, but not for the others that include this forcing factor (see Tables 2.1 and 3.1). For the latter, an aerosol-only sensitivity experiment, that would allow assessing the effects of aerosol parameterizations involved, in each model does not exist. Instead, the values included in the CNRM and IPSLCM4 experiments (Forster et al., 2007) and, on the other hand, the ones shown in AR5 (Myhre et al., 2013) as the current estimations for anthropogenic aerosol forcing, are considered to provide an estimation of the potential effect, in radiative forcing units, of this forcing in these simulations. These two different estimations are shown in Figure 3.3e. The two different alternatives for the radiative forcing associated to the emission of anthropogenic aerosols represent also two different states of knowledge about this factor. There exists a considerable difference between the two series with the conclusion that the radiative forcing associated to the anthropogenic aerosols could have been overestimated previously (Forster et al., 2007), according to recent estimations (Myhre et al., 2013).

Anthropogenic aerosol forcing applied to the pre-PMIP3 (nonPMIP3 except CESM) runs can be represented by the Forster et al. (2007) estimations. In turn,

the estimation shown in AR5 (Myhre et al., 2013) is the one that could be, to a good approximation, considered as the radiative forcing associated to the anthropogenic aerosol loading applied to the PMIP3 (and CESM) experiments. An exception to the latter is the HadCM3, that specifically refers to the Forster et al. (2007) estimations as anthropogenic aerosol boundary conditions (Schurer et al., 2013). These approximations do not take into account either the original sulphate mass loading, or any of the physics involved in the model parametrizations and therefore are not accurate in representing the real forcing in the simulations.

However, including this forcing, even if subject to large uncertainties is arguably more realistic than excluding the effect of anthropogenic aerosols when estimating total anthropogenic forcing in these simulations.

Regarding the land use-land cover (LULC) changes, a single-forcing simulation applying the corresponding LULC forcing (Table 3.1) would be necessary for the proper estimation of the radiative forcing associated to this factor in each model. Nevertheless, as for the anthropogenic aerosol forcing, these series are not always available. The two existing available estimations of the LULC changes in radiative forcing units are shown in Figure 3.3f. Both estimations (Pongratz et al., 2009; Jungclaus et al., 2010) correspond to changes in the surface albedo obtained from two LULC single-forcing simulations using the Pongratz et al. (2008) land cover reconstruction. The Pongratz et al. (2009) estimation is obtained from the ECHAM5 atmospheric model, while the series from Jungclaus et al. (2010) comes from the EC5MP ESM consisting of the atmospheric part ECHAM5 and the MPIOM ocean model and including an interactive carbon cycle. The EC5MP LULC series (Jungclaus et al., 2010) present larger high frequency variability than the Pongratz et al. (2009) estimations, but both evidence a noticeable agreement in contributing a long term cooling, similarly as was evidenced by the anthropogenic aerosol estimates (Figure 3.3e). Apart from the EC5MP, the HadCM3-pre, CNRM and the PMIP3 simulations consider the LULC forcing (see Tables 2.1 and 3.1). There is no a LULC single-forcing for these models, therefore, the only viable approach is to take one of the series in Figure 3.3f as a substitute of the original radiative forcing associated to the LULC changes applied. Based on the agreement at low frequencies between the series from Pongratz et al. (2009) and EC5MP (Jungclaus et al., 2010) in Figure 3.3f, any of them could be considered appropiate. The estimation from Pongratz et al. (2009) is selected herein due to the absence of high frequency variability found in the EC5MP (Jungclaus et al., 2010), that was probably related to the carbon cycle of the model and thus model-dependant.

On the basis of the previous rationale, the total anthropogenic forcing illustrated in Figure 3.3c,d is used as a representation of the actual forcings used in each of the model simulations over the whole millennium for the CCSM3,

CSIRO-pre, ECHO-G, IPSLCM4, FGOALS, CSIRO, IPSLCM5 and also until the 19th century for the CNRM, CSM1.4, EC5MP, HadCM3-pre, CESM, BCC, CCSM4, GISS, HadCM3 and MPI; thereafter, these forcing estimations are subjected to the approximations and limitations described above (dashed lines in Figure 3.3c,d).

During pre-industrial times all simulations in which CO_2 concentration was prescribed, i.e. all except EC5MP, display a very similar evolution of the total anthropogenic forcing. In turn, EC5MP shows less low frequency variability during this period than the others. In contrast, during the 20th century there exists a larger variability among the various estimates of the anthropogenic forcings. Models in which only GHGs are applied (i.e. CCSM3, ECHO-G, FGOALS, CSIRO-pre, CSIRO and IPSLCM5) are, as expected, those showing the largest increase in the total anthropogenic forcing. According to the approximations for the anthropogenic aerosols and land use changes included in Figure 3.3c,d, the other models gradually decrease their anthropogenic forcing during this period. The EC5MP anthropogenic forcing presents the lowest increase during the 20th century. To the latter contributes mainly the estimations of the anthropogenic aerosol forcing. These estimations have suffered significant changes during the last years in the sense that models including the Forster et al. (2007) estimations (nonPMIP3 except CESM and HadCM3 in PMIP3) show less pronounced trends in the anthopogenic forcing than the models including the Myhre et al. (2013) estimations (PMIP3 and CESM). The latter is due to the contribution of the larger negative radiative forcing values associated to the anthropogenic aerosol estimations from Forster et al. (2007) than the one proposed by Myhre et al. (2013); see Figure 3.3e.

EC5MP considers, apart from the Forster et al. (2007) estimations for anthropogenic aerosols, land use changes from Pongratz et al. (2008), thus evidencing the lowest increase of total anthropogenic forcing during the industrial period (see Figure 3.3c).

The temperature response in each model will result from the balance between this anthropogenic effect on the total external forcing, the effect of natural forcings displayed in Figure 3.1 and the climate sensitivity of each model, as it will be shown in Chapter 4.

3.3 Total external forcing

A total external forcing (TEF hereafter) can be obtained by adding for each model all natural and anthropogenic forcings considered in Sections 3.1 and 3.2

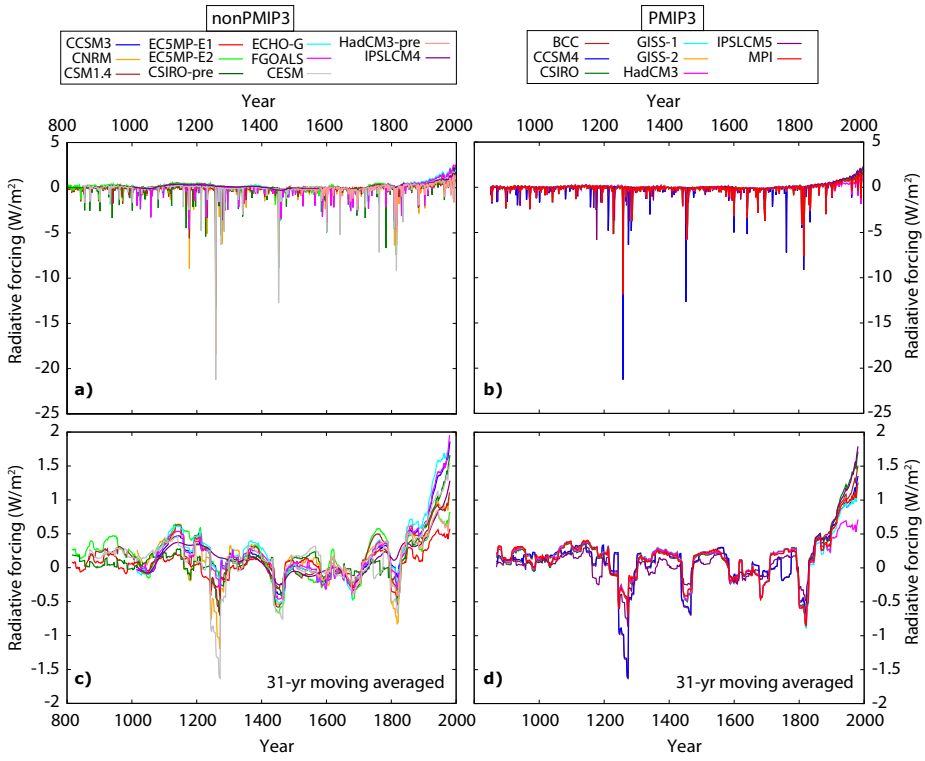

Fig. 3.4: Estimations of TEF including anthropogenic and natural contributions. a, b) Annual data obtained for the nonPMIP3 and PMIP3 models, respectively. c,d) 31-yr moving average filtered outputs of anomalies of TEF wrt 1500-1850 CE for nonPMIP3 and PMIP3 models, respectively. Note: IPSLCM4 does not include volcanic forcing; for the ECHO-G and CSIRO-pre models, estimations of TEF correspond to the simulations with volcanic forcing.

as shown in Figures 3.4 and 3.5. The use of TEF helps us to better understand the temperature response of the models and the assessment of climate sensitivity developed in Chapter 4.

Figure 3.4a,b shows the sum of all the contributions to the external forcing configuration considered by each model in Table 2.1, distingushing between the nonPMIP3 and PMIP3, respectively. For comparison with the analysis in the following chapters, Figures 3.4c,d and 3.5 shows 31-yr filtered outputs of TEF expressed as anomalies relative 1500-1850 CE. Figures 3.4c-f and 3.5 shows all

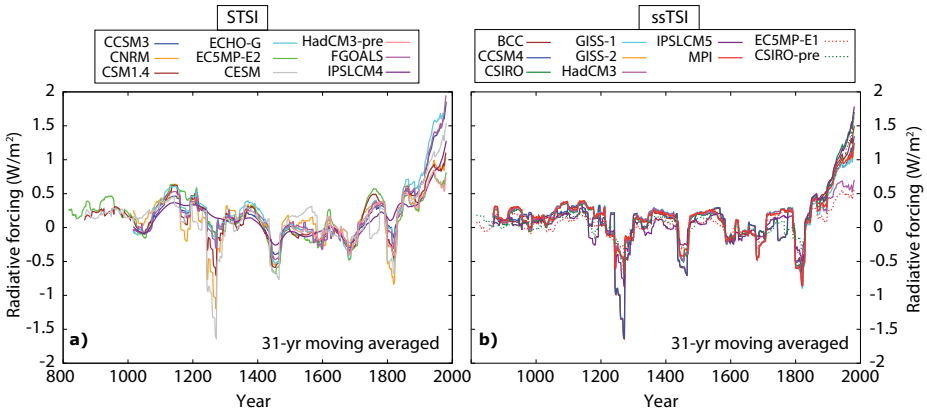

Fig. 3.5: Equivalent to Figure 3.4c,d but dividing the various TEF configurations attending to the level of solar variability, STSI and ssTSI, respectively.

the TEF configurations and classifies them in different panels according to the model (nonPMIP3 and PMIP3 in Figures 3.4c and d) or attending to the level of solar variability forcing considered (STSI and ssTSI in Figures 3.5a and b) in order to analyse better the different behaviours according to the level of solar variability.

The changes in forcing during the pre-industrial period are dominated by solar and volcanic activity. The TEF of the IPSLCM4 model, for which volcanic forcing is not included, serves as an illustration of the non negligible effect of volcanoes at low frequencies (see Figures 3.4c,d and 3.5a,b). This effect is noticeable for instance in the 12th, 13th, 15th and 19th centuries during which decreases in TEF are perceptible during times of recurrent large volcanic events, sometimes also coinciding with minima in solar forcing. Based on the agreement in timing of the volcanic reconstructions mentioned previously (Section 3.1, Figure 3.1c,d), the peaks observed in the TEF series are synchronous among all the scenarios. In turn, the magnitude of the peaks observed in TEF is related to the intensity of the volcanic events in each original volcanic reconstruction (i.e. the largest peaks are related to the Gao et al., 2008, reconstruction). The variability within the TEF series is larger for the nonPMIP3 group than for the PMIP3 (Figure 3.4c,d), due to the larger hetereogenity among the natural forcings contributing to each TEF estimation. Apart from the influence of the volcanic activity, the solar forcing is the other main factor responsible of the low frequency variability of TEF for the pre-industrial times. Indeed, if the TEF series are grouped according to the level of variability of the solar forcing considered (Figure 3.5a,b), the spread of both

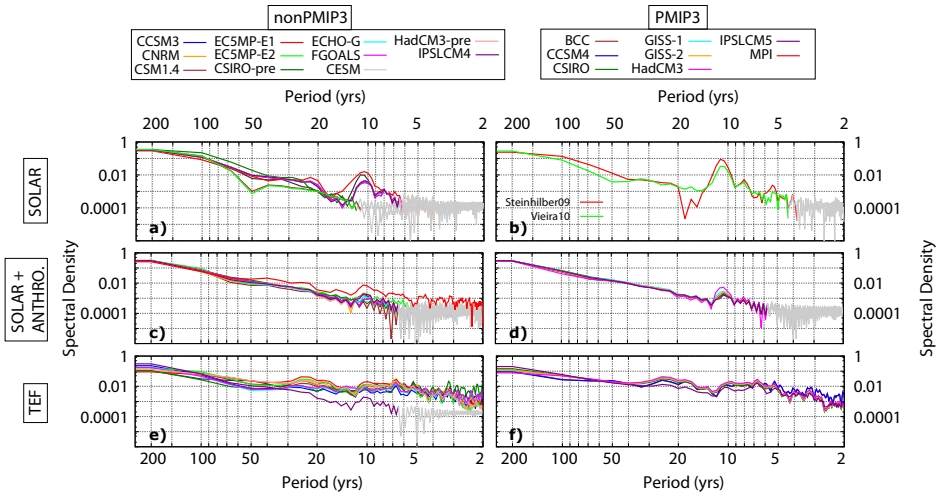

Fig. 3.6: Normalized spectra for various combinations of external forcing for the nonPMIP3 (left) and PMIP3 (right) models: a, b) solar forcing (Figure 3.1a,b); c, d) solar and anthropogenic forcing, i.e. TEF without volcanic activity; e, f) all natural and anthropogenic forcings, i.e. TEF (Figure 3.4). Grey lines correspond to frequency bands affected by Gibbs oscillations (see text for details). For the sake of clarity, spectral curves corresponding to AR1 processes are not shown. Note: IPSLCM4 does not include volcanic forcing; for the ECHO-G and CSIRO-pre models in Table 2.1, estimations of TEF correspond to the simulations with volcanic forcing.

groups becomes more similar. This means that the larger disagreement found within the nonPMIP3 series was based on the coexistence of two groups of solar variability (STSI and ssTSI) in it.

In the 20th century, the distribution of forcing trends is similar to that in Figure 3.3c,d, with the exception of the EC5MP-E1 ensemble that includes ssTSI forcing and is weighted down relative to EC5MP-E2 with STSI forcing, i.e. considering more low frequency variability.

We examine the normalized spectra of the forcing time series in order to identify the forcing signatures in the frequency domain. The relative distribution of variance in the frequency domain is displayed for several forcing combinations in Figure 3.6. Figure 3.6a,b shows the spectra for the various solar forcing configurations for the nonPMIP3 models and those used by the PMIP3 models (Figure

3.1a,b), respectively. The estimation of these spectra are somewhat burdened by the fact that solar forcing specifications do not have variability at high frequencies, which produces Gibbs oscillations in this part of the spectrum (gray lines; Bloomfield, 1976) and precludes for this case a clear comparison of the relative contribution to total variance of low and high frequencies. However, this is useful as an illustration of the relative importance of the variance accumulation at decadal timescales, and for comparison with other forcings in Figure 3.6. Within the nonPMIP3 group (Figure 3.6a), CSM1.4, CSIRO-pre and EC5MP-E2 forcings do not include an 11-yr solar cycle signal and thus do not show any contribution to variance centred around the 11-yr band. Their spectra suffer from Gibbs oscillations below the 20-yr time scale. This problem is reduced in the other simulations that do consider an 11-yr solar cycle. In the cases of EC5MP-E1 and CESM, this variability is imposed through the whole millennium, thus showing maximum variance at these frequencies. CCSM3, FGOALS, CNRM, ECHO-G, HadCM3 and IPSLCM4 use the 11-yr variability in the last few centuries (see Section 3.1). This is reflected in an overlap of their spectra at these timescales, albeit showing a smaller proportion of variance than EC5MP-E1. It is also interesting to note the relative increases of variability from the 20 to 40 yr timescales. Within the PMIP3 group (Figure 3.6b), both solar reconstructions show a significant contribution to variance around the 11-yr band, associated to the 11-yr solar cycle imposed through the whole millennium to both series, and the Gibbs oscillations are observed only below the 5-yr time scale. Steinhilber et al. (2009) evidences an interesting decay of variance close to the 20-yr band, larger than that evidenced by some of the nonPMIP3 models (CCSM3, FGOALS, CNRM, ECHO-G, HadCM3 and IPSLCM4).

Figure 3.6a,b can be compared with Figure 3.6c,d which shows the spectra corresponding to the sum of solar and anthropogenic forcings (i.e. TEF without volcanic activity). Here, the proportion of variability accounted for by the 11-yr solar cycle is importantly diminished and only noticeable in the EC5MP-E1 case (Figure 3.6c) and in the whole suite of PMIP3 models (Figure 3.6d). The radiative forcing associated to the land use changes in the EC5MP introduces variability at high and low frequencies (Jungclaus et al. 2010; Figure 3.3f), thereby contributing to avoid Gibbs oscillations in the EC5MP spectra of anthropogenic forcing (red line in Figure 3.6c).

Figure 3.6e,f shows spectra for TEF (Figure 3.4) for the nonPMIP3 and the PMIP3 models, respectively. Two features are prominent. Firstly, the relative contribution of the 11-yr solar cycle has been greatly diminished in all cases. This suggests that only a very small contribution of the simulated global or hemispheric signal is to be expected at these timescales in the model response for the last millenium. Additionally, the proportion of variance from interannual

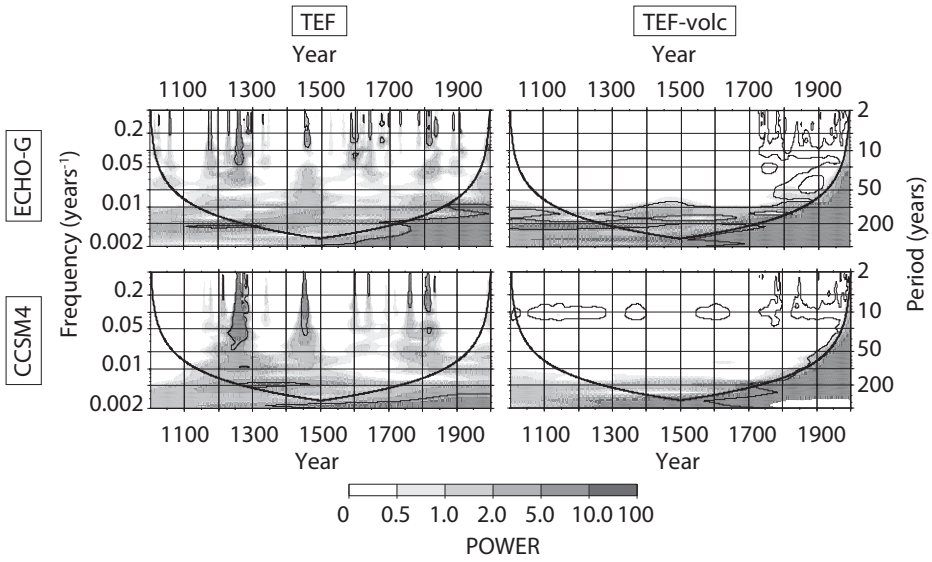

Fig. 3.7: Wavelets for the ECHO-G (top) and CCSM4 (bottom) of two external forcing configurations: TEF (left) and TEF without including the volcanic forcing (right), i.e. anthropogenic forcing plus solar and orbital variability in the case of CCSM4.

to multidecadal timescales, up to 40 yr periods, is increased in comparison to Figures 3.6a-d. This is arguably due to the multidecadal variability associated with the occurrence of large volcanoes. An exception here is the IPSLCM4 case (purple line in Figure 3.6e), which does not include volcanic forcing and serves as a reference that can be compared with the other spectral curves.

Indeed, the pronounced influence of the volcanic activity in the multidecadal variability of TEF is supported also by Figure 3.7 that shows a wavelet spectra analysis. Two different configurations of external forcing are tested, i.e. TEF and TEF without including the volcanic activity, following Figure 3.6c-f, for two different models as an example (CCSM4 and ECHO-G). It is observed that the wavelet spectra of TEF show significative variance peaks spanning from interannual to multidecadal timescales, up to 50 yr periods, coincident with the clustering volcanic eruptions in both models. Instead, at considering TEF withouth the volcanic activity, there is no significant variance contribution at multidecadal timescales, in spite of solar variability being included in TEF estimations. In this cases, i.e. TEF without volcanic forcing, the influence of the 11-yr solar cycle is

observed during the period when it is imposed, i.e. the whole millennium for the CCSM4 and after 1610 CE for the ECHO-G (see Section 3.1).

For longer timescales, Figures 3.6e,f evidence that series showing relatively lower contributions to variance are those using the ssTSI solar forcing (EC5MP-E1, CSIRO-pre and PMIP3) and the ones considering the largest volcanic forcing (CNRM, CESM, BCC and CCSM4, see Figure 3.1c,d). The latter may be due to the higher proportion of high frequency variability associated to volcanic activity.

3.3.1 Hemispheric total external forcing

Figures 3.4 and 3.5 have shown the global estimation of TEF for each model, calculated considering the global radiative effect of all the individual forcing contributors. We suggest herein that if these radiative effects produced by each forcing factor were known over more reagional spatial regions, it would be also possible to calculate regional estimations of TEF over those areas.

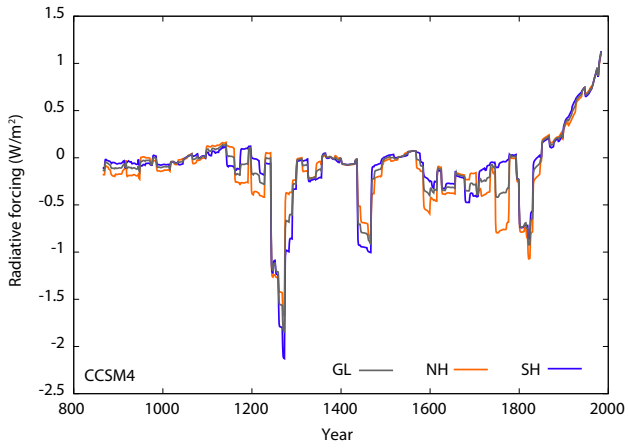

Fig. 3.8: 31-yr moving average filtered outputs anomalies of the TEF wrt 1500-1850 CE for three different spatial regions: GL, NH and SH.

Sub-global estimations of TEF are based on, and limited to, the availability of spatial characterization of each individual forcing in radiative forcing units. Only for the volcanic reconstructions from Gao et al. (2008) and Crowley and

Unterman (2013), suggested for the PMIP3 models (see Section 3.1), various latitudinal bands are available so far, thereby allowing for the calculation of TEF at hemispheric scales. The latter is possible therefore for all the PMIP3 models, except IPSLCM5, and CESM.

Figure 3.8 shows for the CCSM4, as an example, the hemispheric estimations of TEF, together with the global one already shown in Figure 3.4. The behaviour of the three variants of TEF is quite similar and only small differencies regarding the magnitude of some peaks are observed, depending on the volcanic event considered and its geographical location. Interestingly, changes related to the volcanic forcing for the SH are more intense than for the NH during the first half of the millennium, although after 1500 CE, the NH TEF peaks seem to be more pronounced than the SH ones. The small differences observed could be useful to understand, for instance, some possible differences noticed in the hemispheric temperature response by models.

Thus, these hemispheric TEF series will be considered in the following chapters to conciliate the radiative forcing with the spatial domain of a given temperature reconstruction. This approach was not included in Fernández-Donado et al. (2013). Nevertheless, it is worth to mention that since only the volcanic and not other forcings are available at different spatial scales, the hemispheric TEF estimations proposed herein are just mere approximations. For a proper hemispheric, or even regional, TEF estimation, other forcings, such as the GHG, the anthropogenic aerosols or the land use cover changes should also provide information at different spatial scales. Thus, conclusions extracted from this regional TEF should always be considered cautiously.

3.4 Conclusions

A comprehensive review of the forcing factor reconstructions used as boundary conditions to drive AOGCM simulations of the CE has been presented and discussed. Since the suite of simulations considered among the present manuscript include PMIP3 and nonPMIP3 experiments, this chapter complements the information from Schmidt et al. (2011) in which the suitable PMIP3 forcings are described and Fernández-Donado et al. (2013) focused on nonPMIP3 ones. Also, it updates the information provided in Masson-Delmotte et al. (2013).

Each individual forcing factor has been characterized and discussed. The large heterogenity found within the various solar forcing reconstructions applied to the models is noticeable. Indeed, two different levels of solar variability forcing, high and low, are distinguished, i.e. STSI and ssTSI, respectively. The STSI group includes reconstructions with an increase in TSI larger than 0.20 % from LMM to present, while the ssTSI shows an increase of less than 0.1 %. The ssTSI is

applied in the PMIP3 simulations and in a couple of models of the nonPMIP3 group, while the STSI forcing is generally devoted to those simulations developed before the PMIP3 project. Indeed, simulations will also be distinguished and classified according to the solar forcing variability scenario applied (STSI vs. ssTSI).

Regarding other forcing factors, the largest differences among the various reconstructions applied are found in the volcanic activity and the anthropogenic aerosols. The overall agreement among the various reconstructions in the timing of the volcanic events is evidenced, however their magnitude is still uncertain when comparing the various sources. Specifically the Gao et al. (2008) reconstruction, corresponding with one of the series suggested by Schmidt et al. (2011) for the PMIP3 experiments, presents the largest volcanic activity. In turn, the estimations of the radiative forcing asssociated to the anthropogenic aerosols have been significantly reduced from previous evaluations to current ones. In general, PMIP3 runs apply an anthropogenic forcing whose radiative effect is lower (in absolute value) than the corresponding to the nonPMIP3 simulations.

Additionally, for each simulation, the TEF that accounts for the total external forcing applied to each model when including all the individual forcing contributors was estimated. The TEF allows to describe the role and the impact of each individual forcing to the total forcing configuration, and it will be used in Chapters 4 and 5 to better understand the relation between the temperature response and forcing applied. The large influence of the volcanic activity is evidenced not only at interannual, as expected, but also at multidecadal timescales. Indeed a considerable part of the multidecadal variability observed in TEF series is associated to the volcanic forcing during the complete LM.

During the industrial period, the set of anthropogenic forcing factors included in each simulation has an influence in the TEF variability. Thus, models applying a complete set of anthropogenic forcings, i.e. GHG, aerosols and land use, evidence a modest positive trend of TEF during the industrial period, while those not including aerosols or land use forcings render a more pronounced trend. Nevertheless, due to the lack of estimations of the radiative forcing associated to some of the external forcing contributors (i.e. anthropogenic aerosols and land use changes) during the industrial period, estimations of TEF should be analyzed cautiously during that period.

4

Simulated temperature response*

This chapter presents an analysis of the simulated temperature evolution during the CE, at global and hemispheric scales, based on the ensemble of state-of-the-art paleoclimatic experiments introduced in Chapter 2. The heterogeneity of the simulations considered in this Thesis allows an evaluation of the uncertainty that arises as a consequence of using different models to produce the temperature simulations and and also in some cases, the uncertainty associated to the different initial and boundary conditions within a single model.

The climate variability of the simulated temperature during the CE can be separated into components attributable to externally forced response and internally induced climate variability (Schurer et al., 2013). Within this context, and based on the description of the external forcing configurations from Chapter 3, an analysis of the relation between simulations and external forcing is presented. Furthermore, a quantitative assessment of the linear response of temperature to the external forcing during the LM is provided herein for the suite of model experiments and compared to ECS and TCR values as alternative metrics of the sensitivity of climate to changes in forcing.

The chapter is structured as follows. Section 4.1 presents the temporal evolution of the LM simulations at global and hemispheric scales. Section 4.2 analyses the relationship between the temperture response from the ensemble of model

* The main contents of this chapter are included in:
- Fernández-Donado, L., J. F. González-Rouco, C. C. Raible, C. M. Ammann, D. Barriopedro, E. García-Bustamante, J. H. Jungclaus, S. J. Lorenz, J. Luterbacher, S. J. Phipps, J. Servonnat, D. Swingedouw, S. F. B. Tett, S. Wagner, P. Yiou and E. Zorita, 2013: Large-scale temperature response to external forcing in simulations and reconstructions of the last millennium. *Climate of the Past*, **9**, 393-421. DOI 10.5194/cp-9-393-2013.

simulations and the external forcing. Section 4.3 establishes a straightforward metric to quantify this response during the LM in the basis of the results in Section 4.2. Section 4.4 summarizes the main conclusions inferred in this chapter.

4.1 Simulated temperatures at global and hemispheric scales during the Common Era

This section makes use of the simulations described in Chapter 2 (detailed information was provided in Table 2.1) in order to assess the evolution of the temperature during the CE. As it was stated in Chapter 2, most of the analysis carried out herein focuses on the last 12 centuries (800 - 2000 CE) as this is the time interval that most simulations span (see Table 2.1). Note also that the series considered are usually 31-yr low-pass filter outputs as the interest of this work focuses on the variability at multidecadal and longer timescales, where the imprint of the external forcing can be distinguised from the internal variability of the climate system.

Figure 4.1 shows the global temperature anomalies with respect to the period 1500-1850 CE for all simulations (Table 2.1). Figure 4.1a shows the whole suite of simulations in Table 2.1, with nonPMIP3 and PMIP3 simulations highlighted with different linestyles. Figure 4.1b,c group the experiments, for the sake of clarity, according to their corresponding high (STSI) or low (ssTSI) solar variability scenario. The choice of 1500-1850 CE as a reference relies on the fact that this period is covered by all simulations (recall that some of them only span the last 500 yrs., Table 2.1). Additional rationale for such a selection is the similarity of the forcing conditions during the period considered (Chapter 3; Figures 3.4 and 3.5) in comparison to the industrial period.

The temperature evolution of the simulations in Figure 4.1 is consistent for all models and experiments and evidences of the three key periods of the LM (higher temperatures in the MCA and the industrial period and a relative minimum in the LIA) are noticeable. Moreover and in spite of the relative differences among the inter-model forcing configurations, the trajectories of all simulations show a high degree of similarity also at multidecadal timescales. Most of the simulations show minima during the Wolf, Spörer, Maunder and Dalton intervals, following the solar forcing estimations in Figure 3.1, albeit modulated by the presence of volcanic activity.

Overall, it can be said that the simulated trajectories shown in Figure 4.1 follow closely the TEF in Figures 3.4 and 3.5. The distribution of warming trends in the last two centuries of the simulations follow also a similar arrangement, in spite of the limitations discussed regarding the estimation of TEF (Chapter 3).

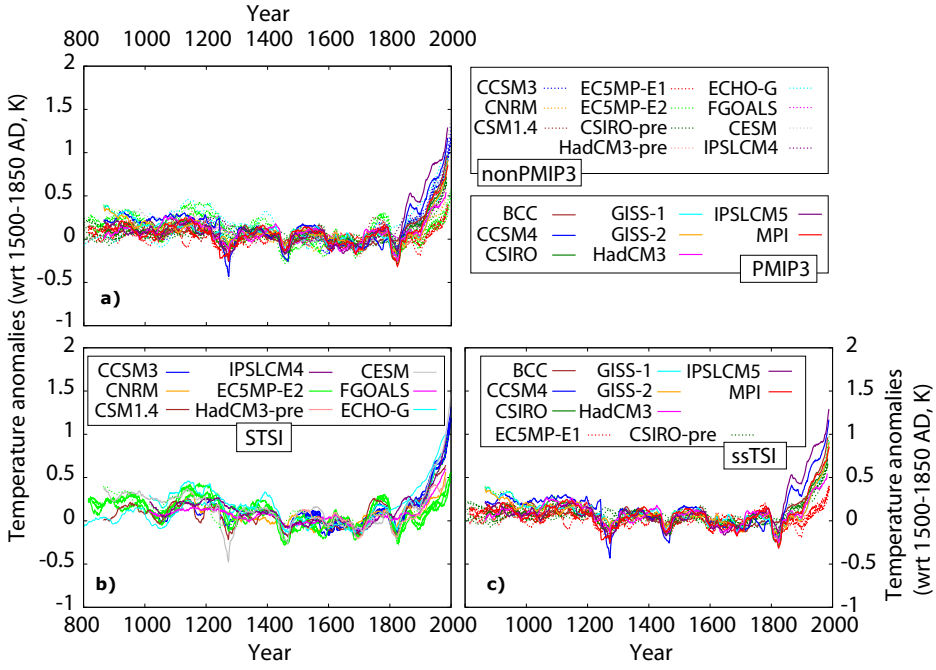

Fig. 4.1: Global simulated temperature anomalies wrt 1500-1850 CE. Series are 31-yr low pass filter outputs. In a) all the available simulations are presented together (dashed lines for the nonPMIP3 runs and solid lines for PMIP3). b) and c) present, for clarity, separate panels for the experiments driven by STSI and ssTSI solar forcing, respectively.

The largest temperature increases are simulated by the runs incorporating only GHGs and natural forcing and decrease according to the inclusion of additional factors (aerosols, land use) that contribute with negative forcing during this period. The 20th century trends are, however, not solely a function of the applied external forcing but also of model sensitivity, as it will be analysed in Section 4.3.

During preindustrial times, the ssTSI group of simulations (Figure 4.1c) shows reduced low frequency variability compared to the STSI group (Figure 4.1b). Indeed, the high solar varibility suite of simulations shows larger changes in amplitude, particularly during the MCA warming. Nonetheless, it is worth noting that the STSI ensemble is characterized by a larger spread of estimations than the ssTSI during preindustrial times. The latter is supported by Figure 4.2a

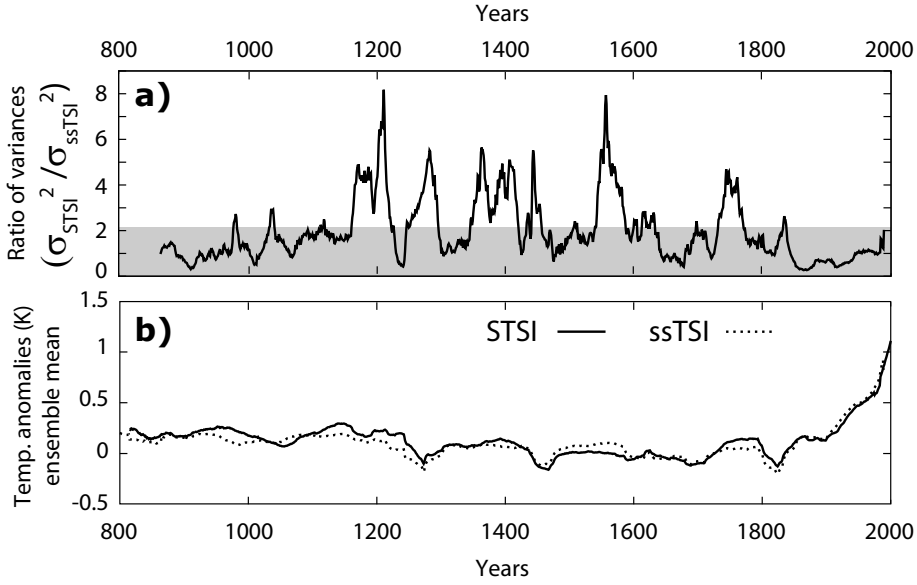

Fig. 4.2: a) Ratio of variances between the STSI and the ssTSI ensembles for each year. White background indicates significant values according to a Fisher test (with $\alpha = 0.05$ significance level). b) Anomalies of global temperature simulation ensembles averages for the STSI and ssTSI groups. Averages are calculated after a 31-yr moving average filter of each member of the ensemble.

that shows the ratio of variances between the STSI and ssTSI ensembles for each year. A significant larger variance is observed intermitently for the STSI ensemble during 1200 - 1800 CE. It is pertinent therefore to relate the differences in the amplitude of the temperature response to the corresponding differences in the external forcing applied. Indeed, if the TEF calculated in Chapter 3 is used now to represent the ratio of variances between high *vs.* low solar scenarios in Figure 4.3a, a comparable sequence of episodes where the ensemble of strong solar variability evidences larger spread of TEF amplitudes can be observed between 1200 - 1800 CE. The later is tantamount to relate the amplitude of the temperature response to the configuration of the external forcings considered.

Figure 4.2b shows the ensemble averages of the simulated global temperature for the two solar variability groups, STSI and ssTSI. The averaging of all simulations within each group considerably reduces the internal variability as well as the relative differences in the external forcing. The signal of the common re-

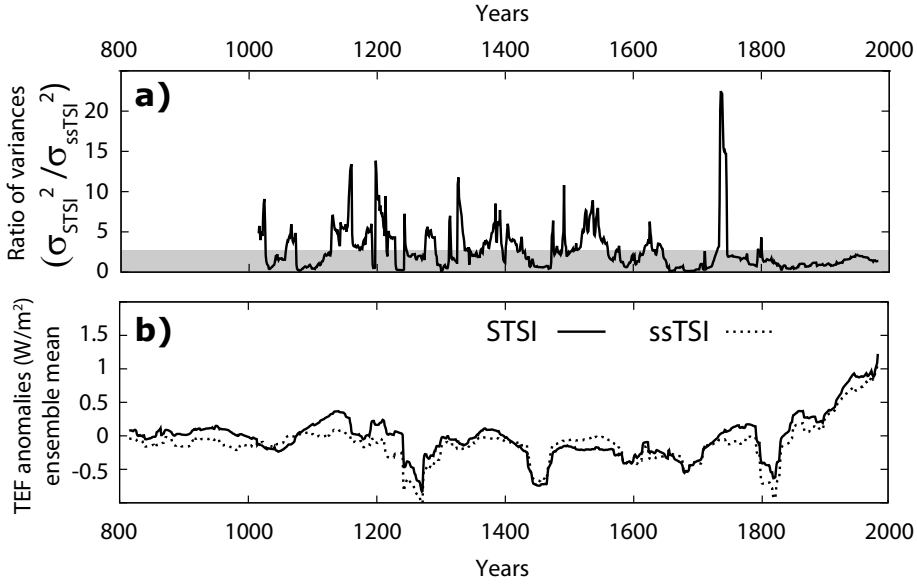

Fig. 4.3: As Figure 4.2 but for TEF configurations instead of simulations.

sponse to the external forcing is, in contrast, emphasized. It is noticeable that there are no significant differences (at the 95% confidence interval) in the amplitude of temperature changes of both ensemble averages, despite the warmer MCA evidenced by the STSI ensemble mean. The same can be said in the case of the TEF averages for each ensemble (Figure 4.3b). Consequently, the spread of temperature responses to the corresponding external forcings can be linked to differences in initial and boundary conditions of the simulations that is cancelled out when the ensemble mean is considered.

Figure 4.4 is equivalent to Figure 4.1 but for the NH (Figure 4.4a) and the SH (Figure 4.4b) temperatures. Qualitatively, the behaviour of the temperature response at hemispheric scales in Figure 4.4 is comparable to what was observed at the global scale in Figure 4.1. The key periods of the LM are also present if the hemispheric variations of temperature are considered. The most striking feature if the NH and SH are intercompared in Figure 4.4 is the attenuated level of low frequency variability of the SH simulated temperature. The latter can be illustrated with Figure 4.5 where the standard deviation of the temperature during the LM is represented for each model simulation and region (GL, NH and SH). Higher values are in general apparent for all simulation in the case of the

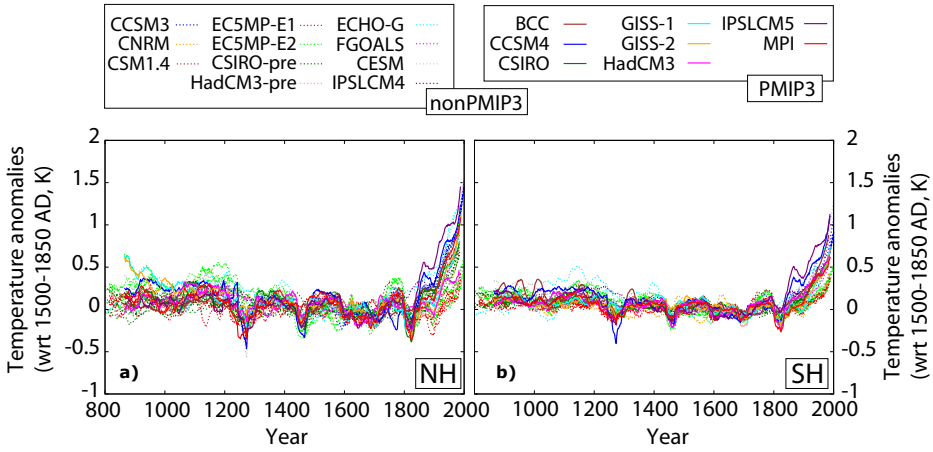

Fig. 4.4: NH (a) and SH (b) simulated temperature anomalies wrt 1500-1850 CE. All available simulations are shown (dashed lines for the nonPMIP3 runs and solid lines for the PMIP3 ones). Series are 31-yr low pass filter outputs.

NH compared to the SH and also to the global temperatures. The latter can be associated to the larger proportion of land included in the NH in relation to GL or SH targets, that favours a response to the forcing of larger amplitude (Zorita et al., 2005).

4.2 Response to changes in external forcing

The previous section evidenced the similarity, suggestive of a linear relation, found at multidecadal timescales and above between the evolution of the simulated temperatures (Figures 4.1 and 4.4) and the TEF (Figure 3.4). This section aims at analysing in more detail the relationship between both variables.

A first approach is based on calculating the correlation between both, temperature and forcing for each simulation. Figure 4.6a shows the correlation values and associated uncertainties obtained for 31-yr filtered temperatures and global TEF for all the simulations spanning the whole millennium. These correlations are calculated for GL, NH and SH temperatures, therefore three different values are obtained for each simulation.

All the correlation values obtained are significant ($\alpha = 0.05$ significance level), supporting the existence of some linear relationship between temperature simulations and TEF. Most of the correlation values are relatively high. The mean value

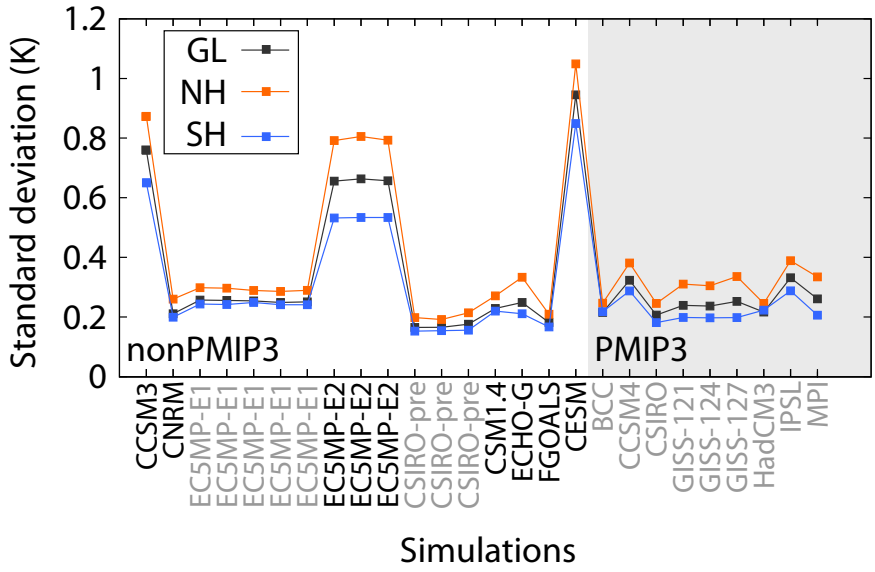

Fig. 4.5: Standard deviation calculated over the LM of each temperature simulation from Table 2.1 for global and hemispheric scales (GL, NH and SH). The nonPMIP3 simulations are plotted over a white backgroud, while the gray one corresponds to the PMIP3 runs. The name of the simulations driven by a STSI forcing are written in black while those using the ssTSI scenario are in gray.

of the correlations calculated in Figure 4.6a is around 0.8. In general, considering the whole ensemble of simulations, no bias towards higher or lower correlations due to the spatial domain considered (GL, NH or SH) can be noticed. Nevertheless, slightly larger values are often obtained for the SH within the nonPMIP3 suite of simulations. The PMIP3 runs exhibit however somewhat lower values compared to the other spatial domains. The correlation values obtained between temperature and the forcing applied are, in general, considerably high and independent on whether they are PMIP3 runs or not and even on the solar scenarios they consider. The influence of the internal variability is visible in the distribution of correlation values in Figure 4.6 when different runs of the same model are analysed. In particular for the EC5MP ensembles, the differences between the correlation values among ensemble members are larger than the statistical uncertainties associated to each correlation value. Minimum correlation values are attained for the CNRM, particularly for the SH, for the NH in some of the sim-

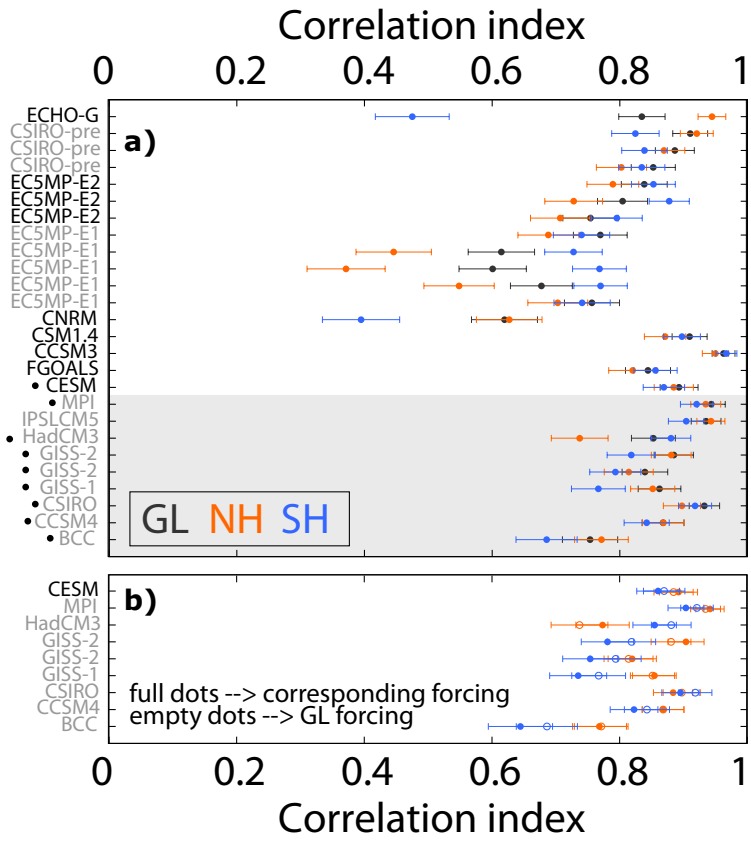

Fig. 4.6: Correlation indexes between the 31-yr moving average simulated temperature and its corresponding TEF for each simulation and for different spatial targets (GL in dark gray; NH in orange; SH in blue). Top panel (a) shows the correlation values obtained using invariably the GL TEF. Correlation indexes obtained for nonPMIP3 simulations are plotted over a white background, while the gray shaded area corresponds to the PMIP3 runs. The name of the simulations driven by a high variability solar forcing scenario (STSI) are written in black while those using a low variability solar forcing scenario (ssTSI) are in gray. A black dot tagging the model name indicates that this model has available TEF configuration at the hemisperic scale (see Chapter 3) and the correlation indexes are therefore calculated considering the same spatial domain in the simulation and in TEF. Bottom panel (b) shows, for the simulations tagged with the black dot in a), the correlation values obtained between the global TEF values (empty dots) and the hemispheric corresponding TEF (full dots).

ulations of the EC5MP-E1 ensemble and for the SH in the ECHO-G run. In the case of the CNRM there may be at least two reasons that contribute to this. On one hand, the temperature response of the model is not proportional to the large volcanic forcing applied if compared to other model experiments (see Figures 3.4, 4.1 and 4.4). On the other hand, this linear relationship to the forcing may be reduced by a relatively high internal variability and strong feedbacks in atmosphere dynamics (Swingedouw et al., 2010). The latter, together with an existing drift in the SH, may be the cause of the low values for the ECHO-G simulation. The EC5MP-E1 ensemble shows correlations for the NH in the range of 0.38 to 0.71, highlighting again the large influence of the different initial conditions, i.e. internal variability, on the temperature response at hemispheric scales.

Chapter 3 raised the possibility of calculating an hemispheric TEF (Figure 3.8) for some of the simulations including recent volcanic forcing estimations with latitudinal distribution of aerosol loading (Gao et al., 2008; Crowley and Unterman, 2012). Figure 4.6b shows for these simulations the correlation values obtained between simulated temperatures for the hemispheric domains and the corresponding external forcing calculated over identical domain. The correlation in the case of the global TEF is also shown (empty dots) for comparison. Overall, the differences between correlation values for the various spatial domains are not significantly larger than the statistical uncertainty for each simulation. It can be observed that for the SH temperature, slightly higher values are obtained when considering the GL TEF, while for the case of the NH, higher values are observed when considering the corresponding NH TEF. Nevertheless these differences between correlations calculated with the GL or the corresponding hemispheric TEF are not significant ($\alpha = 0.05$). The lack of a clear improvement when considering analogous domains for the TEF and the temperature response may be related to the fact that hemispheric estimates are not considering specific information at these scales for forcings other than volcanic. A sound relation between hemispheric temperature response and the global external forcing has been evidenced allowing for estimates of the relation between temperature and forcing not only at global but also at hemispheric scales, even in those situations where the hemispheric forcing is not available. i.e., nonPMIP3 simulations.

The high correlation values between the temperature simulations and the TEF in Figure 4.6 suggest that a considerable amount of the temperature response to the external forcing is linear. This relationship is also illustrated in Figures 4.7 and 4.8 where regression estimates (black in Figures 4.7 and 4.8) calculated from a simple linear regression between simulated GL temperatures (green) and TEF for all simulations in the nonPMIP3 and PMIP3 groups are shown. Grey shading represents the uncertainty associated with the regressed estimates. Therefore, the GL temperature response can be considered, to a good approximation, lin-

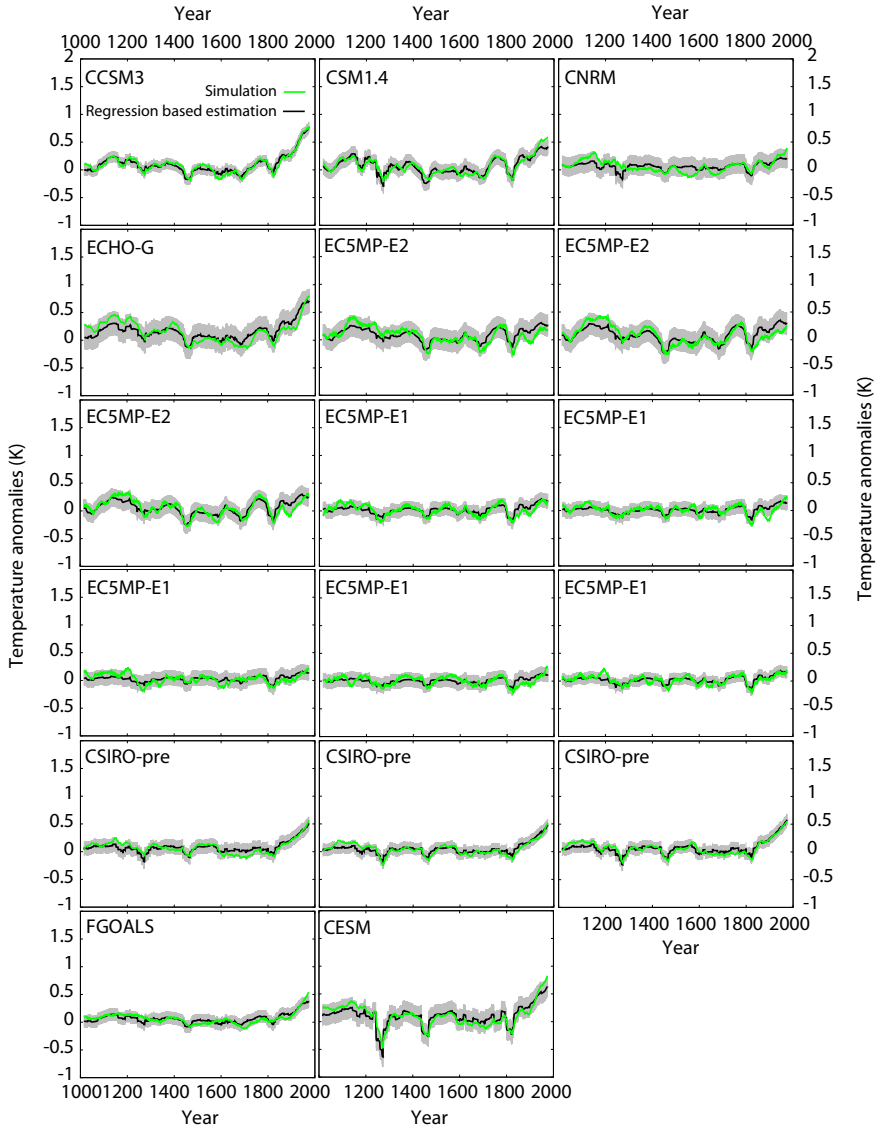

Fig. 4.7: Simulated GL mean temperature (green line) and the corresponding linear regression estimates (black) obtained from the global TEF. Calculations are based on 31-yr moving averages.

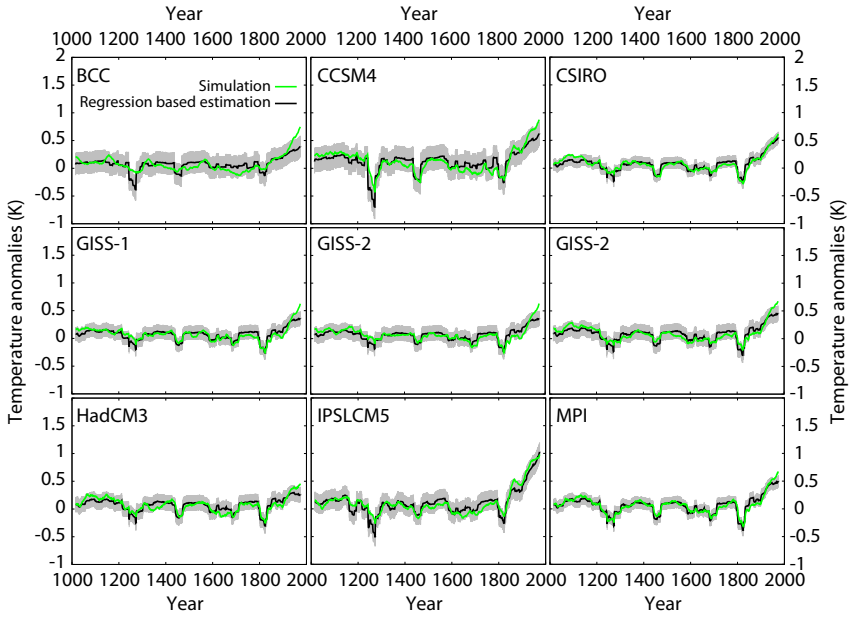

Fig. 4.8: As in Figure 4.7 but for the PMIP3 simulations.

early related to the imposed TEF forcing at these timescales. Similar results are obtained also for the NH and SH (not shown).

The agreement between simulations and estimated temperatures is evident for all models, i.e. the simulations tend to fall well within the estimated uncertainty band, except for the most recent period in some cases. These slight discrepancies are found in some models, mainly for simulations from the PMIP3 group (Figure 4.8), for the last half-century, when the warming trends of the simulations are more pronounced than those obtained from the regression. Noticeably, these disagreements occur just for the simulations that include anthropogenic aerosol forcing (see Table 2.1), likely evidencing some limitation in the approximation of the anthropogenic aerosol forcing assumed for the TEF calculation.

The discrepancies observed, however, are limited to a short time span (ca. 50 years) that is not unique in the context of the complete period analysed, i.e. the LM. This is supported by the temperature residuals shown in Figure 4.9, obtained as the differences between the global averages and the regression-based estimates from Figures 4.7 and 4.8. Figures 4.9a and b divide the residuals according to the STSI and ssTSI groups, respectively. There exist a significant

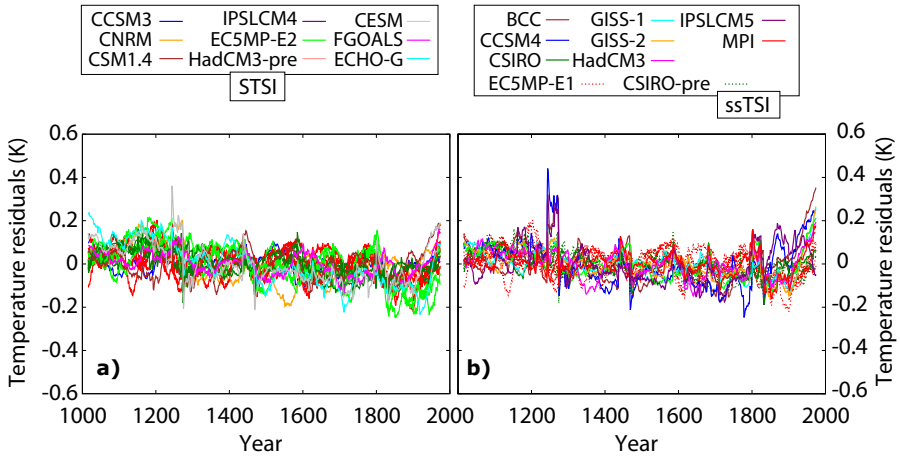

Fig. 4.9: Temperature residuals obtained from substracting the global temperature averages from the regression-based estimates. Panels a) and b) divide the residuals according to the STSI and ssTSI groups, respectively.

discrepancy around ca. 1258 CE on the models that applied the Gao et al. (2008) volcanic forcing, related to the fact that the temperature response saturates for large volcanic eruptions (Timmreck et al., 2009). Nevertheless, in general these residuals resemble a white noise signal, particularly in pre-industrial times, i.e. the series are centered around zero and there is not any significant long-term trend. This means that the assumption of a linear relationship between TEF and temperature is consistent during the whole period. For the last century, however, a slightly positive trend is observed in several simulations, in agreement with the discrepancies observed in Figures 4.7 and 4.8 discussed above. Nevertheless, these discrepancies confined within a short period in the context of the LM do not invalidate the overall linear relationship between TEF and temperature.

4.3 Last Millennium Transient Climate Response

The sensitivity of the climate variability to changes in external forcing can be characterised in the context of future climate by two quantities (see also Section 2.1): equilibrium climate sensitivity (ECS) defined as the temperature change, after reaching equilibrium, due to a doubling of atmospheric CO_2 above pre-industrial levels (Schneider et al., 1980); and the transient climate response (TCR), defined as the change in global surface temperature in a 1% CO_2 increase

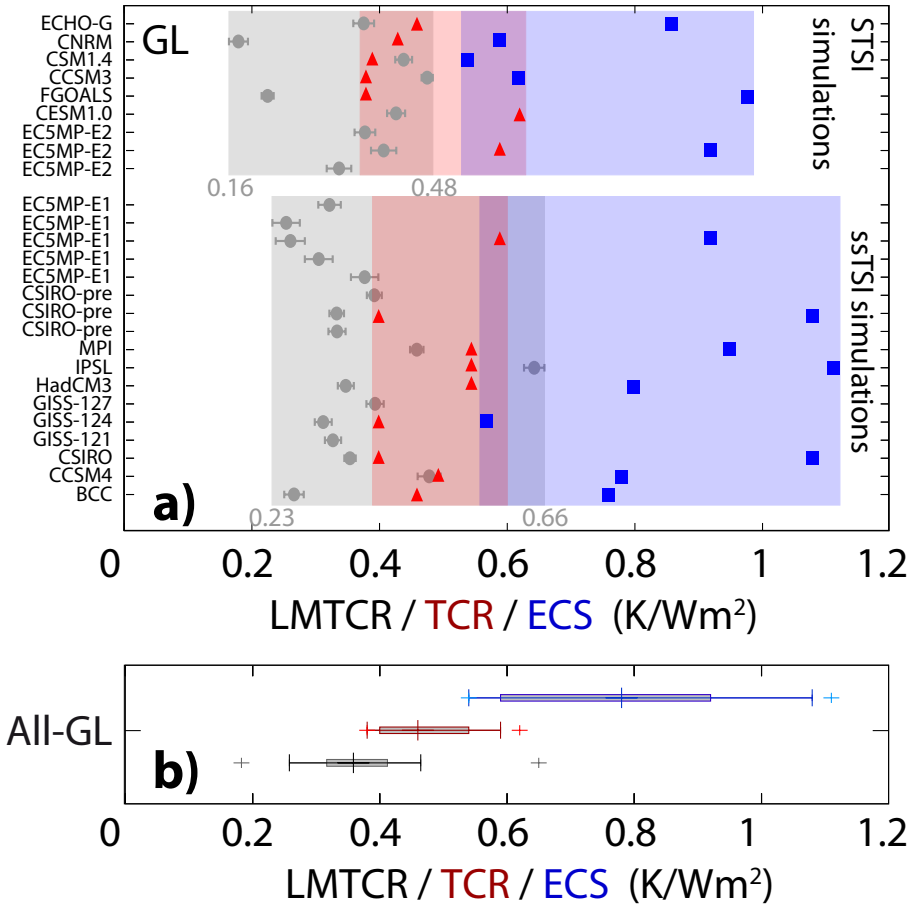

Fig. 4.10: a) LM Transient Climate Response (LMTCR, grey) values for the STSI and ssTSI simulations estimated from GL temperatures. Transient Climate Response (TCR, red) and Equilibrium Climate Sensitivity (ECS, blue) are also shown for each model and obtained from original references (Chapter 2). Colored backgrounds indicate the range of existing values, based on the minimum and maximum estimates, for each variable. b) Box-whiskers plots of the LMTCR, TCR and ECS values obtained from the ensemble of simulations of a). Maximum and minimum values are indicated with a cross, 10, 50 and 90 (25 and 75) percentiles are represented by the whisker bars (boxplots).

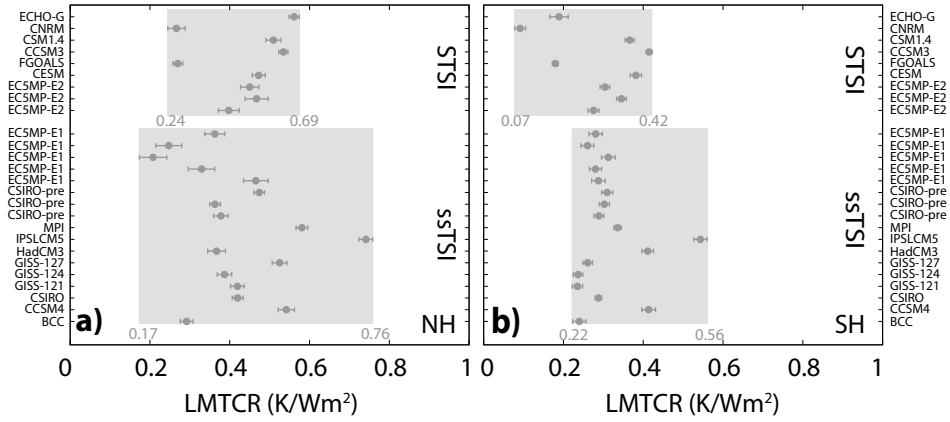

Fig. 4.11: LM Transient Climate Response (LMTCR) values for the STSI and ssTSI simulations for the NH (a) and SH (b) spatial domains.

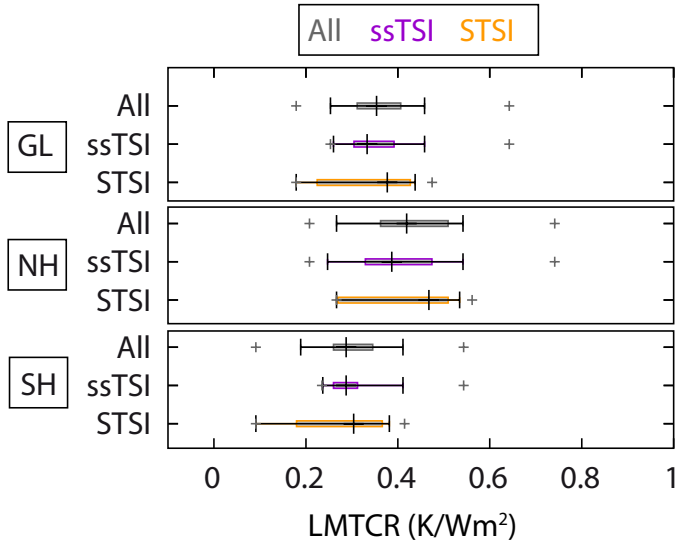

Fig. 4.12: As in Figure 4.10b but for the three different spatial domains, i.e., GL, NH and SH.

experiment at the time of atmospheric CO_2 doubling (Knutti et al., 2005). Here, we define a new quantity that measures the response of temperature to changes in external forcing during the LM. This quantity will be based on regression estimates between TEF and temperature in transient full forcing simulations of the LM. This metric will be extended in Chapter 5 also to temperature reconstructions of the CE.

Based on the linearity of the relationship between external forcing and temperature assessed in the previous section, the analysis is expanded to calculate the rate of changes in temperature relative to forcing. We denote the rate of linear change of temperature as a response to the external forcing as the LM transient climate response (LMTCR hereafter). These values correspond to the linear regression coefficient between temperature and forcing and represent the estimations of the climate sensitivity that integrate the response of the climate system to different forcings operating from multidecadal to multicentennial timescales. It should be kept in mind that, because of its definition being based on linear regression, LMTCR addresses only the quasi-instantaneous response of temperature to forcing changes. Any non linear feedbacks or delayed adjustments of temperature do not fall into this definition. In turn, TCR and ECS can account for non linear and delayed responses in the system in a process of continuous warming during several decades (TCR) or in a sudden warming due to CO_2 doubling and subsequent relaxation to equilibrium.

Figure 4.10a shows the LMTCR values with estimated regression errors obtained from simulated GL temperatures and TEF. Results for the models and forcings within the STSI and ssTSI groups are shown on the same panel and compared to the available TCR and ECS values of the different models (Table 2.2), that also refer to GL temperatures. We only consider herein simulations that include, at least, a minimum set of external forcing factors such as solar variability, volcanic and GHG estimations (Table 2.1).

We establish ranges for LMTCR based on the minimum and maximum values obtained within the STSI and ssTSI groups. The resulting range of LMTCR values spanned by the ssTSI group of simulations is wider (0.23-0.66 K/Wm^2) compared to the STSI ensemble (0.16-0.48 K/Wm^2), although the LMTCR values that fall in the lower end of the range are comparable in both cases.

It is interesting to note that LMTCR estimates do not change when grouing the simulations into STSI and ssTSI, since for instance, for the EC5MP is observed an overlap of the LMTCR estimates when considering the STSI (EC5MP-E2) or ssTSI (EC5MP-E1) experiments, as expected since these LMTCR values come from the same model (EC5MP) characterized by the same climate sensitivity.

The distribution of all LMTCR, TCR and ECS values in Figure 4.10a is represented in the box-whiskers plot of Figure 4.10b. The LMTCR ranges are, as expected, lower than ECS values and are also lower than or they overlap with the range of TCR values. This is a reasonable feature since, as commented above, TCR and particularly ECS, include system readjustments that involve nonlinear relationships, either in a monotonously warming climate simulation (TCR) or in equilibrium simulations (ECS), in contrast to the quasi-instantaneous response considered in LMTCR. The difference observed between the LMTCR and the other sensitivity values could be considered, in a sense, as the amount of response to TEF changes that is not linear. Attending to the latter and based on the mean LMTCR value of ca. 0.4 K/Wm^2 and the 0.8 K/Wm^2 corresponding to the ECS, it can be argued that about half of the ECS could be attributted to linear processes.

Figure 4.11 shows, following Figure 4.10, the individual LMTCR values estimated from simulated NH and SH temperatures and GL TEF. In general the LMTCR values obtained for each simulation are lower for the SH than the NH, i.e. based on the LMTCR estimates, the SH evidences a weaker response than NH to the external forcing. The latter seems reasonable since LMTCR evaluates the quasi-instantaneous response of the system to the forcing and the SH contains a larger proportion of oceanic areas relative to continental regions that can induce a delayed response. Additionally, LMTCR values are more clustered in the SH than in the NH. As in Figure 4.10b, box-whiskers plots showing the distribution of LMTCR values for the three spatial domains are also represented in Figure 4.12.

If the 10-90 percentiles are considered, the LMTCR ranges described by the two ensembles of simulations, STSI and ssTSI, are quite similar. The main differences found in Figures 4.10 and 4.11 can be associated to the minimum and maximum values. Specifically, IPSLCM5 evidences the largest LMTCR values (Figures 4.10 and 4.11) within the ssTSI group while the CNRM presents always the lowest values within the STSI ensemble.

It is noteworthy that the range of values spanned by the STSI and ssTSI groups of simulations is to a reasonable extent comparable to the interval of LMTCR values expanded in the case where no segregation according to the solar forcing scenario is considered (grey in Figure 4.12). Therefore the mean LMCTR based on all simulations yields a robust estimate of LM climate sensitivity provided that this statistic is calculated over a larger population compared to the cases where the simulations are divided depending of their corresponding solar forcing factor.

All the spatial scales (GL, NH and SH) render positive LMTCR values, i.e. models show a positive and quasi-instantaneous response to TEF at mul-

tidecadal and longer timescales by the positive correlations in Figure 4.6. Ranges of LMTCR estimates for GL, NH and SH are consistent and not significatively different ($\alpha = 0.05$). The mean response is found to lie between 0.3 and 0.4 K/Wm^{-2}, with the largest individual LMTCR values being obtained for the NH (ca. 0.7) and the lowest for the SH (ca. 0.1).

4.4 Conclusions

The temperature response of the suite of LM AOGCM simulations has been assessed. All simulations agree in showing a similar temperature evolution during the LM with a warm period during the MCA, followed by a colder LIA and the recent warming during the industrial period. Despite all the differences observed among the simulations, mostly attributable to internal variability, there is a common signal related to a response to the external forcing applied. Indeed, all the experiments agree on evidencing a multidecadal variability during preindustrial times driven by the solar and volcanic activity. In turn, the simulated temperature trends during industrial times are a direct response to the set of forcing configurations applied to each model during this period.

Despite the role of the internal variability, at multidecadal and longer timescales the simulated temperature response evidences a robust linear relationship with the forcing imposed to drive the simulation. The simulated temperature response can be reproduced by linear regression estimates between the temperature and the TEF series. We examined the system response to the external forcing based on regression estimates of the rates of temperature-to-forcing changes, the so-called LMTCR. This estimate of LM climate sensitivity does not consider other system feedbacks and delayed responses that are implicit in the definitions of ECS and TCR in future climate change experiments. In fact, the ranges of calculated LMTCR are always lower than ECS and at the most overlap in some cases with TCR.

LMTCR values are positive, for all simulations and the different spatial scales considered (GL, NH and SH), evidencing that all the models show a positive quasi-instantaneous response to forcing changes. This response is quantified by the mean LMTCR value from all experiments whose value is around 0.3-0.4 K/Wm^{-2} for GL, NH and SH scales. LMTCR values tend to be larger for the NH than for the SH, related to the larger amount of oceanic areas in the SH that induces therefore a delayed temperature response to the forcing. Interestingly, LMTCR values obtained for the ssTSI and STSI groups are quite similar, suggesting that no significant differences in the response of the models due to the different level of solar variability forcing applied can be identified.

A limited number of works have also used the past temperature variations to assess the climate sensitivity. Hegerl et al. (2006) estimated the climate sensitivity of the system based on LM temperature reconstructions and arrived at values compatible with the LMTCR values estimated herein. Alternatively, the Last Glacial Maximum and the transition to the Holocene has been more often explored to analyse the temperature change recorded in proxies and to identify the external forcing imprint with the aim of constraining the ECS estimates (see Masson-Delmotte et al., 2013, for a review). The LM evidences comparatively smaller values of climate sensitivity, i.e., the response to the external forcing is weaker if compared with longer timescales where the orbital forcing changes are larger and the base climate state is different, i.e. glacial. Nonetheless, the analysis presented in this chapter provides evidences of a detectable and measurable response of the climate system to the external forcing during the LM. Next chapter will assess the temperature variation associated to changes in the forcing for the LM reconstructions.

5

Reconstructed temperature response*

Chapter 4 analyzed the temperature response of the AOGCM/ESM simulations during the LM, characterizing their temporal evolution and assessing their relationship with the external forcing configurations applied. This chapter complements this information by performing this analysis with temperature reconstructions of the CE. An special focus will be placed on special focus on the spread of the ensemble of reconstructions and our understanding of ensemble uncertainty. As introduced in Chapter 1 (Figure 1.2), this ensemble of reconstructions and their uncertainty represents the current state of knowledge of the temperature evolution in the real climate system. Additionally to the characterization of uncertainty for the global and hemispheric scales, a more detailed analysis of factors contributing to the ensemble spread is provided.

A better understanding of the LM reconstructed temperatures may allow for insights on the relative influence of external forcing and internal variability within the system. Thus, considering the available estimations of TEF applied to the AOGCMs/ESMs in Chapter 3 as the best available estimates of the real external

* The main contents of this chapter are included in:
- Fernández-Donado, L., J. F. González-Rouco, C. C. Raible, C. M. Ammann, D. Barriopedro, E. García-Bustamante, J. H. Jungclaus, S. J. Lorenz, J. Luterbacher, S. J. Phipps, J. Servonnat, D. Swingedouw, S. F. B. Tett, S. Wagner, P. Yiou and E. Zorita, 2013: Large-scale temperature response to external forcing in simulations and reconstructions of the last millennium. *Climate of the Past*, **9**, 393-421. DOI 10.5194/cp-9-393-2013.
- Fernández-Donado, L., J. F. González-Rouco, E. García-Bustamante, J. E. Smerdon, S. J. Phipps, J. Luterbacher and C. C. Raible, 2015: Northern Hemisphere temperature reconstructions of the Common Era: Ensemble uncertainties and their influence on model-data comparisons. *Geophysical Research Letters*, **In review**

forcing, a quantitative assessment of the temperature response to the forcing is proposed also for reconstructions at global and hemispheric.

The chapter is structured as follows. Section 5.1 presents the temporal evolution of the ensemble of reconstructions considered, including the individual series but also the associated uncertainty band. Within this section, 5.1.1 analyses, for the NH, the impact of several factors contributing to the ensemble uncertainty. Section 5.2 analyses the relationship between temperature reconstructions and the available external forcing configurations. Finally, Section 5.3 summarizes the main results and conclusions obtained within the chapter.

5.1 Reconstructed global and hemispheric temperatures during the Common Era

Similarly to Figure 1.2c, Figure 5.1 shows the temporal evolution of the collection of large-scale temperature reconstructions detailed in Table 2.3 over the CE. Specifically, Figure 5.1a,b,c shows the temperature anomalies for the NH, SH and GL scales, respectively, using 1850-1990 CE as the reference period. This interval encompasses most calibration periods used for the reconstructions. Therefore, the spread before the 19th century is considered as a measure of uncertainty in the estimation of past temperatures.

Figure 5.1a evidences the larger availability of NH temperature reconstructions, most of them spanning the LM but also with some series covering the last2k. In general, reconstructions in Figure 5.1a show a broad agreement among them as discussed in Section 1.2, depicting a warm MCA extending between ca. 900 to 1250 CE, followed by a colder LIA during ca. 1400 to 1700 CE and a subsequent warmer industrial period. Temperatures in the first half of the 20th century are comparable to those in the MCA in most of the series except for Christiansen and Ljungqvist (2012b) and Loehle and McCulloch (2008). The extratropical NH reconstruction of Christiansen and Ljungqvist (2012b) displays larger low frequency amplitude changes than any other reconstruction in the ensemble, noticeably enlarging the spread during the MCA and mid-20th century. As in the case of Christiansen and Ljungqvist (2011), reconstructions using the LOC method aim to preserve low frequency variability, perhaps to the detriment of high frequencies (Christiansen, 2011). An overestimation of variability in this method can not be ruled out and some studies suggest these reconstructions may be taken as an estimation of maximum bounds for low frequency amplitude changes during the LM (Chapter 1; Masson-Delmotte et al., 2013; Tingley and Li, 2012; Moberg, 2013; Christiansen, 2012; Christiansen and Ljungqvist, 2012a). This reconstruction shows noticeably more variance at low frequencies not only in the pre-industrial period but also during the 20th century.

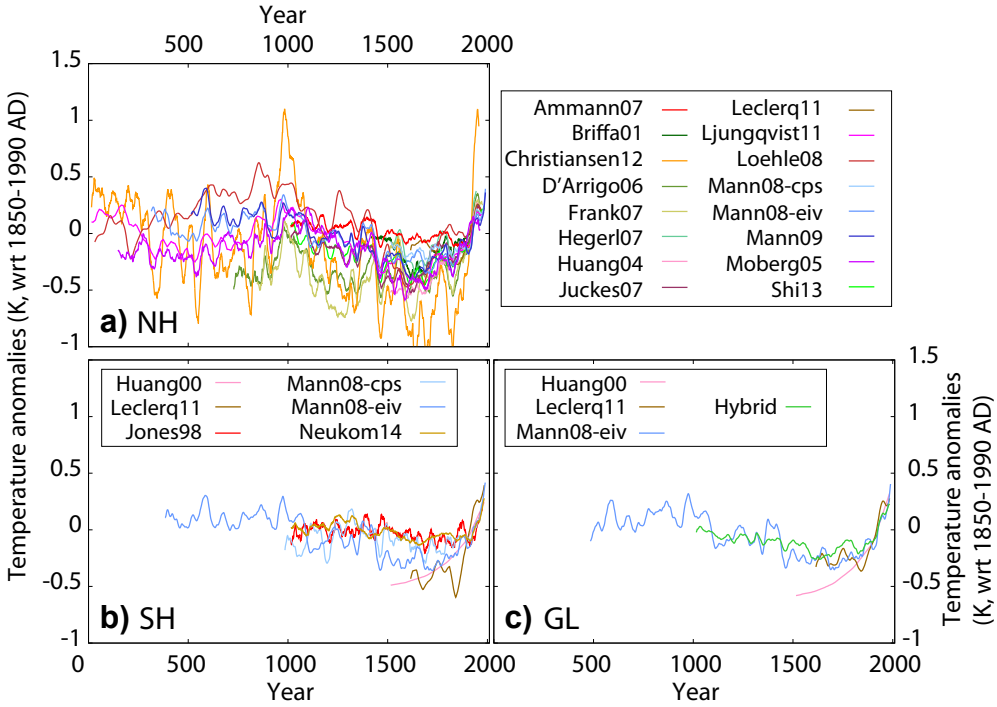

Fig. 5.1: NH (a), SH (b) and GL (c) temperature anomalies wrt 1850-1990 CE from reconstructions listed in Table 2.3. Series are 31-yr moving average filter outputs.

In contrast to the NH, SH and GL scales gather very few reconstructions. Figure 5.1b presents an update of the assessment of the SH temperatures from Masson-Delmotte et al. (2013) that additionally includes the Neukom et al. (2014) reconstruction. The six existing records reflect very similar multicentennial trends, whereas multidecadal variability can be quite different among the records. It should be kept in mind that three of the records considered share identical or overlapping proxy information or use comparable methods as in the case of the CPS approach (Jones et al., 1998; Mann et al., 2008). The evidence shown in Figure 5.1b is suggestive of relatively higher temperatures before the 15th century, comparable to those at the beginning of the 20th century and a colder interval spanning the period between the 15th and 19th centuries. Prior to 1000 CE, the only available evidence (Mann et al., 2008), suggests warmer temperatures than in any other time interval over the preindustrial period. In

turn, at global scales (Figure 5.1c) only four records are available, including the hybrid reconstruction developed within the present work (see Chapter 2). There is a broad agreement among these series over the common period in representing multidecadal and multicentennial variability.

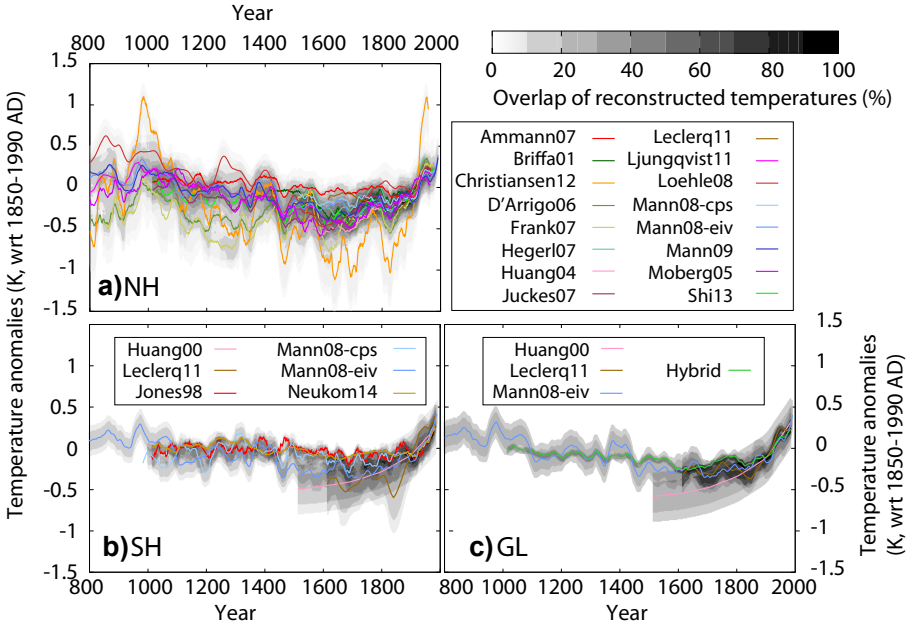

Fig. 5.2: NH (a), SH (b) and GL (c) reconstructed temperature anomalies wrt. 1850-1990 CE (colors; see Table 2.3. The grey shaded areas on the background are the overlap of uncertainty calculated from the errors provided with each reconstruction and following Masson-Delmotte et al. (2013). SH and GL are shown for the sake of illustrating the available information for these domains. Note that the small number of reconstructions for SH and GL precludes a reliable estimation of ensemble uncertainties. All series are 31-yr moving average filtered ouputs.

Figure 5.2 complements Figure 5.1 by showing, together with the individual reconstructed series, a measure of the uncertainty using the overlap of the series. This is represented by the gray shading and constructed as in Jansen et al. (2007), Fernández-Donado et al. (2013) or Masson-Delmotte et al. (2013). For each year, the temperature axis is divided in 0.01°C pixels that receive a score of 1 (2) if they lie within the range of the reconstructed temperatures ±1.645

(±1) standard deviation (STD). The scores are summed over all reconstructions considered and scaled to range within 0 and 100% of overlap. Figures 5.2 shows the reconstructions and their uncertainty bands for the NH, SH and GL scales over the period 800 to 2000 CE, with a larger availability of series and that also allows comparing climate simulations in Figures 4.1 and 4.4. Figures 5.2b,c are shown only only for the sake of presenting the available information at the SH and GL scales, but the estimation of their uncertainty bounds is limited by the very few reconstructions available. On the contrary, Figure 5.2a offers a more reliable estimation with a diversity of proxy types and methods that offers a more robust metric of uncertainty in the knowledge of the temperature evolution. Indeed, the resulting distribution could be understood as the best estimate of the temperature evolution during the LM. The maximum of overlap is obtained during the second half of the LM, indicating that the agreement among reconstructions is higher and the uncertainty is more constrained. In turn, the first half is characterized by lower levels of overlap illustrating lower agreement within the reconstruction ensemble as evidenced in Figure 5.1a. This larger spread in NH temperatures during the MCA was partially attributed to a reorganization of spatial climate patterns (Jansen et al., 2007) and argued later to be associated with low sample replication of proxies rather than regional heterogeneity (Esper and Frank, 2009).

As it was already mentioned in Chapter 1, there can be multiple contributions to the spread of reconstructions. The following section offers more insights about the spread of NH temperature reconstructions in Figure 5.2a and about the factors contributing to it.

5.1.1 Ensemble uncertainties

The relative influence of different methods and proxy quality on reconstructed NH temperatures has been largely assessed by means of pseudoproxy experiments (e.g. Mann and Rutherford, 2002; Buerger et al., 2006; Von Storch et al., 2004; Mann, 2007; Lee et al., 2008; Lehner et al., 2015). Therein, AOGCM simulated LM temperatures are used as a surrogate for the real past temperature evolution and allow for systematic experiments testing the influence of proxy network characteristics on the performance of reconstruction methodologies (see Smerdon, 2012, for a review). Additionally, a few real-proxy reconstruction efforts (Buerger and Cubasch, 2005; Rutherford et al., 2005; Esper et al., 2005) analyse the sensitivity of methods to various decisions during the design and training of employed statistical reconstruction methods. One example is the use of different seasonal windows as targets to accommodate potential seasonal biases in the proxy signal (Jones et al., 2009), such as warm season biases in the case of tree-ring based proxies. This effect may be important depending on the mixture of different proxy archives in multi-proxy approaches (e.g. Pauling et al., 2003) and their influence

on hemispheric-scale temperature reconstructions is open to debate (e.g. Rutherford et al., 2005; Phipps et al., 2013). Assessments that have tested the impact of changes in the target spatial domain using a single reconstruction method suggest comparatively more substantial changes in the variability of the reconstructed temperatures (e.g. Rutherford et al., 2005; Esper et al., 2005). In fact, the currently available ensemble of reconstructions involves a variety of targets that include either the complete NH (e.g. Moberg et al., 2005), or focus on NH land only areas (e.g. Shi et al., 2013) or NH land-only extratropical regions (e.g. Briffa et al., 2001) as it was evidenced in Table 2.3. The distribution of the proxy records is also relevant and is largely biased to continental areas, with relatively few proxies distributed over coastal and ocean regions (Mann et al., 2009). This has a potentially large impact on regional skill, particularly if climate field reconstruction methods are applied (Smerdon et al., 2011, 2015). Finally, the selection of calibration/verification intervals also has a considerable impact on the reconstructed variance (Buerger et al., 2006; Frank et al., 2010b).

The use of different methods, target domains and calibration/verification intervals can impact the spread of the ensemble of NH temperature reconstructions. Understanding and, if possible, narrowing uncertainties in the ensemble of LM reconstructions is relevant for improving our knowledge of mechanisms contributing to climate variability (Goosse et al., 2012b). Also, improved estimates of the range of reconstructed temperature responses to external forcing from decadal to multi-centennial timescales are relevant to model evaluation and as constraints on climate sensitivity estimates (Fernández-Donado et al., 2013; Hegerl et al., 2006). This section addresses these issues and specifically the question of whether sound conclusions can be drawn from comparisons among reconstructions in the absence of uniform criteria focused on the design and training of the reconstruction model.

Figure 5.3a updates Figure 5.2a distinguishing (linestyles) the various spatial targets included within the NH series (Table 2.3), following Masson-Delmotte et al. (2013). Within the NH two main groups are determined: the land-only (NH-LO); and the complete NH (NH-C). Additionally, some of the reconstructions target the temperature signal of extratropical regions whereas others include the tropics. Within the NH-LO ensemble, a NH-LOE subgroup identifies those reconstructions targeting the land-only extratropical domain (Table 2.3). There is no subgroup within NH-C referring to extratropical-only series as only the reconstructions of Ljungqvist (2010) and Christiansen and Ljungqvist (2012b) fulfill this definition. Thus, considering these target subgroups, Figure 5.3a shows in dark (light) green the NH-LOE (NH-LO) series and in red/pink the NH-C ones. The NH-LO reconstructions consistently indicate lower temperatures during the pre-industrial period. The highest overlap is attained after 1500 CE, when the NH-C and the NH-LO subgroups show a more comparable range of cooling. The

reconstruction of Christiansen and Ljungqvist (2012b, dashed purple line) shows a remarkably different behavior to all NH-C counterparts depicting the largest amplitude of decadal to multi-centennial changes in the ensemble (consistent with previous documentation, e.g. Masson-Delmotte et al., 2013).

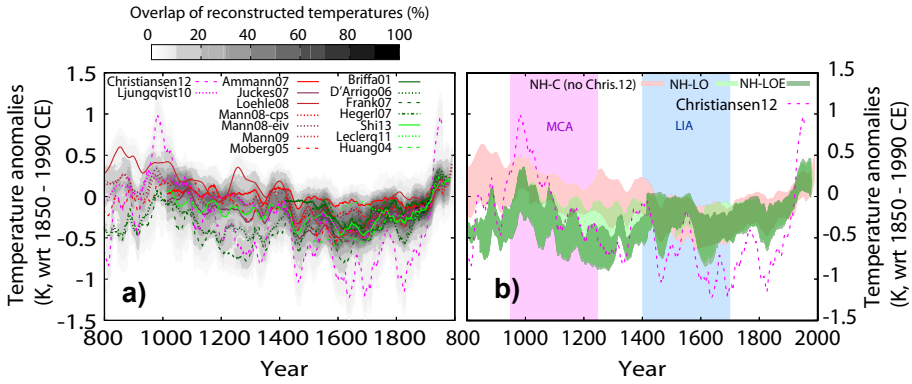

Fig. 5.3: (a) NH temperature anomalies with respect to the 1850-1990 CE average for the reconstructions specified in Table 2.3. Complementary to Figure 5.2a this representation emphasizes the different reconstruction targets: NH-C (red/pink); NH-LO (green); NH-LOE (dark green). The gray shaded areas represent the overlap of the reconstructions uncertainties as in Figure 5.2. (b) As the gray shading in (a) but segregated according the original spatial target in Table 2.3. The red (light/dark green) shaded area stands for the 10-90% range of the ensemble uncertainty of reconstructions assigned to the NH-C (NH-LO/NH-LOE). The reconstruction from Christiansen and Ljungqvist (2012b) (purple-dashed) has not been used in the calculation of uncertainties. All series are 31-yr moving average filter outputs.

The ensemble of reconstructions is further separated for each spatial target in Figure 5.3b. The reconstruction of Christiansen and Ljungqvist (2012b) is not included in the calculations of the NH-C uncertainty envelope, as it potentially overestimates low-frequency variability (Moberg, 2013).

The uncertainty bands in Figure 5.3b reveal differences in the range of temperature changes corresponding to the three target domains. The NH-LOE displays higher multi-centennial scale variability and presents the largest cooling since 800 CE. The largest differences among the three groups are found between the 9th and the 15th centuries during which NH-LOE is on average about 0.5 K cooler

than NH-C. It is also noteworthy that the width of each uncertainty envelope (NH-C, NH-LO and NH-LOE) is steady throughout the whole period (Figure 5.3b) in contrast to the case in Figure 5.3a, where the spread increases substantially between the 11th and 13th centuries. The differences between the MCA and the LIA change in the various subgroups, indicating larger MCA-LIA changes in the NH-C group with the MCA temperatures being more comparable to present times. In fact, both NH-LO and NH-LOE depict only a short period of relative warmth around 1000 CE.

These findings imply that comparisons between NH reconstructions that do not account for different methodological approaches and specifically different spatial targets may artificially inflate the spread of reconstructions during the CE and, in particular, the MCA. This implication is added to the previous findings regarding the potential influence of a more heterogeneous climate during the MCA or the impact of potential biases in the distribution of proxies (Rutherford et al., 2005; Jansen et al., 2007; Esper et al., 2005)

As discussed above, several factors can play a role in increasing the spread of the reconstruction ensemble, e.g. the selection of different methodologies and specific implementations of target variables and training periods. Figures 5.4 and 5.5 explore these influences by recalibrating all reconstructions using uniform methods, targets and calibration/verification intervals.

Figure 5.4 shows the impact of recalibrating each reconstruction to its original spatial target (Table 2.3) by using two simple methods: variance matching (VM; Jones et al., 1998; Frank et al., 2010b), thus having a tendency to inflate the variance of the original series at the expense of increasing the calibration error; and forward regression (FR; Lee et al., 2008), causing variance underestimation in the mean estimate (von Storch, 1999). While this decision underrepresents the richness of methodological variants in Table 2.3, it provides a fairly simple and common framework for analysing the variability of temperature reconstructions and serves as a reference for the subsequent analysis. The instrumental NH indices for the various target domains are obtained from the HadCRUT4 database (Brohan et al., 2006). All reconstructions are calibrated to the same period, i.e., 1865-1960 CE, except for Loehle and McCulloch (2008), since this series ends in 1935 CE. Prior to the recalibration, the effective temporal resolution of each reconstruction is identified by exploring its spectral density. In doing so, several series with annually resolved information were found to contain only variability at decadal and longer timescales (Chapter 2, e.g. Loehle and McCulloch, 2008; Mann et al., 2008)). Thus, the effective resolution of each reconstruction was emulated in the instrumental data with low pass filters prior to calibration.

The effect of scaling (VM) and regressing (FR) all reconstructions to their adscribed targets is shown in Figure 5.4a and Figure 5.4d, respectively. A de-

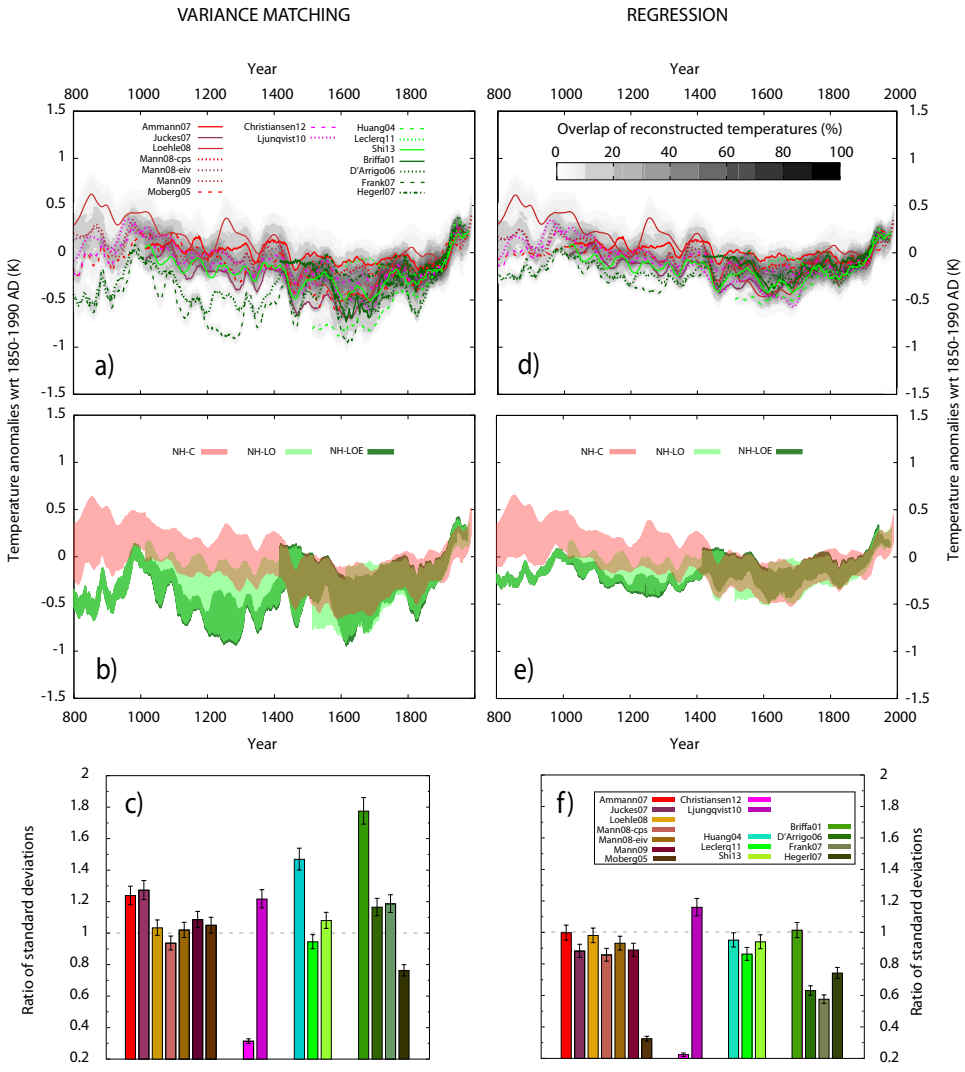

Fig. 5.4: (a, b) and (d, e) As in Figures 5.3a,b but for the VM (left) and FR (right) calibration methods, respectively. (c) and (f) show for each NH reconstruction the recalibrated *vs.* original STD ratio when the VM and FR methods are applied, respectively. All series are 31-yr moving average filtered outputs.

crease in the amplitude of the gray shaded area, i.e. the overlapping band, in Figure 5.4d with respect to Figure 5.4a is noticeable: as expected, the regression method induces a loss of variance with respect to the variance matching procedure. Additionally, the overall shape of the uncertainty envelope is comparable to that of Figure 5.3a.

Shaded bands in Figures 5.4b and 5.4e show the 10-90% range of uncertainty envelopes of all reconstructions segregated according to their original spatial target, as in Figure 5.3b and after applying the VM and FR based recalibration methods, respectively. The different ranges of variability of reconstructions along with their spatial targets observed in Figure 5.3b, particularly in the first centuries of the LM, are still present after the recalibration process. The differences are emphasized in the case of the VM method (Figure 5.4b). It is worth noting that the recalibration process does not equally affect the three spatial groups. Indeed, the spreads corresponding to the continental-only series (NH-LO and NH-LOE) are further influenced by recalibration with a sensible decrease of the width of the corresponding band, especially in the case of the regression method. This can be expected from the reduced variability of oceans relative to land due to their greater heat storage and evaporation capacity. This effect has been explored in observations (Sutton et al., 2007) and simulations for a range of past and future climate scenarios (Zorita et al., 2005; Lâiné et al., 2009; Byrne and O'Gorman, 2013).

The impact of each recalibration method on the individual variance of each reconstruction can be found in Figures 5.4c and 5.4f, where the ratio of the original *versus* the recalibrated STD is represented for each reconstruction and method. Most STD ratios in both panels distribute close to one. However, it is reasonable to expect some changes in the variance of the series after recalibration since the method applied in the recalibration of the series may differ from that originally used by each author. There are some cases where a systematic over/underestimation is identified in Figures 5.4c,f regardless of the recalibration method applied. The largest deflation of variance is found for the Christiansen and Ljungqvist (2012b) reconstruction both for VM and FR. Indeed, variance of this series is under debate (Moberg, 2013; Christiansen, 2012) and it has been noticed in the previous section that this reconstruction presents anomalous variance compared to the rest of the ensemble members. Based on this evidence, this reconstruction has been excluded from the uncertainty envelopes in Figure 5.3b. In addition to Christiansen and Ljungqvist (2012b), both VM and FR suggest that the reconstruction of Hegerl et al. (2007b) overestimates variance relative to the instrumental NH-LOE target period and deflate it accordingly by ca. 25%. Note that Total Least Squares is used in Hegerl et al. (2007b) which assumes some error (i.e. more variance) in the target, which is not considered in the VM

or FR cases. D'Arrigo et al. (2006) and Frank et al. (2007) are also deflated by regression by about 40%. Both VM and regression inflate Ljungqvist (2010) by ca. 20% and VM largelly inflates Huang (2004) and Briffa et al. (2001) by 40% and 80%, respectively. These changes basically have unaltered the spread of the NH-C group as a robust feature both for the VM and FR cases. The spread of land only groups in VM seem quite similar to Figure 5.3b while narrowing considerably after FR is applied. The segregation of NH-C and land only regions before the 15th century persists as a robust feature after VM and FR.

Figure 5.5 shows the sensitivity of recalibrating all reconstructions to the same target, using both VM and FR, as an illustration of the influence of each method and the variability associated with each target. Shadings stand to identify the original target of the recalibrated pool of reconstructions. Both scaling (VM) and regression (FR) to a single target produces larger spread for the continental extratropical case and the land only and shorter for NH-C. This illustrates that there is an effect of the target spatial domain in Figure 5.3a which needs to be taken into consideration; the narrowing of the uncertainty bands of NH-C is remarkable in comparison with Figure 5.3a. FR outputs for each of the three targets narrow the spread of the NH-LOE bands in all three recalibration exercises for the same reasons as stated when discussing Figure 5.4. From Figure 5.5 it is evident that all reconstructions spread over the same ranges of values after the 15th century. Nevertheless, the bias observed in the ensemble that implies different ranges of variability and the absence of overlap between the land-only and complete NH groups at the beginning of the LM remain when recalibrating reconstructions to a common spatial target.

Finally, a last contribution to spread is assessed by using different calibration periods. This is indicated in Figure 5.5 by using several lines for each reconstruction, each identifying the results for a different training period (see caption). The calibration period shows some relatively minor contribution to variability in the ensemble by inflating or deflating variance depending on the reconstruction and time period used.

Collectively, the results of the reconstructions indicate that method selection and training periods contribute to inflations or deflations of the variance of individual reconstructions, while the choice of target domain has a systematic bias towards increasing or decreasing the spread of the ensemble uncertainty. Regardless of this, however, the differences identified during the the last centuries of the first millennium and the first centuries of the LM persist independently from homogenizing methodological decisions.

During this time interval (800-1200 CE), the NH-C and NH-LO subgroups comprise 8 and 3 series, respectively. The reduced NH-LO subgroup is indeed formed by only two NH-LOE reconstructions (D'Arrigo et al., 2006; Frank et al.,

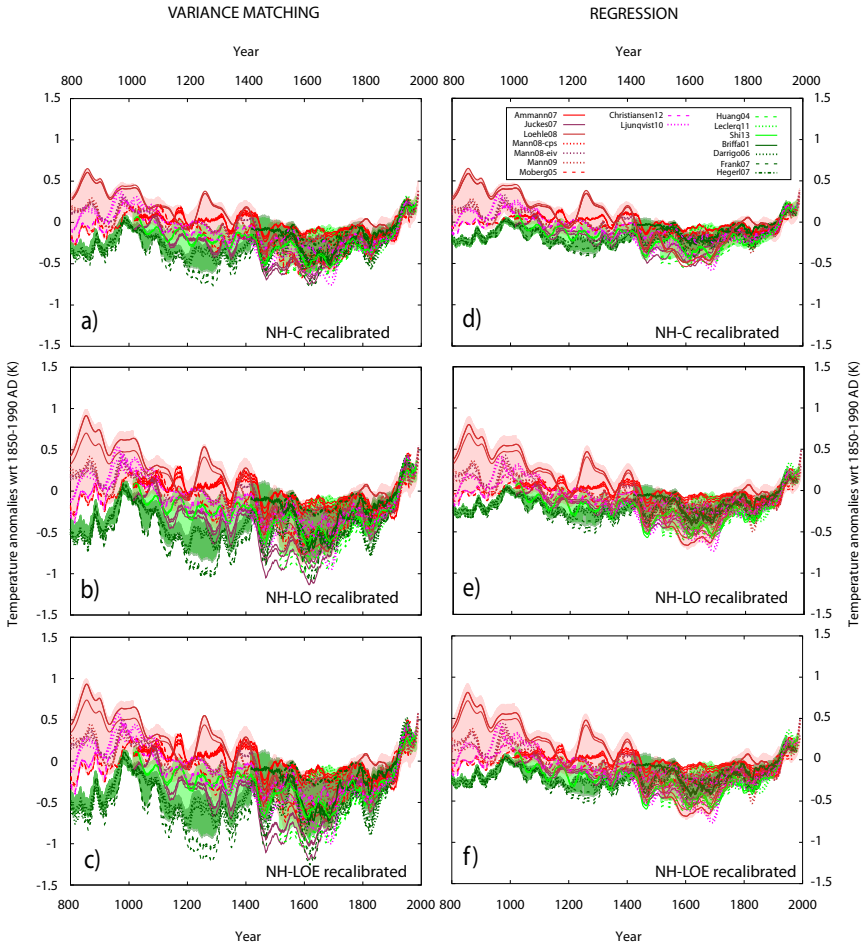

Fig. 5.5: As in Figure 5.3b shading represents the spread of NH reconstructions, after recalibrating to a common spatial target (see panels) using VM (left) and FR (right). Colours/shadings indicate the original target of the reconstructions. Lines represent individual reconstructions recalibrated to a common spatial target using as calibration periods: 1865-1960, 1865-1945 and 1880-1960. All series are 31-yr moving average filtered outputs.

2007) back to 1000 CE, both of them based exclusively on tree rings (Table 2.3). Indeed, during this early period of the LM, the two reconstructions share almost the same individual proxy records that are characterized by a diminishing level of replication (D'Arrigo et al., 2006) and relatively biased to high regions of North America and Eurasia. Thus, influences from the availability of the proxies and a regional bias cannot be excluded. The other NH-LO series within this period (Shi et al., 2013) is a multiproxy reconstruction that considers a larger amount of proxy records located over a wider area in the NH at the beginning of the second millennium and is in better agreement with the NH-C ensemble than the other NH-LO series. This therefore could be an indication that a larger variety of proxy archives (e.g. ice cores), distributed over a larger domain may reduce regional biases and lead to more comparable behavior of both NH sub-ensembles. Nevertheless, the reduced number of proxy records at the beginning of the LM is a common feature of all reconstructions. Even within the larger NH-C group where more proxy sites are considered, reconstructions share many of the same proxy records and regional biases cannot be ruled out. Proxy quality, abundance and sampling are thus still candidates that can account for the large differences between NH temperature reconstructions when partitioned into groups defined by different spatial targets.

While the differences discussed above can indeed be related to specific aspects of the reconstructions and the proxy information used to produce them, the influence of selecting different spatial domains as targets for the reconstructions has been shown to have an influence on the spread of the ensemble of reconstructions. The next section explores this effect in other simulated and observed sources of information.

5.1.2 Variability of NH temperatures: domain dependance

This section focuses on analysing the different level of temperature variability shown by simulations, instrumental series and a climate field reconstruction when considering different spatial domains within the NH. The objective pursued is twofold: 1) to support the evidenced of a larger level of variability in continental series found in the reconstructions; and 2) to explore whether the systematic difference found during the MCA between the various spatial domains in reconstructions is also evidenced by any other source of temperature information.

The ensemble of NH reconstructions analysed previously consists of hemispheric indices, except for the climate field reconstruction (CFR) of Mann et al. (2009). This allows the calculation of temperature averages for this specific CFR discriminating among the NH-C and the NH-LO and NH-LOE domains as in Figure 5.3b. Recall that this CFR does not extend over the complete NH, as northern latitudes and some tropical areas in Africa are not well represented. The series

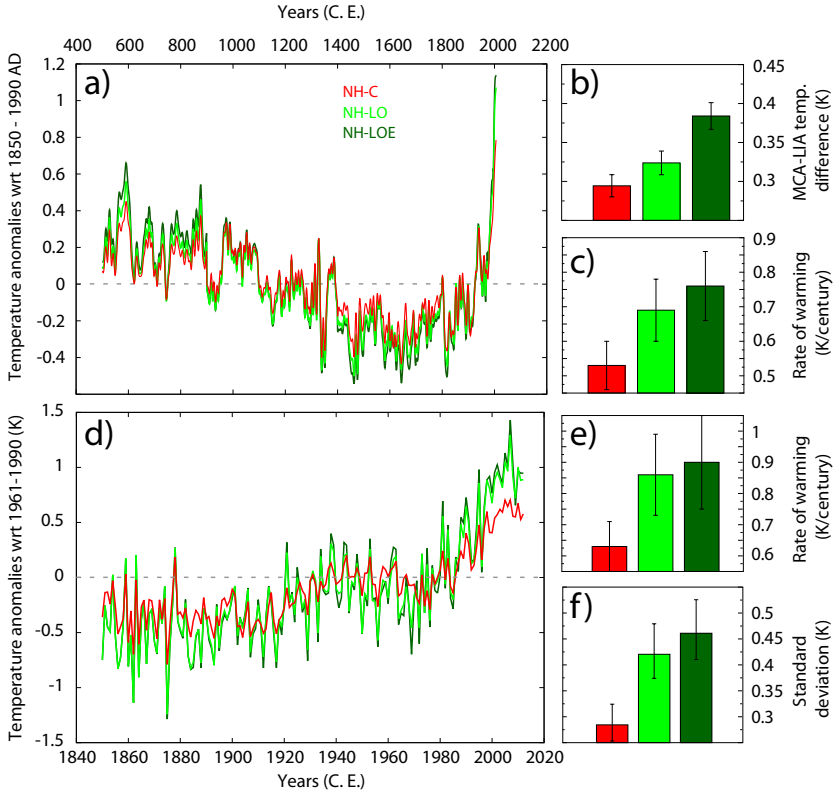

Fig. 5.6: Left panels: a) NH temperature anomalies wrt 1850-1990 CE from CFR reconstruction (Mann et al., 2009) averaged over the three regions corresponding to the spatial targets specified in this work, i.e. red for NH-C group, green for NH-LO and dark green for NH-LOE ensemble; d) as in a) but series are averages from the HadCRUT4 instrumental data (Brohan et al., 2006). Right panels: b) MCA (950–1250 CE)–LIA (1400–1700 CE) temperature difference for the three spatial target averages from Mann et al. (2009) CFR reconstruction in (a); c) the centennial rate of warming over each of the three target groups for the 1880-2006 CE period from the CFR reconstruction; e) as in c) but from HadCRUT4 data; and f) as in c) but the standard deviation of each target group averages from the HadCRUT4 data is represented instead. Color criteria of right panels as in a) and d).

averaged over the three domains (NH-C, NH-LO, NH-LOE) are shown in Figure 5.6a. Larger decadal to multi-centennial temperature variability can be found in the case of the land-only extratropical (NH-LOE) average, followed by the land-only series (light green shading in Figure 5.6a). As expected, the variance throughout the CE narrows in the NH-C (red) case with respect to the others. Figures 5.6b and 5.6c indicate the MCA-LIA temperature difference and the rate of warming during the last century, respectively, calculated also for the three different target averages from the reconstruction of Mann et al. (2009). Both panels confirm that the temperature range of variability is larger for the series that do not include the marine regions in their targets during calibration.

In turn, Figure 5.6d shows the most recent version of instrumental gridded data, HadCRUT4 (Brohan et al., 2006), averaged over the same three target groups, as in Figure 5.6a. Indeed, the averaged observations from HadCRUT4 behave similarly to the reconstructions, with the land-only regions, NH-LOE and NH-LO, showing larger overall variance than the complete hemisphere series. Figure 5.6e shows the estimated linear regression trends for the period 1880–2006 CE for the three target groups. The corresponding standard deviations over the same period are also shown in Figure 5.6f. Both representations display larger warming trends and variances when the hemispheric region considered excludes the ocean. Nevertheless, apart from the different level of variability, there in not found a systematic period of temperature differences such as the one evidenced by the reconstructions during the MCA.

Additionally, the LM simulations explored in Chapter 4 have been averaged also over the three spatial domains according to the three different regions. The respective STDs over the period 1001-1990 CE are represented in Figure 5.7 for NH-C (red), NH-LO (light green) and NH-LOE (dark green), similarly to Figure 4.5. For each simulation STD values are invariably smaller for the NH-C series compared to the land-only cases. The difference between NH-LO and NH-LOE is not as strong as with the NH-C but the inclusion of the tropics induces a relative reduction of variance with respect to the NH-LOE case. In any case, the relative variations between the spatial domains shall only be considered in the context of each simulation as there are important differences across models and experimental configurations that prevent direct comparisons among different simulations.

Therefore, the sources of temperature information tested within this section have not reproduced the differences observed among the reconstructions during the MCA, supporting the rationale of an intrinsic characteristic of the reconstructions as a plausible cause of such behaviour. Independently of the origin of the differences observed among the reconstructions, they have a direct impact

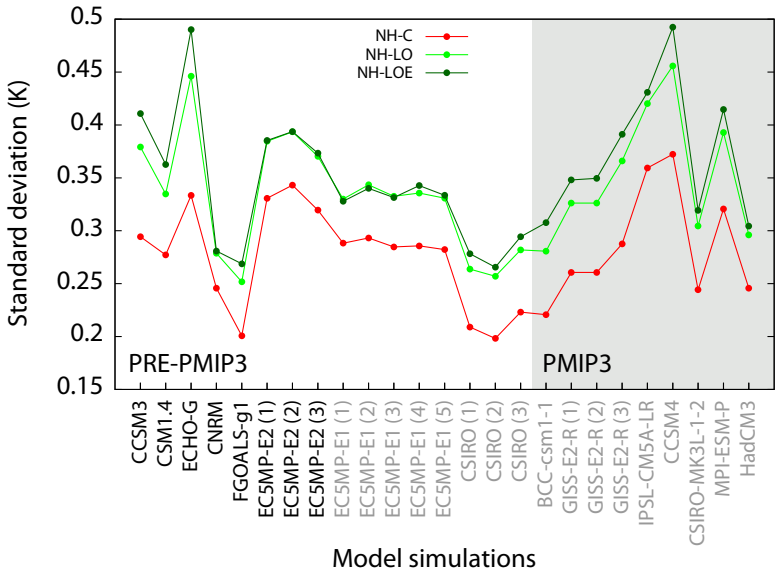

Fig. 5.7: Standard deviation (STD) of each millennial-length AOGCM simulation from Table 2.1 for the three spatial target groups. The STD is calculated over the common period 1001-1990 CE. The nonPMIP3 simulations are plotted over a white background, while the gray one corresponds to the PMIP3 runs. The names of those simulations that are driven by STSI are written in black while those using ssTSI are in gray.

on their intercomparison, as evidenced above. Indeed, it would be desirable that future intercomparison exercises consider the various spatial target subgroups as each one evidences a different level of variability. The latter is also applicable, based on previous results, to the comparisons between simulations and reconstructions (Chapter 6) and to analyse the influence of the external forcing in the different reconstructions, as it will be discussed in the next section.

5.2 Response to changes in to external forcing

The ensemble of temperature reconstructions from Table 2.3 (Figure 5.1) can be considered as the most representative evidence available of the real climate system. Thus, the series in Figure 5.1 allow us to explore the response of climate

to external forcing in a similar fashion as it was done for model simulations in Sections 4.2 and 4.3. Specifically, we focus in this section on analysing the linear relationship between global and hemispheric reconstructions and forcing configurations.

Figure 5.8 shows correlations that quantify the linear relationship between reconstructions and external forcing. According to results in previous section, reconstructions are classified attending to their spatial target (Table 2.3). Since it is not known which one of the available forcing specifications better represents the real past forcing, reconstructions are cross-compared with all TEF configurations described in Chapter 3 (Figure 3.4). Due to the large amount of TEF configurations available, for a better visualization of the results, the correlation values are divided, following previous analysis, into two panels according to the level of the solar variability included in the TEF configuration (STSI and ssTSI). Thus, for each reconstruction a set of correlation values is obtained, each value corresponding to a different TEF configuration. Correlation values lower than the minimum correlation value obtained between simulations and TEF in Figure 4.6 are highlighted by a gray background.

All the correlation indexes for each reconstruction group into a continuous range of values, and although there is no a predilect TEF configuration that dissociates its correlation value from the rest of evidences, the TEF configurations not including anthropogenic aerosol forcings tend to show higher correlation values than others. Nevertheless there is no a significant and systematic bias observed at considering STSI or ssTSI TEF configurations.

A bias to higher correlation values is observed for the continental series (NH-LO, NH-LOE, SH-LO and GL-LO), while lower values are systematically found at considering the reconstructions targeting also the oceanic areas (NH-C, SH-C and GL-C) in Figure 5.8. This effect is specially pronounced in the SH (Figures 5.8c,d), presenting larger differences between the correlation values associated to the complete and to the continental reconstructions than in the cases of NH (Figures 5.8a,b) and GL (Figures 5.8e,f) series. This behaviour could be related to a nonlinear response of the oceans to the external forcings, which indeed are characterized by a delayed response to any external influence or to methodological issues in the reconstructions including oceanic information. Nevertheless it should also be noticed that within the continental targets there is a larger amount of shorter lenght series than in the complete regions, series that do not span the complete LM, specially at considering the SH and GL scales (e.g. Huang et al., 2000; Leclercq and Oerlemans, 2012, see Table 2.3). When the correlation values are calculated for the last 500 years instead of the whole common period between TEF and reconstructions, as in Figure 5.8, differences within the correlation values obtained for the continental and complete reconstructions are reduced,

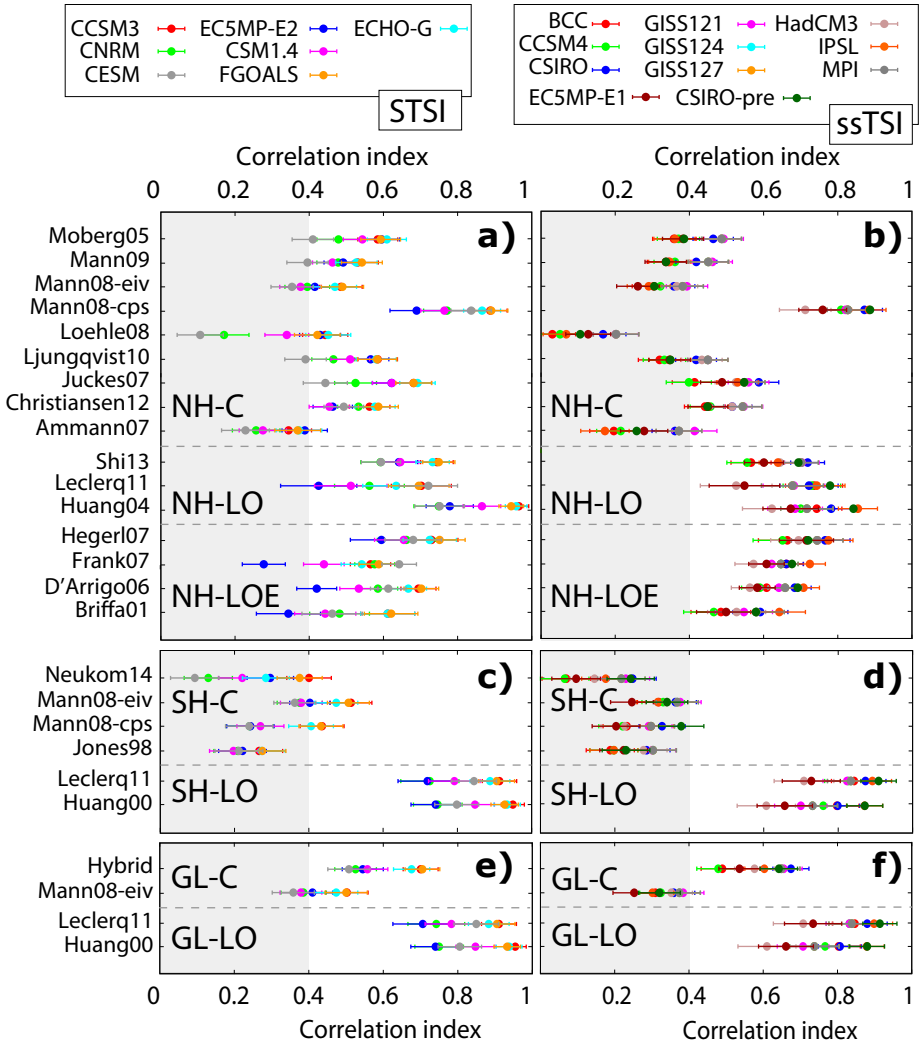

Fig. 5.8: Correlation values between temperature reconstructions and the various global TEF configurations. Reconstructions are distributed according to the NH (a, b), SH (c, d) or GL (e, f) targets and compared to forcing configurations involving stronger (STSI; left) and weaker solar forcing variability (ssTSI; right). Gray background indicates lower correlation values than those obtained between simulations and TEF in Figure 4.6.

i.e. the correlation values for the continental reconstructions increase. Thus, the lower correlations values obtained within the SH-C reconstructions and TEF configurations seem to be more associated with the discrepancies found beteween both variables during the first half of the LM, period already evidenced to be characterized by a large uncertainty in the reconstructions (see previous sections).

Nevertheless, although some of the correlations values of reconstruction-forcing pairs are lower than the minimum values obtained between simulations and forcing (grey shaded area), most correlations are above this threshold, with maximum values reaching 0.98. Indeed in almost all the reconstructions they are still high enough to consider an important part of the variance in reconstructions, the square of correlation between temperature and forcing, is a linear response to external forcing over decadal timescales.

Attending to this linearity between reconstructions and TEF and similarly to the process followed with the simulations in Chapter 4, an estimation of temperature sensitivity to forcing can be calculated in terms of the LMTCR (see Section 4.3) using the temperature reconstructions. Figure 5.9 shows the LMTCR values obtained for each reconstruction from the different TEF configurations. These values correspond, as in the case of the simulations, to the regression coefficients from the linear regression of the reconstruction over each TEF configuration. Following Figure 5.8, in Figure 5.9 values obtained for each series are divided into two panels according to the level of TSI variability (STSI, ssTSI) included in the TEF.

All the reconstructions show a positive value of LMTCR, in consistency with the positive correlations in Figure 5.8. The range of the plausible LMTCR values for each reconstruction, based on the various values obtained for different TEF configurations, depends the series analysed. It is interesting to note that the narrowest ranges of LMTCR values correspond to the reconstructions evidencing the lowest LMTCR values. These series (Ammann et al., 2007; Neukom et al., 2014; Jones et al., 1998) are, in turn, the ones showing the lowest correlation values with TEF (Figure 5.8). Since LMTCR is based on regression estimates, the resulting values will depend both on the correlation between temperatures and TEF and on the ratio of temperature and TEF standard deviations (Fernández-Donado et al., 2013). This implicitly establishes the requirement that reconstructions and TEF show covariability in time, a feature that is supported by many of the temperature-TEF pairs in Figure 5.8. However, if the correlation between both variables is low, the linearity is not secured and this will bias the regression between reconstruction and forcing to low values, like the estimates obtained for SH-C series in Figure 5.9. Alternatively, if the ratio of STD of temperature and forcing is too high, as in the Christiansen and Ljungqvist (2012b) reconstruction (Moberg, 2013), a bias to large LMTCR values is noticed.

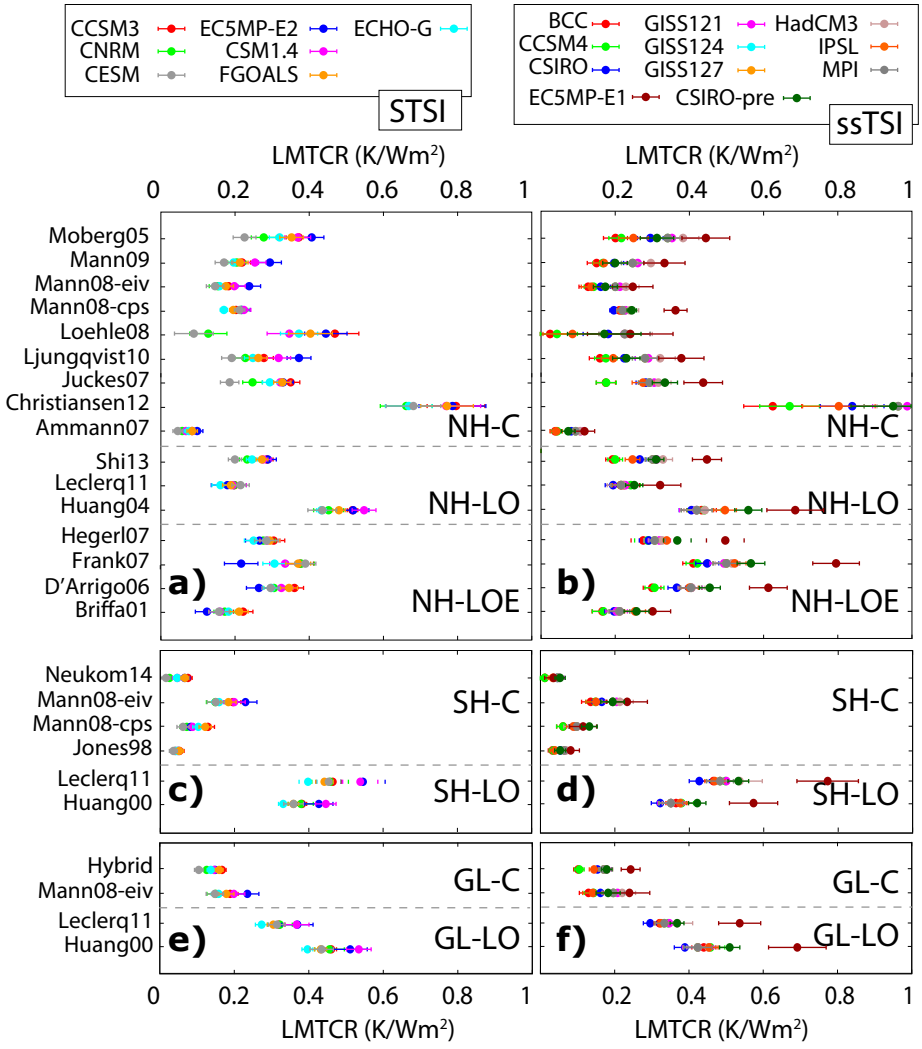

Fig. 5.9: LMTCR values for the reconstructions based on the linear regression between reconstructions and the various TEF configurations. Gray shadings indicate the 10-90 percentiles of the LMTCR ranges obtained for the simulations. LMTCR values are classified according to the target of the reconstruction, NH (a, b), SH (c, d) or GL (e, f). LMTCR values and shadings are divided into STSI (left) and ssTSI (right) groups corresponding to the different level of solar variability.

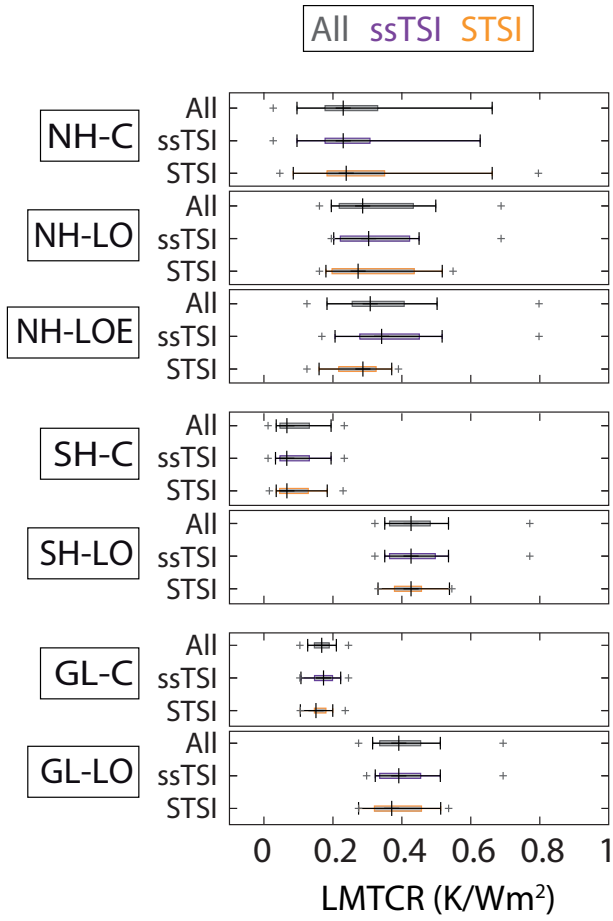

Fig. 5.10: Boxplots of the LMTCR values obtained for the reconstructions from Figure 5.9 for the STSI, ssTSI and all TEF configurations together. LMTCR values are grouped according to the spatial target of the reconstructions. Maximum and minimum values are indicated with a cross. 10, 50 and 90 percentiles are indicated by the whisker bars and the 25 and 75 percentiles by the size of the box.

There is a tendency to larger LMTCR values when using the EC5MP-E1 TEF configuration, although these values should be analysed cautiously since they come from sistematically slightly lower correlation values than other TEF configurations. Overall, there is no significant difference in the LMTCR values when considering the STSI or ssTSI TEF configurations, i.e. LMTCR values obtained for the suite of available TEF configurations are quite similar. The latter is supported by Figure 5.10 in which the LMTCR values in Figure 5.9 are synthesized for the various domains and solar forcing configurations. The distribution of the LMTCR values results independent from the level of solar variability forcing included in the TEF. Figure 5.10 shows that reconstructions targetting the NH evidence similar LMTCR values with a slight bias to larger LMTCR values in the continental reconstructions. Overall, the LMTCR values obtained for all the NH regions are around 0.2-0.3 K/Wm^2. This bias to larger LMTCR values in continental reconstructions is much more pronounced in the SH and GL and in agreement with the results of Section 5.1.2. Indeed, SH-LO and GL-LO show LMTCR values around 0.4 K/Wm^2 while their corresponding complete versions show values around 0.1 and 0.2 K/Wm^2 respectively. Since SH-C and GL-C areas include a large proportion of oceanic regions, the ratio of response to forcing is much lower than when only continental regions are considered. Continental reconstructions from SH and GL, in turn, evidence larger responses to the forcing than the NH reconstructions.

5.3 Conclusions

This chapter has assessed the current state of knowledge of the evolution of temperature during the CE based on the available reconstructions for NH, SH and GL scales, thereby updating Figure 1.2c and Masson-Delmotte et al. (2013). Conclusions extracted for the SH and GL scales should be, nevertheless, analysed cautiously due to the low amount of available reconstructed evidences over these regions.

The available NH reconstructions consistently evidence a warm period during the first half of the LM, followed by a cooler interval during the so-called LIA and followed by the recent industrial warming. They also agree on depicting minima of solar variability as well as the major volcanic eruptions (Chapter 3) as it was observed for the simulations in Chapter 4. The reconstructions spread is larger during the period between the 9th and the 15th centuries.

The conventional approach for defining the reconstruction ensemble in the NH is discussed, analysing and assessing the influence of different methodological decisions using various target domains and calibration periods. The analysis

demonstrates that changing methods and calibration periods contribute to arbitrary inflations or deflations of variance for each reconstruction in the ensemble. In turn, when considering various target domains together, different ranges of NH temperature variability are mixed, with NH-LO (NH-C) targets favoring larger (lower) variability. Independently of the reconstruction method, using different spatial domains as calibration targets has an impact on the reconstructed variability. The segregation of the reconstructions according to their spatial target in the representation of CE uncertainty demonstrates not only different ranges of variability, but also large differences between the average temperature of the NH-C and NH-LO subensembles during the end of the first millennium and the first half of the LM. Such differences among the various spatial targets observed within the reconstructions, nor any other indication of such systematic biases, either in simulations or in instrumental data, is reported. Therefore, it is unlikely that these differences can be attributed to mechanisms of internal or externally forced variability producing a different mean climate for the NH-C and the NH-LO groups. Additionally, this different behavior among the reconstruction ensembles remains after recalibrating the series with a common method, using a common instrumental target or a common calibration period. Therefore, other factors must contribute to the NH-C and NH-LO differences during the first centuries of the LM. Within this context proxy quality, abundance and sampling are factors that should not be disregarded as candidates accounting for such differences.

The relationship existing between the temperature reconstructions and TEF configurations it has been also assessed in an attempt to characterize the temperature response of the real system to the external forcing. Most of the available temperature reconstructions show high correlation values with the existing TEF configurations. The latter implies that a linear relationship between both variables can be assumed at multidecadal scales during the LM for most of the reconstructions. This linear relationship allows to quantify the temperature response to the forcing in the reconstructions through the LMTCR. All the reconstructions show positive values of LMTCR, i.e. a response to the forcing is evidenced in all the series representing the real climate system. There are not observed significant differences between the various existing TEF configurations, not being able to distinguish a predilect TEF configuration for the real climate evidences. NH reconstructions tend to show an homogeneous level of response to forcing, around $0.3\ K/Wm^2$, with a slight bias to larger LMTCR values in the continental reconstructions. SH-C and GL-C reconstructions agree on evidencing lower ratios of linear response to forcing due to the delayed response of the oceanic areas. Continental SH and GL reconstructions show LMTCR values around $0.4\ K/Wm^2$. These values of response to forcing shown by the reconstructions seem to be consistent with those evidenced by simulations in Chapter 4. An in-depth analysis

of the model-data consistency at global and hemispheric scales in, nevertheless, presented in the next chapter.

6

A comparison of simulations and reconstructions at global and hemispheric scale*

Chapters 4 and 5 have assessed the temperature responses of AOGCM/ESM LM simulations and reconstructions to changes in external forcing. Both reconstructions and simulations evidence a linear response to TEF above multidecadal timescales that is quantified through the LMTCR value. This chapter further discusses previous results and explores additionally the model-reconstruction consistency over the CE. Similarities found between both sources of information will be suggestive of a common response to external forcing, while the temperature variability observed in reconstructions and not reproduced by the simulations will be an indication of internal variability processes, problems in reconstructed variability or in modeling it.

* The main contents of this chapter are included in:
- Fernández-Donado, L., J. F. González-Rouco, C. C. Raible, C. M. Ammann, D. Barriopedro, E. García-Bustamante, J. H. Jungclaus, S. J. Lorenz, J. Luterbacher, S. J. Phipps, J. Servonnat, D. Swingedouw, S. F. B. Tett, S. Wagner, P. Yiou and E. Zorita, 2013: Large-scale temperature response to external forcing in simulations and reconstructions of the last millennium. *Climate of the Past*, **9**, 393-421. DOI 10.5194/cp-9-393-2013.
- Fernández-Donado, L., J. F. González-Rouco, E. García-Bustamante, J. E. Smerdon, S. J. Phipps, J. Luterbacher and C. C. Raible, 2015: Northern Hemisphere temperature reconstructions of the Common Era: Ensemble uncertainties and their influence on model-data comparisons. *Geophysical Research Letters*. **Under review**.
- González-Rouco,J. F., L. Fernández-Donado, C. C. Raible, , D. Barriopedro, J. Luterbacher, J. H. Jungclaus, D. Swingedouw, J. Servonnat, E. Zorita, S. Wagner and C. M. Ammann, 2011:Medieval Climate Anomaly to Little Ice Age transition as simulated by current climate models. *Medieval Climate Anomaly. Pages News.*, **19**, 7-8.

Model-reconstruction comparison is developed through this chapter by implementing various approaches. Section 6.1 compares the temporal evolution of both reconstructions and simulations integrating the results of Sections 4.1 and 5.1, thereby updating Figure 1.2d. We also incorporate the results of Section 5.1.1 to assess the influence of the various spatial targets included in the NH reconstructions spread (Chapter 5) on model-data comparison exercises. Section 6.2 compares the LMTCR ranges described by both sources of information (Section 4.3 and 5.2), offering a quantitative assessment of the linear response to forcing by both simulations and reconstructions. Section 6.3 focuses on analysing the temperature response to external forcing at interannual and at decadal timescales through the superposed epoch analysis approach. The latter is based on generating composites of reconstructed and simulated temperature sequences corresponding to the timing of a given forcing during periods of volcanic events or low solar variability. Section 6.4 analyses the changes during key periods of the LM like the MCA and the LIA. Finally, Section 6.5 summarizes the main conclusions.

6.1 Simulated and reconstructed temperature changes during the Common Era

Figure 6.1 shows the hemispheric and global temporal evolution of the simulated and reconstructed evidences since 800 CE. Simulated temperature anomalies with respect to the period 1500-1850 CE are plotted over the spread of reconstructions. Note that the reconstructions spread changes from Figure 5.2 to Figure 6.1 due to the choice of a different reference period. Specifically the grey shading for the NH (Figure 6.1a) is somewhat narrower during 1500-1850 CE due to having this time interval as reference, and slightly wider over the rest of the millennium relative to what is shown in Figure 5.2. Hence, care must be taken not to interpret the spread in Figure 6.1 directly as a measure of the reconstruction ensemble uncertainty of past temperature changes. The shape of the spread can be used for the purpose of verifying consistency between simulations and reconstructions.

Simulated temperatures lie mostly within the spread of the available reconstructions (Masson-Delmotte et al., 2013) for all domains, albeit some differences are noticeable. For the GL and SH domains (Figure 6.1b,c), trends after 1500 CE seem consistent with the reconstructions. Before that time, the reconstructions suggest a larger range of variability with higher temperatures until the 11th century and during the 13th and 14th centuries. Nevertheless, the amount of proxy evidence for the SH is limited as well as the number of reconstructions both for the GL and SH domains. Therefore, comparisons should be done carefully.

The evolution of simulated NH temperatures is embedded within the area of uncertainty defined by the reconstruction spread (Figure 6.1a), although some

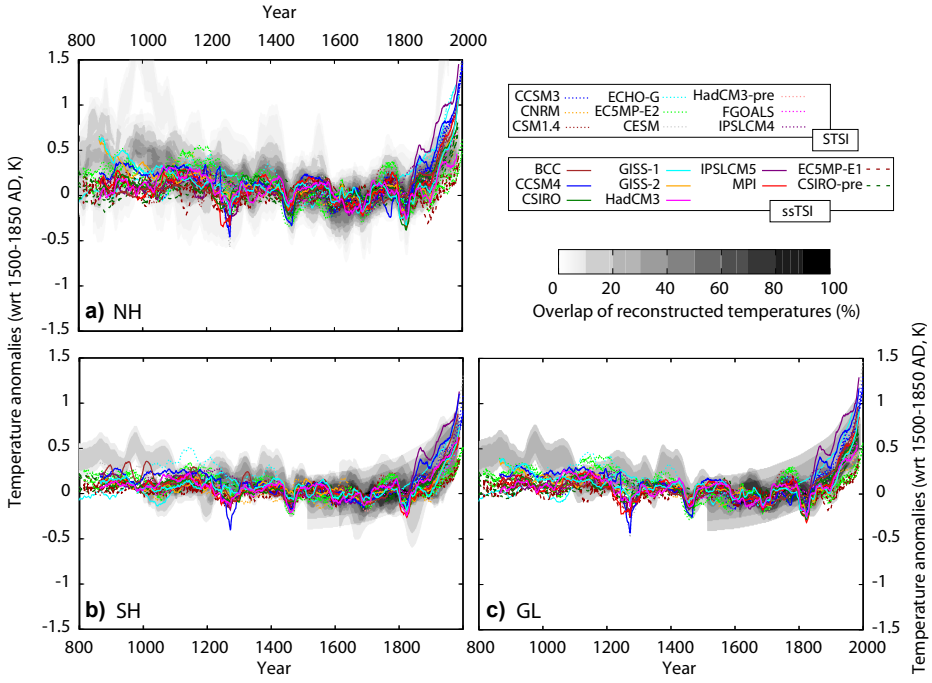

Fig. 6.1: NH (a), SH (b) and GL (c) simulated temperature anomalies with respect to 1500-1850 AD (colors) over the uncertainty envelope of reconstructions (grey shading) shown in Figure 5.2. Note the change of the shape of the grey shaded area between this figure and Figure 5.2 due to the selection of the different reference period. The anomalies are calculated wrt 1500-1850 AD. All series are 31-yr moving average filter outputs.

differences may be noted and briefly discussed herein. Firstly, the areas of larger density in the reconstruction spread tend to show more low frequency variability than the simulations. As for the case of the SH and GL domains, simulated temperatures up to the 11th century and during the 13th and 14th centuries are below the area of higher density of probability of reconstructed temperatures. Thus, reconstructions indicate a warmer MCA relative to the reference period than the models do. These areas of larger density of spread are mostly a contribution of the reconstructions showing larger amplitude changes at multi-decadal and centennial timescales (Figure 5.2). The STSI ensemble (dashed lines) seems to show

a slightly larger range of low frequency changes relative to the ssTSI ensemble, particularly in the MCA transition to the LIA (see also Figure 1.2d; Masson-Delmotte et al., 2013; Fernández-Donado et al., 2013). It is difficult to ascertain whether this could indicate a problem on the side of the models or on the side of the reconstructions. If the latter were to be taken as a reliable estimate of the amplitude of past low frequency variability, this would suggest that either models underestimate the real world sensitivity or that forcing changes in pre-industrial times are underestimated, particularly in the ssTSI group (Figure 3.4f). Alternatively, low frequency multi-decadal and centennial changes produced by internal variability at hemispherical scales could optionally also contribute to the reconstructed spread. However, such low frequency variability is not simulated in any of the available control runs (not shown) and would mean that AOGCMs/ESMs underestimate internal variability at these timescales.

Secondly, the simulations show a larger spread than the reconstructions during the industrial period . The experiments from the CCSM3, the CCSM4, the ECHO-G, the IPSLCM5 and the EC5MP ensembles evidence the largest discrepancies. CCSM3, IPSLCM5 and ECHO-G clearly suffer from not including aerosols and land use and overestimate the warming in comparison to the reconstructions. CCSM4 also shows this overestimation of the temperatures, but the reasons for this increase are different since it includes the anthropogenic aerosols forcing (Table 2.1). This increase could be referred to the climate sensitivity of the model, and despite the TCR value associated to CCSM4 (Table 2.1) is not really high, its LMTCR value corresponds to one of the largest (see Figure 4.10). In turn, the EC5MP temperature increase is lower than in the reconstructions. The reasons for this could be also related to the lowest values of LMTCR evidenced in Figure 4.10 despite the highest transient climate responses in future scenario simulations (see Table 2.2). Arguably, the physics related to the treatment of aerosols or land use changes may exacerbate the related cooling in this model during the 20th century.

Thirdly, the largest differences between the reconstructions and the simulations take place in the 800 - 1100 CE period, during which the simulations are well below the area of maximum probability of the reconstructions as for the SH and GL domains. The weak model response during this period is due to the relatively low TEF values (Figure 3.4) during the 10th and 11th centuries. Thus the discrepancy can be essentially established on the ground of differences between reconstructions and forcings. Although forcing factors over the last millennium are not perfectly constrained (e.g. Plummer et al. 2012; Schmidt et al. 2011, 2012), the quality of their reconstructions (Chapter 3) can be considered homogeneous through time, and thus it can be argued that this discrepancy points to a problem in the reconstructions during this period (Fernández-Donado et al.,

2013). An alternative argument would be that multicentennial climate variability during this time would be detached from external forcing and dominated by internal processess. If we consider the ensemble of reconstructions, there are multiple issues that can influence the ensemble spread, as discussed in Section 5.1.1, that can influence the comparison of reconstructions and model simulations and should be considered at this stage.

Indeed, the sensitivity of the ensemble uncertainty to methodological decisions described in Chapter 5 can also influence model-data comparisons in Figure 6.1. Large differences in NH reconstructions like those highlighted in Figure 5.3 for different target domains in reconstructions of the NH can be considered in comparisons like that of Figure 6.1. While some previous model-data comparison assessments have taken the segregation of the spatial targets into account (Schurer et al., 2013), this has often not been the case (e.g. Masson-Delmotte et al., 2013; Fernández-Donado et al., 2013; Jungclaus et al., 2010; Servonnat et al., 2010; Landrum et al., 2013).

In this context, two different possibilities can be used to provide more adequate comparisons: 1) assess model and data after recalibrating reconstructions to a single target as in Figure 5.5 or 2) focus on simulated hemispheric averages that sample the same domain as the reconstructed targets. The first option has the disadvantage that the spread of the final ensemble is highly sensitive to the recalibration method. The second option implies a reduced number of estimates for each NH target while preserving the original character of the reconstructions. This latter option is the one that is applied in Figure 6.2 where climate simulations are spatially sampled and averaged over each target area. These simulations are shown over the corresponding 10–90% uncertainty envelopes built by considering only the reconstructions that originally target the same region (Table 2.3). A version of Figure 6.1a including all the reconstructions together is provided in Figure 6.2d for comparison.

Figure 6.2a-c indicates that simulations and reconstructions show more noticeable differences in the temperature amplitude across the CE when segregated in different regional targets, thus questioning the overall agreement in previous literature (Jansen et al., 2007; Fernández-Donado et al., 2013; Masson-Delmotte et al., 2013) based on comparisons like that of Figures 6.1a and 6.2d. The uncertainty bands representing the NH-C, NH-LO and NH-LOE temperature evolution no longer embrace the ensemble of simulations throughout the whole period as in Figure 6.2d. This is especially the case during the end of the first millennium and the first half of the LM for the NH-C group when simulations underestimate the reconstructions. Also, NH-LO and NH-LOE (Figure 6.2b,c) show noticeable differences, with reconstructions suggesting higher (lower) temperatures than models in the 11th (12th-13th) centuries. These results thus indicate that

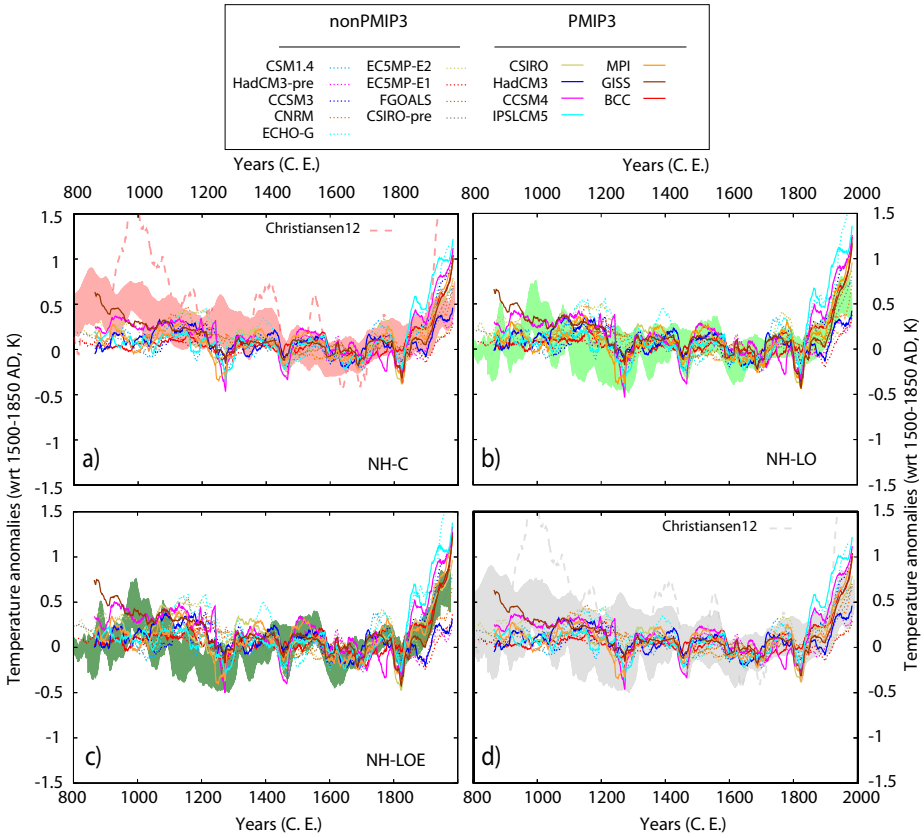

Fig. 6.2: Simulated temperature anomalies (solid/dashed lines) and reconstructed ensemble uncertainty (shading) wrt 1500 - 1850 CE for the NH-C (a), NH-LO (b) and NH-LOE (c) target groups. In (d) no spatial domain segregation is considered for simulations and the gray shaded area includes all reconstructions in Table 2.3 as in Figure 6.1. The reconstruction of Christiansen and Ljungqvist (2012b) is shown with dashed lines in panels (a) and (d). All series are 31-yr moving average filter outputs.

focusing the comparison on different spatial targets discloses prominent discrepancies between simulations and NH-C reconstructions during the period from 800 to 1400 CE. Temperatures in the NH-C (NH-LO) reconstructions peak around

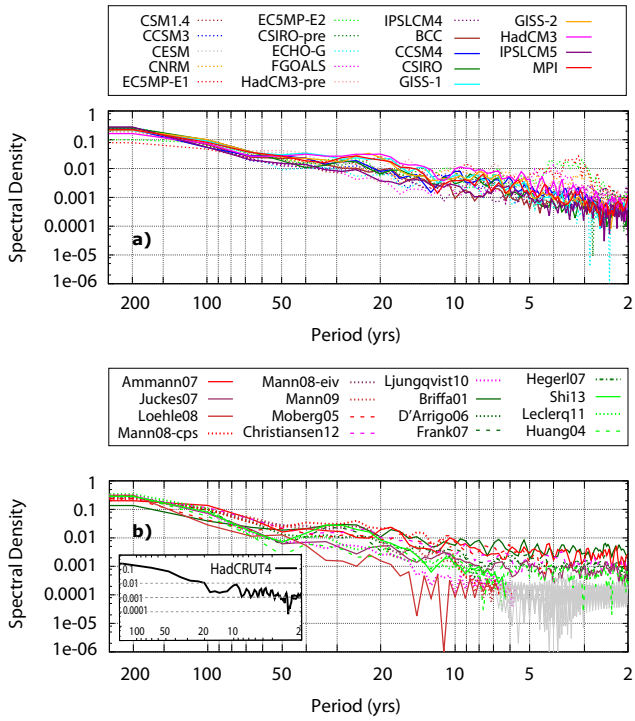

Fig. 6.3: Normalized spectra of simulated (a) and reconstructed (b) NH temperatures. For the models for which several simulations (CCSM3,ECHO-G,CSIRO) were available, averages of the runs including volcanic forcing were considered (Figure 6.1). Gray lines depict Gibbs oscillations in reconstructions missing high frequency variability. The inset in b) shows the normalized spectra of the NH instrumental data (Brohan et al., 2006).

ca. 900 CE (1000 CE) while in the simulations the maximum is found around ca. 1100-1200 CE (Masson-Delmotte et al., 2013) independently of the spatial domain considered.

It is interesting to extend the comparison of temporal variability of the simulations and reconstructions to the spectral domain. Figure 6.3 shows normalized spectra for NH temperatures in both the simulations and reconstructions. The reconstructions that do not provide variability at high frequencies (2-10 yr timescales) suffer from spectral noise (spectra with dashed lines; Gibbs os-

cillations depicted with grey colour). The spectra of the other reconstructions compare well with simulations at all frequency ranges. For the high frequencies (interannual timescales) the reconstructions tend to show a somewhat stable level of variance density whereas most of the simulations (except for EC5MP, CNRM and FGOALS) show a continuous decay. Interestingly, the spectra of TEF show a similar decay for all models except for the CSIRO-pre (Figure 3.6) at high frequencies. This suggests either a possible noise contamination at high frequencies on the proxy side or an underestimation of interannual variability by the models. The inset in Figure 6.3b shows the normalized spectra of the NH instrumental data (Brohan et al., 2006), evidencing a behaviour that is more similar to that of the proxies and thereby suggesting an underestimation of interannual variability by the models. Additionally, some of the simulations (CNRM, FGOALS, EC5MP) show an anomalously high accumulation of variability at 3-5 yr timescales which is not visible in the reconstructions. This is produced in these models by enhanced variability in the Tropics at these timescales (not shown, Zhang et al., 2010) and is also not supported by instrumental data.

A small decay of the spectral density between 10 and 20 yr timescales is apparent in the simulations, reconstructions and observations, as well as in TEF. At multidecadal and longer timescales, the spectra of the simulations and the reconstructions shows a similar shape, which also resembles that of TEF (Figure 3.6). Indeed at these timescales, the relation to forcing is more evident and larger differences are observed between the unforced and forced simulations (Flato et al., 2013). In Chapter 3 we argued that the relative increase of variance at 20-40 yr timescales was due to both solar and volcanic variability. The spectra of EC5MP-E1 ensemble shows the lowest proportions of low frequency variability, which lies below those of the reconstructions. The factors that may contribute to this are the lower solar forcing variability and the small warming trends in the 20th century simulated by this model related to the low LMTCR values. The latter factor seems to play a major role since all the other ssTSI simulations (CSIRO-pre and the complete set of PMIP3 experiments) show similar levels of low frequency variability than models of the STSI group and they also simulate larger trends in the 20th century.

6.2 Response to external forcing above multidecadal timescales

The response to the external forcing in simulations and reconstructions was assessed using the concept of the LMTCR in Chapters 4 and 5, respectively. LMTCR quantifies the linear response to the TEF at multidecadal timescales and above based on a simple linear regression between the temperature signal

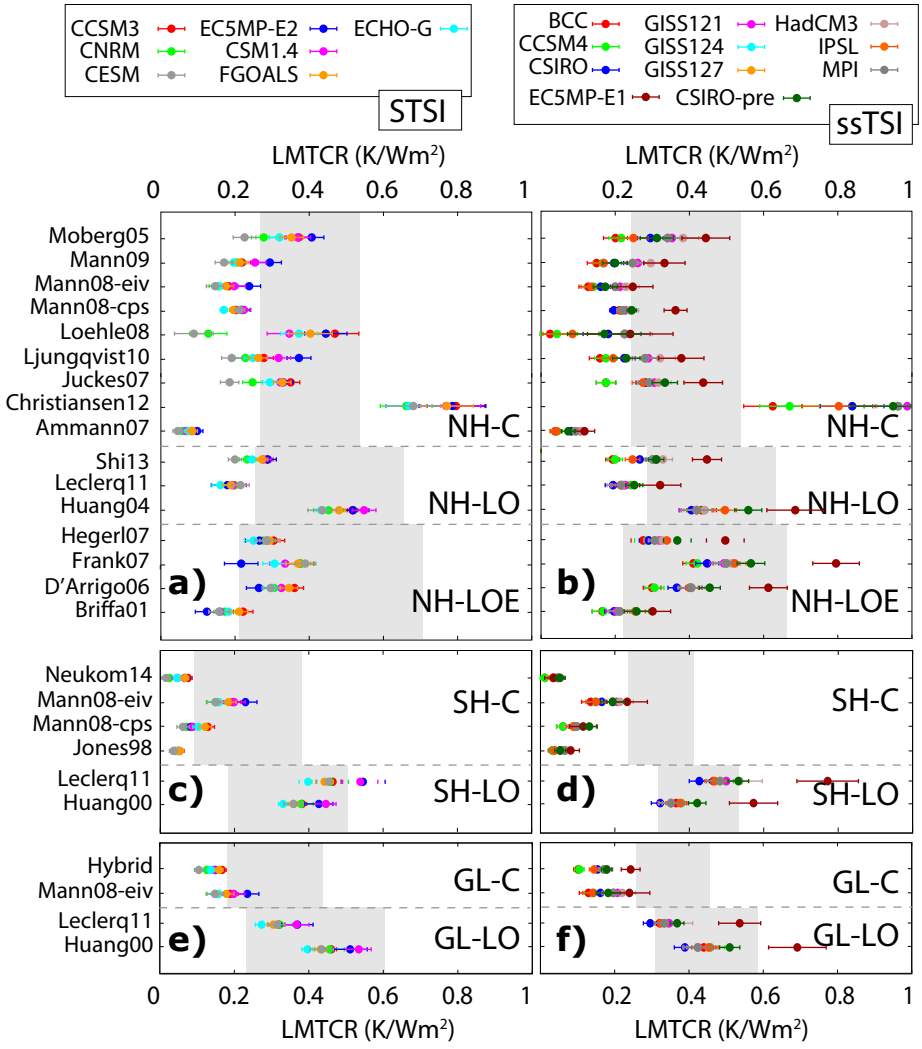

Fig. 6.4: LMTCR values for the reconstructions based on the linear regression between reconstructions and the various TEF configurations. Gray shadings indicate the 10-90 percentiles of the LMTCR ranges obtained for the simulations. LMTCR values are classified according to the target of the reconstruction within the NH (a, b), SH (c, d) and GL (e, f) scales (Table 2.3) and divided into STSI (left) and ssTSI (right) groups corresponding to the origin of the TEF configuration applied.

and TEF (Fernández-Donado et al., 2013). LMTCR allows for analysing the consistency of both sources of information from this new perspective.

Figure 6.4 shows the LMTCR values obtained from the reconstructions over the range of LMTCR values (shaded areas) indicating the 10-90 percentiles of LMTCR range obtained for the simulations (see Figure 4.12). LMTCR values are organized according to the level of solar variability included in the TEF configurations, STSI and ssTSI. Thus, the LMTCR values from the reconstructions obtained with the STSI (ssTSI) TEF configurations are plotted over the LMTCR range developed from the STSI (ssTSI) simulations. In addition, based on the influence of the different spatial targets of the reconstructions on model-data comparison exercises described in previous section for the NH, the LMTCR ranges from the simulations are also calculated at distinguishing in the runs the various spatial target regions included within the large-scale reconstructions (Table 2.3). Thus, the LMTCR ranges shown for the simulations in Figure 6.4 vary within the same panel according to the spatial target indicated. As in the reconstructions (Chapter 5), LMTCR values obtained in simulations for complete regions, i.e. including oceans, are sistematically a bit lower than the ones corresponding to continental areas.

In general, LMTCR values obtained for the reconstructions in Figure 6.4 are in the range of the ones obtained for the simulations, suggesting a consistent and positive linear response of the simulations and reconstructions to the external forcing. Nevertheless, the level of agreement between the LMTCR values coming from the two paleoclimatic tools, as well as the reliability of this result, depends on the spatial target considered and also on the different ranges of TSI variability.

Due to the fewer amount of reconstructions available for the SH and GL domains, the LMTCR values, and thus the model-data comparison exercises, are characterized by a lower level of reliability than the NH. Despite this lower reliability, SH (Figure 6.4c,d) and GL (Figure 6.4e,f) scales agree with NH (Figure 6.4a,b) in depicting for the land only targets, in general, a better agreement between the LMTCR values from reconstructions and simulations than for the complete regions, i.e. including also the oceans.

Although most of the LMTCR values derived from individual reconstructions lie within the simulated NH ranges (Figure 6.4a,b), there are some noticeable discrepancies. For instance, the simulated range is not consistent with some of the reconstructions, like Ammann et al. (2007), Mann et al. (2008), Leclercq and Oerlemans (2012) and Christiansen and Ljungqvist (2012b) for any, or almost any, of the available TEF configurations. The reasons for the disagreement of these series are, nevertheless, different. The Ammann et al. (2007) reconstruction was already characterized by a really low LMTCR value based on the low correlation value (Chapter 5). However, the rest of the reconstructions present

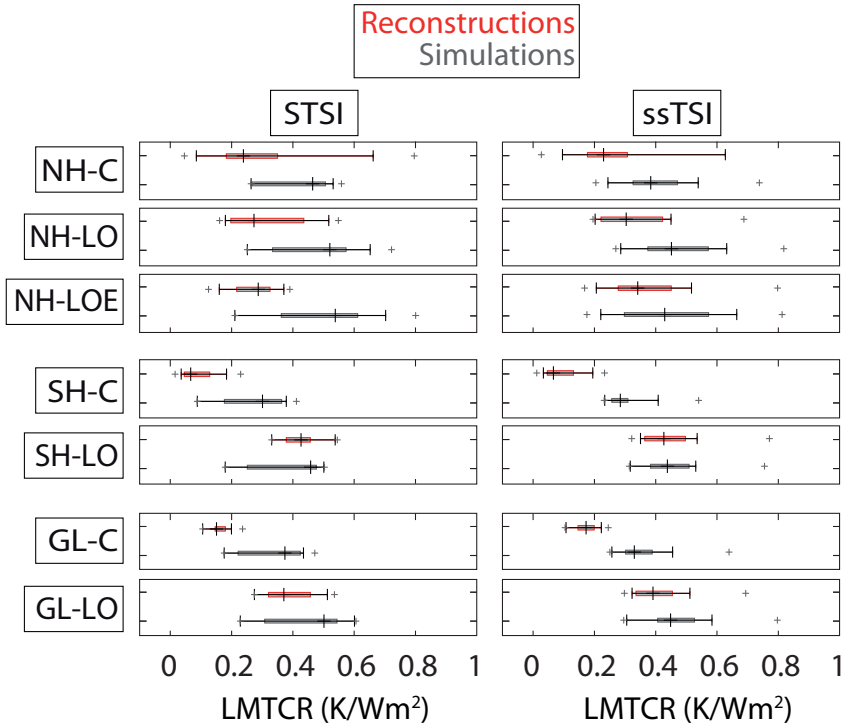

Fig. 6.5: Boxplots of the LMTCR values obtained for the ensemble of simulations (gray) and reconstructions (red) from each spatial target for the STSI and ssTSI TEF configurations. Maximum and minimum values are indicated with a cross. 10, 50 and 90 percentiles are indicated by the whisker bars and the 25 and 75 percentiles by the size of the box.

relatively high correlations (Figure 5.8), and these discrepancies can be traced to different ratios of temperature vs. TEF variability (standard deviation) relative to those in the simulations: lower in the case of the reconstructions of Mann et al. (2008) and Leclercq and Oerlemans (2012) and much higher in the case of Christiansen and Ljungqvist (2012b). Therefore, this analysis highlights discrepancies between reconstructed and simulated climate that report different rates of temperature response to forcings.

Figure 6.5 presents the corresponding boxplots of LMTCR values from Figure 6.4 in an attempt to explore whether the STSI or ssTSI TEF configuration may

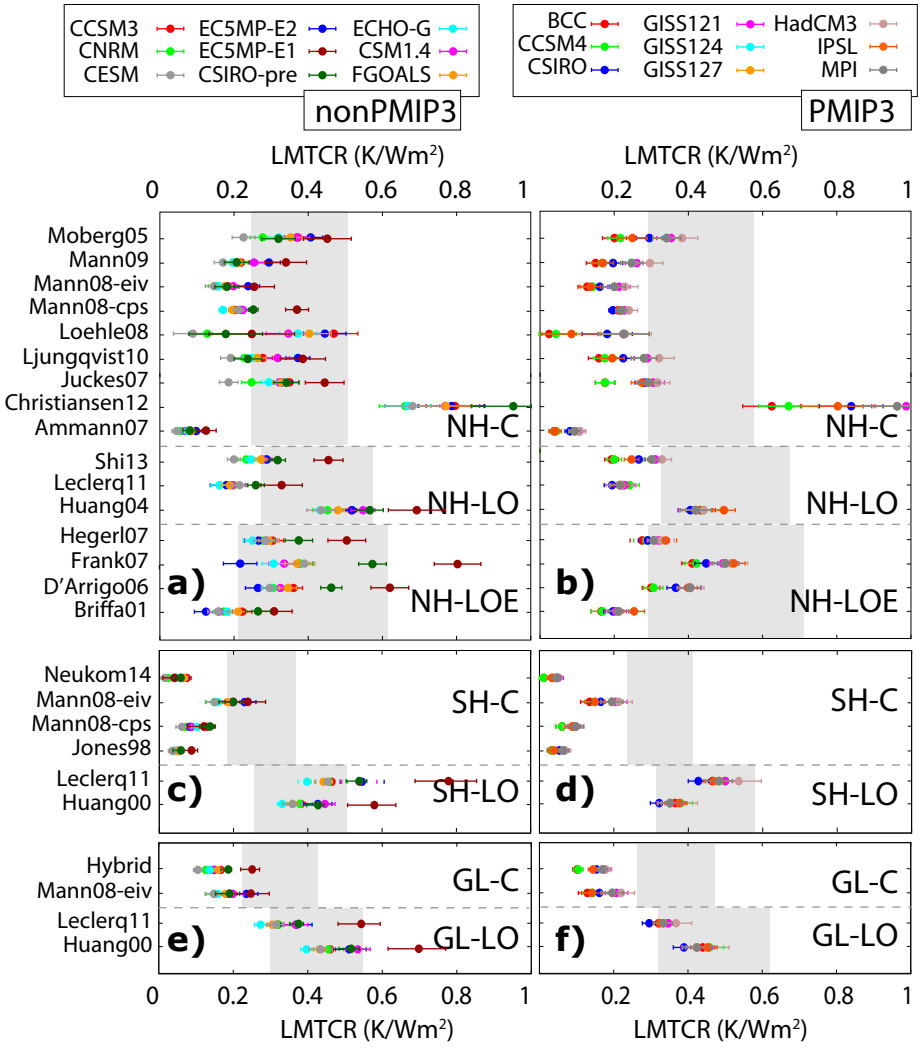

Fig. 6.6: As Figure 6.4, but distinguishing between nonPMIP3 (left) and PMIP3 (right) TEF configurations

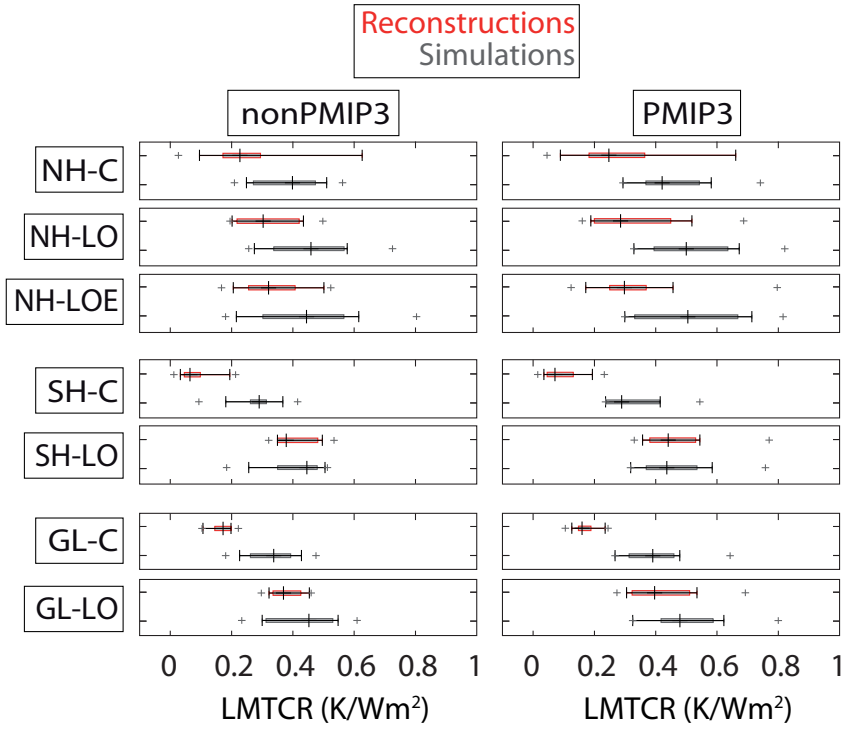

Fig. 6.7: As Figure 6.5, but distinguishing between nonPMIP3 (left) and PMIP3 (right) TEF configurations.

lead to larger model-data consistency. The ssTSI TEF configuration seems to evidence slightly more similar LMTCR mean values in reconstructions and simulations for most domains. Nevertheless, when considering the complete LMTCR ranges, there is not a significant difference in the level of model-data consistency illustrated by both TEF configurations.

Figure 6.6 is equivalent to Figure 6.4 but classifying the LMTCR values from the reconstructions according to PMIP3 and nonPMIP3 TEF configurations considered. The somme reconstructions that showed discrepancies with simulations in Figure Figure 6.4, also show it here, namely Ammann et al. (2007), Mann et al. (2008), Leclercq and Oerlemans (2012) and Christiansen and Ljungqvist (2012b). Figure 6.5 shows the boxplots corresponding to the LMTCR ranges obtained in Figure 6.6, similarly to Figure 6.5 but distinguishing between the

PMIP3 and nonPMIP3 TEF configurations. This Figure attempts to illustrate whether there exist a larger model-data consistency with the new generation of model experiments or not. As it occurred in Figure 6.5, there is not a predilect TEF configuration for a better consistency between models and reconstructions, evidencing indeed for both generations of experiments a very similar level of consistency.

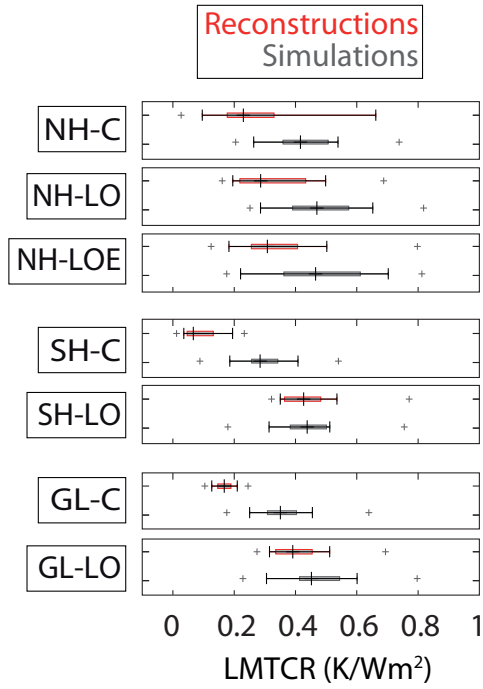

Fig. 6.8: As Figures 6.5 and 6.7, but considering the complete ensemble of TEF configurations together.

Based on the similarities found in previous classifications of the LMTCR ranges, Figure 6.8 shows the LMTCR ranges obtained for simulations and reconstructions when all the TEF configurations are considerd together. Figure 6.8 shows the larger consistency found within the continental targets than when also oceans are considered. Additionally, a sistematically larger response in the NH by the simulations than the reconstructions is also evidenced.

6.3 Response to external forcing at interannual and decadal timescales

This section offers additional insights into model-data consistency by analising the temperature response at interannual and decadal timescales. This is done by applying a superposed epoch analysis (Masson-Delmotte et al., 2013). The latter is based on generating composites of reconstructed and simulated temperature sequences corresponding to the timing of the strongest volcanic events and of the intervals characterized, at decadal timescales, by a weaker solar forcing (PAGES2k-PMIP3 group, 2015).

Figure 6.9 shows the results of superposed composites (time segments from selected periods positiones so that the years with peak negative forcing are alined; Masson-Delmotte et al., 2013) obtained from analysing the Gao et al. (2008) and the Crowley and Unterman (2012) volcanic forcings as well as their effect on NH temperature reconstructions and simulations. All the panels show, for each forcing series analysed (colours), a shading indicating the complete range of values attained by the various time intervals selected and the mean composite value through a dashed line. Upper panels show the composites of the volcanic forcing when considering: 12 of the strongest volcanic events after 1400 CE (Figure 6.9a); 5 clusters of eruptions from multi-decadal volcanic forcings (Figure 6.9b); 7 intervals characterized by a low solar irradiance at multidecadal timescales (Figure 6.9c). The latter also includes the composite of the corresponding solar forcing for the years of the minima. While the years for the strongest volcanic eruptions at annual or multidecadal timescales vary from Gao et al. (2008) to Crowley and Unterman (2012) reconstruction (see caption for details on the selection of the years), the years selected for the minima of solar forcing are common for both reconstructions and follow the selection of Masson-Delmotte et al. (2013). Middle (Figure 6.9d-f) and lower (Figure 6.9g-i) panels show the composites of the temperature responses to the corresponding selection of periods for the NH reconstructions and simulations, respectively. For analysing the NH temperature reconstructed responses all the NH series starting, at least, in 1400 CE, are considered (see Table 2.3) for both collections of time intervals, i.e. Gao et al. (2008) and Crowley and Unterman (2012). In turn, simulations considered for each collection of years are only those applying the corresponding Gao et al. (2008) or Crowley and Unterman (2012) volcanic forcing (see Table 3.1).

The mean peak values of the composites are approximately -4 W/m^2 and -3 W/m^2 for the Gao et al. (2008) and Crowley and Unterman (2012) volcanic forcing estimations, respectively (Figure 6.9a). Gao et al. (2008) evidences, as it was already mentioned in Chapter 3, the largest volcanic forcing estimations, reaching -12 W/m^2 for a single event (1258 CE), in comparison with the -6 W/m^2 reached by Crowley and Unterman (2012). The associated temperature responses

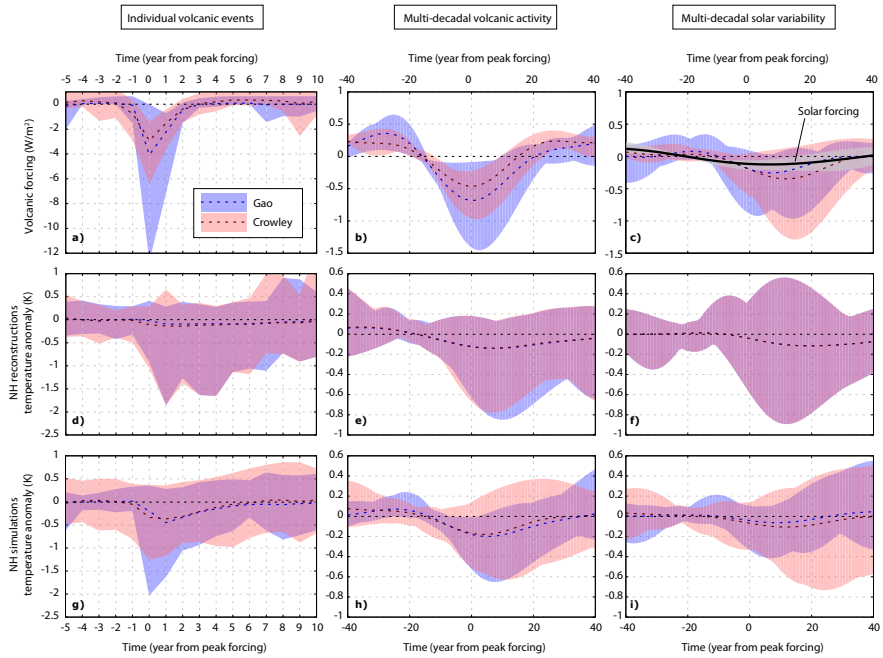

Fig. 6.9: Superposed composites of the volcanic forcing (a, b, c) and reconstructed (d, e, f) and simulated (g, h, i) temperature responses to: (left panels) 12 of the strongest volcanic events after 1400 CE; (middle panels) multi-decadal changes in volcanic activity; (right panels) multi-decadal changes associated to periods of low solar variability. Panel c) also includes the superposed composites of solar forcing at selecting the years of solar minima. Each shading shows the complete range of forcing or temperature values attained by the various time intervals selected and the composite mean by a dashed line. Gao et al. (2008) volcanic forcing variations and temperature responses associated are shown by light red shading and lines, while the ones corresponding to Crowley and Unterman (2012) are evidenced in blue, obtaining magenta in overlapping areas. A detailed description of the criteria followed for selecting the years and intervals in each reconstruction can be found in Masson-Delmotte et al. (2013). For the Gao et al. (2008) reconstruction there have been considered the years 1452, 1584, 1600, 1641, 1673, 1693, 1719, 1762, 1815, 1883, 1963 and 1990 for the individual volcanic composites(left); and the 1259, 1453, 1601, 1694 and 1816 for the decadal analysis (middle). Selection of years from Crowley and Unterman (2012) is coincident with the one presented in Masson-Delmotte et al. (2013): 1442, 1456, 1600, 1641, 1674, 1696, 1816, 1835, 1884, 1903, 1983 and 1992 fro the individual events (left) and 1259, 1456, 1599, 1695 and 1814 for the multidecadal exercises (middle). Intervals selected for the low solar minima follow Masson-Delmotte et al. (2013): 1044, 1177, 1451, 1539, 1673, 1801 and 1905.

are larger for the simulations (Figure 6.9g) than reconstructions (Figure 6.9d), without noticeable differences existing if the Gao et al. (2008) or Crowley and Unterman (2012) forcing estimation are used. The peak of the composite averages obtained for the simulations is around -0.5 K while for the reconstructions is below -0.25 K. The temperature perturbation typically lasts longer than the forcing itself, with a recovery to pre-eruption temperatures after 3 to 5 years in the simulations and almost to 8 years in the reconstructions (Masson-Delmotte et al., 2013; PAGES2k-PMIP3 group, 2015).

At multidecadal timescales, the composite mean volcanic forcing (Figure 6.9b) peaks around -0.75 W/m^2 for the Gao et al. (2008) reconstruction and around -0.5 W/m^2 for the Crowley and Unterman (2012) estimation. Interestingly, the temperature responses from reconstructions (Figure 6.9e) and simulations (Figure 6.9h) are comparable in magnitude, with a mean value close to -0.2 K. Nevertheless, simulations show a faster recovery than reconstructions, reaching the pre-eruption temperature values after 20 years, in contrast to the 40 years evidenced by the reconstructions.

The analysis of the solar variability at multidecadal timescales (Figure 6.9c) evidences weaker variations than volcanic forcing for the same periods selected. The composite temperature responses to these variations show a mean value of -0.1 K in both simulations (Figure 6.9i) and reconstructions (Figure 6.9f). Nevertheless, these temperature variations should be analyzed cautiously, since they may be influenced by the coincident variations in volcanic forcing shown in Figure 6.9c (Masson-Delmotte et al., 2013).

6.4 Temperature changes from the MCA to the LIA

Additional model-data comparison exercises are provided within the current section by analysing the temperature differences between key periods of the CE. Considering the overlap in time span of the reconstructions and simulations, these periods correspond to the warmest and coolest centennial or multi-centennial time intervals since 800 CE, i.e. the MCA (950 - 1250 CE), the LIA (1400 - 1700 CE) and 20th century (1900 - 2000 CE). The convention adopted for the definition of the MCA and the LIA periods is the same as in Mann et al. (2009).

Figure 6.10 presents the temperature differences evidenced by NH reconstructions and simulations for the MCA-LIA (Figure 6.10a) and present-LIA (Figure 6.10c) transitions. In addition to the temperature changes, the corresponding variations of TEF within the configurations applied to each model are also illustrated in Figure 6.10b and d. Thus, changes observed in temperature can be analysed together with the associated variation in forcings for the same transitions.

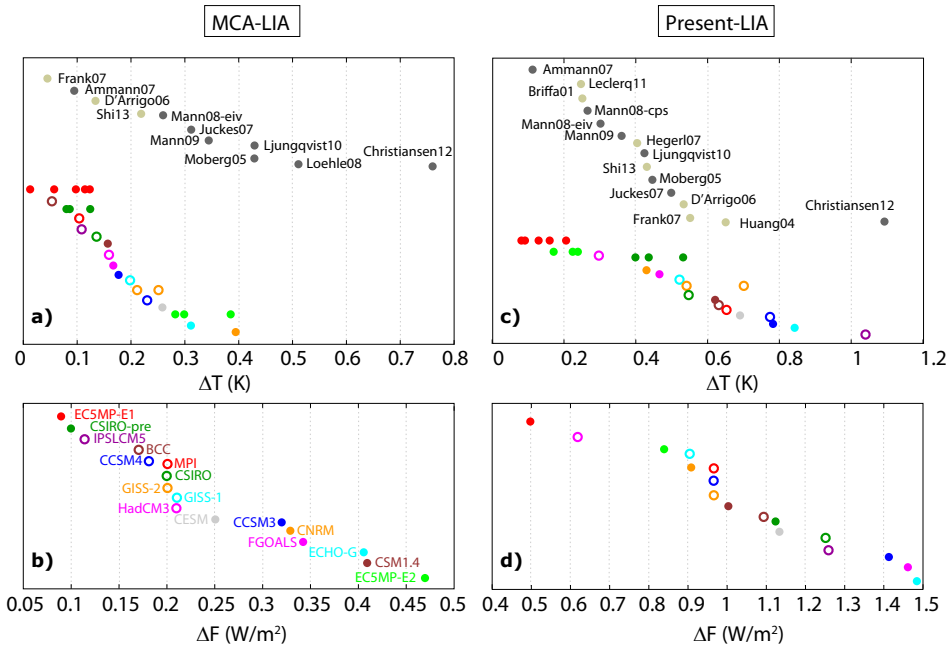

Fig. 6.10: a) Temperature change in the MCA (950 - 1250 CE) - LIA (1400 - 1700 CE) transition for NH simulations and reconstructions in Tables 2.1 and 2.3, respectively. Reconstructions are labeled, distinguishing between the NH-C series (grey dots) and the NH-LO and NH-LOE (dark yellow dots). Temperature changes from simulations are represented by coloured dots, distinguishing between the PMIP3 (solid dots) and nonPMIP3 (hollow dots), see lengend in panel b). b) MCA-LIA TEF change corresponding to the forcing applied to the simulations in a). c) and d) Equivalent to a) and b) but for the present (1900 - 2000 CE) - LIA transition.

The MCA-LIA transition (Figure 6.10a,b) is characterized by positive differences in forcing as well as in simulated and reconstructed temperature changes. Reconstructions evidence a larger range of temperature changes (0.05 - 0.8 K) than simulations (0 - 0.4 K). Changes in forcing (Figure 6.10b) are clearly separated into the ssTSI (0.1 - 0.2 W/m^2) and STSI (0.25 - 0.45 W/m^2) groups, with the lowest values of forcing change being attained by the two ssTSI models from the nonPMIP3 ensemble, i.e. CSIRO-pre and EC5MP-E1. The CESM sim-

ulation, although officially considered as an STSI experiment, due to its peculiar solar forcing estimation (Chapter 3) represents an intermediate state from ssTSI and STSI groups with a variation of 0.25 W/m^2, located closer to the ssTSI variations than the rest of the STSI ones. Simulated temperature changes tend to be organized also according to the level of solar variability with the lowest temperature changes evidenced by the ssTSI simulations (PMIP3, CSIRO-pre and EC5MP-E1) and the largest values corresponding to STSI experiments. The latter suggest that solar forcing could be a major player in the model simulations of the MCA-LIA transition. Nevertheless, there is not such a clear separation between the two groups of simulated temperature changes as it is evidenced in the forcing, highlighting the influence of other factors on the temperature response such as the internal variability and model climate sensitivity. For instance, the CSM1.4 simulation, despite presenting one of the largest forcing variations, evidences the lowest temperature change within the STSI ensemble, probably associated to its low climate sensitivity (Table 2.2)

Interestingly, members of a model ensemble sharing the same external forcings (CSIRO-pre, EC5MP, GISS) show a spread of temperature changes due to internal variability. This spread may be larger than differences among various models, for instance, intra-model variability in the EC5MP and the CSIRO-pre ensembles is larger than inter-model differences between the CSM1.4, CCSM3 or CESM simulations. Therefore, this suggests that internal variability could have had major impacts on the temperature response at hemispheric scales during the MCA to LIA transition.

Regarding the temperature differences shown by the reconstructions, a tendency to larger MCA-LIA differences in the NH-C series than in the continental targets is observed. Indeed, the larger range of temperature differences, in comparison to the simulated evidences, is associated to these NH-C reconstructions. These values could be already expected from previous evidences of a warmer MCA period than the ones presented in other reconstructions (Section 5.1.1) as well as in the simulations (Figure 6.2).

The changes from the LIA to present are also characterized by positive differences (present-LIA; Figure 6.10c,d) in forcing and in temperature changes both in simulations and reconstructions, evidencing in all cases a larger range of differences than in the MCA-LIA transition. Figure 6.10c shows a similar range of change both for simulated and reconstructed temperatures and, in contrast to Figure 6.10a, there is no clear difference between the STSI and ssTSI experiments (Masson-Delmotte et al., 2013). Changes in forcing (Figure 6.10d) are also not grouped according to the level of solar variability forcing relate more to the inclusion of anthropogenic forcings. For example the ECHO-G, FGOALS or CCSM3, that do not consider anthropogenic aerosols forcing (see Table 2.1) evidence the

largest forcing variations, while those simulations including land use changes as well as the Forster et al. (2007) estimation of anthropogenic aerosols (Section 3.2) show the lowest differences (e.g. EC5MP, HadCM3). For the present-LIA transition the number of available reconstructions is larger than for the MCA-LIA (Table 2.3) and the range of temperature differences is in good agreement with the one from the simulations. A slight tendency to larger temperature differences is observed for the continental reconstructions (e.g. Frank et al., 2007; Huang, 2004) than for the NH-C ones (e.g. Mann et al., 2008; Moberg et al., 2005).

6.4.1 Spatial variability of temperature changes

The relative roles of internal versus forced variability can be also discussed by considering the spatial distribution of simulated and reconstructed temperature changes associated to the MCA, LIA and the 20th century transitions described above. Figure 6.11 shows, similarly to Masson-Delmotte et al. (2013), the spatial temperature differences associated to the MCA-LIA, present-MCA and present-LIA transitions for the STSI and ssTSI simulations as well as for the Mann et al. (2009) reconstruction. The latter corresponds to the only global climate field reconstruction (CFR) available that spans the periods of interes. The temperature changes of the simulations are represented by the average of all the maps of temperature differences in all memebers of the ensemble (see caption). Figure 6.11 is virtually identical to the one shown in AR5 (Masson-Delmotte et al., 2013). Only one additional simulation (CESM; Lehner et al., 2015) is added.

The three transitions are characterized by a more spatially homogeneous response in the simulations than in the reconstruction, in which various regional features are evidenced. Indeed, the simulated difference patterns for all the transitions show a spatially homogeneous warming that illustrates larger differences, particularly in the STSI group, over the continents than the oceans, and larger over the sea-ice boundary at high latitudes, in agreement with the temperature response pattern described in Zorita et al. (2005).

Within the three transitions analyzed, the largest simulated and reconstructed changes are between the LIA and the present (Figure 6.11g-i). Reconstruction support the general warming evidenced by the simulations (Figure 6.11g,h) for this period transition, except for a cooling south of Greenland (Figure 6.11i; Masson-Delmotte et al., 2013). For the differences between the MCA and present a similar pattern, although less intense, of temperature changes to the present-LIA differences is observed in Figure 6.11d-f. Larger differences with respect to the present-LIA transition are observed for the reconstruction (Figure 6.11f), in which the cooling south of Greenland as well as other regional coolings at mid-latitudes are more pronounced. The latter could be expected from the larger

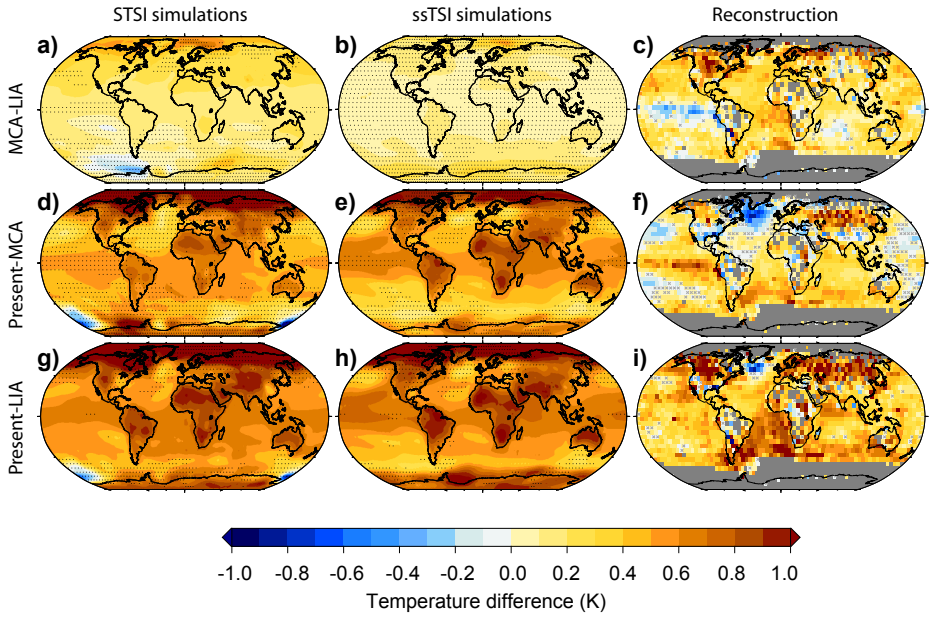

Fig. 6.11: Simulated and reconstructed temperature changes for key periods over the LM. Annual mean temperature differences for the: (top, a-c) MCA (950-1250 CE) minus LIA (1400-1700 CE); (middle, d-f) Present (1900-2000 CE) minus MCA; (bottom, g-i) Present minus LIA. Model temperature differences are average temperature changes in the ensemble of available model simulations, grouped into those applying STSI (left column) and ssTSI forcing (middle column), respectively. The right column shows the differences for the Mann et al. (2009) climate field reconstruction. Dotted areas indicate non-significant differences according to a two sided t-test ($\alpha < 0.05$) in reconstructed fields or that <80% of the simulations showed significant changes of the same sign. Only simulations spanning the whole millennium and including at least solar, volcanic and GHG have been used (Table 2.1). For the simulations starting in 1000 AD (CCSM3, ECHO-G, CNRM, FGOALS) the period 1000 to 1250 was selected instead to define the MCA.

temperature changes observed in reconstructions in comparison to the simulations during the MCA-LIA transition (Section 6.4). Indeed, a greater regional variability is expected in the reconstructions compared with the mean of mul-

tiple model simulations (Masson-Delmotte et al., 2013). Nevertheless, regional differences over areas with limited proxy data availability, such as the oceans, should be analyzed cautiously.

Figures 6.11a,b show the pattern of mean temperature differences for the MCA-LIA obtained for STSI and ssTSI simulations, respectively. A widespread warming is simulated, although of less amplitude than in other periods due to the smaller variations in forcing during the MCA-LIA transition. This warming is also less intense than the Mann et al. (2009) reconstruction and not singnificant in most of the global domain. Apart from the larger temperature differences, the reconstruction also shows negative anomalies in the Tropical Pacific not reported by the simulations.

Many studies (e.g. Seager et al., 2007; Mann et al., 2009; Graham et al., 2011) suggest that during this period there was a pattern of coordinated temperature and hydrological anomalies evidencing an increased zonal gradient in the tropical Pacific produced by anomalous cooling in the eastern Pacific and anomalous warmth in the western Pacific and Indian Ocean. Additionally, a broad expansion of the Hadley Cell with an associated northward shift of the zonal circulation might have led to a more positive North Atlantic Oscillation (NAO) like signature (Graham et al., 2011; Trouet et al., 2009, 2012; Ortega et al., 2015). The relative importance of forcing and internal variability in producing this coordinated pattern of anomalies are not clear. Mann et al. (2009) shows in Figure 6.11c a reconstructed pattern of MCA-LIA temperature change indicating enhanced and pervasive cooling in the eastern equatorial Pacific cold tongue region, often referred to as La Niña-like background state, as well as positive anomalies dominating at mid and high latitudes of the NH. Mann et al. (2009) also showed that the negative anomalies in the eastern equatorial area were not reproduced by forced simulations with the GISS-ER and CSM1.4 models. Extratropical warmth was also reported by Ljungqvist et al. (2012) and was found to be consistent with results of assimilation experiments (Goosse et al., 2012b,c) in response to a weak solar forcing and a transition to a more positive Arctic Oscillation state. AOGCM experiments without data assimilation however do not seem to support an enhanced zonal circulation during medieval times (Lehner et al., 2012; Yiou et al., 2012), in contrast also with the most recent reconstructions of North Atlantic variability (Ortega et al., 2015).

As well as the mean model simulations shown in Figure 6.11a,b, that do not reproduce the regional features illustrated by Mann et al. (2009), the spatially distributed temperature differences for the MCA-LIA transition for each single model simulations are shown in Figures 6.12 and 6.13 (González-Rouco et al., 2011). This allows analysing the regional features evidenced by the individual simulations for a proper comparison to the reconstructed pattern.

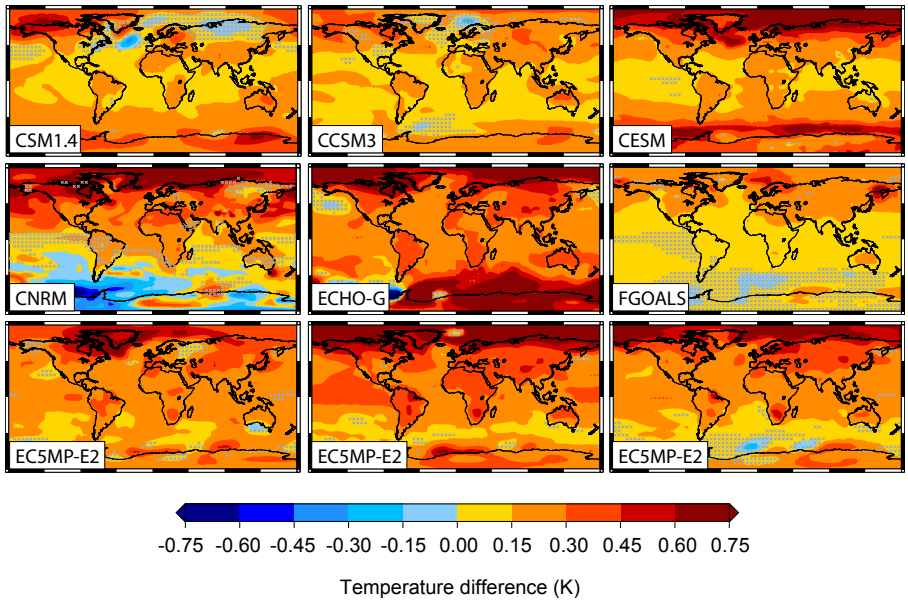

Fig. 6.12: MCA-LIA (950-1250 AD minus 1400-1700 AD) annual mean temperature difference in the STSI forced simulations used in Figure 6.11 (Table 2.1). Hatched areas indicate non significant differences according to a two sided t-test ($\alpha < 0.05$).

All simulations tend to produce an MCA warming that is almost globally uniform, particularlly high for STSI (Figure 6.12). In spite of this, there are considerable differences among the simulations, which highlight a potential influence of initial conditions as an expression of internal variability. For the ssTSI group (Figure 6.13), the temperature response for the MCA-LIA is less uniform in sign and more likely influenced by internal variability. Moreover, it also presents lower values than in the Mann et al. (2009) pattern. Therefore, under both high and low TSI change scenarios, it is possible that the MCA-LIA reconstructed anomalies would have been largely influenced by internal variability (González-Rouco et al., 2011; Fernández-Donado et al., 2013).

The agreement among the suite of simulated MCA-LIA temperature differences can be quantified by the values of inter-model spatial correlations between all the possible combinations of two simulations within the STSI and ssTSI groups. Slightly larger correlation values are obtained for ssTSI (r=0.7 between

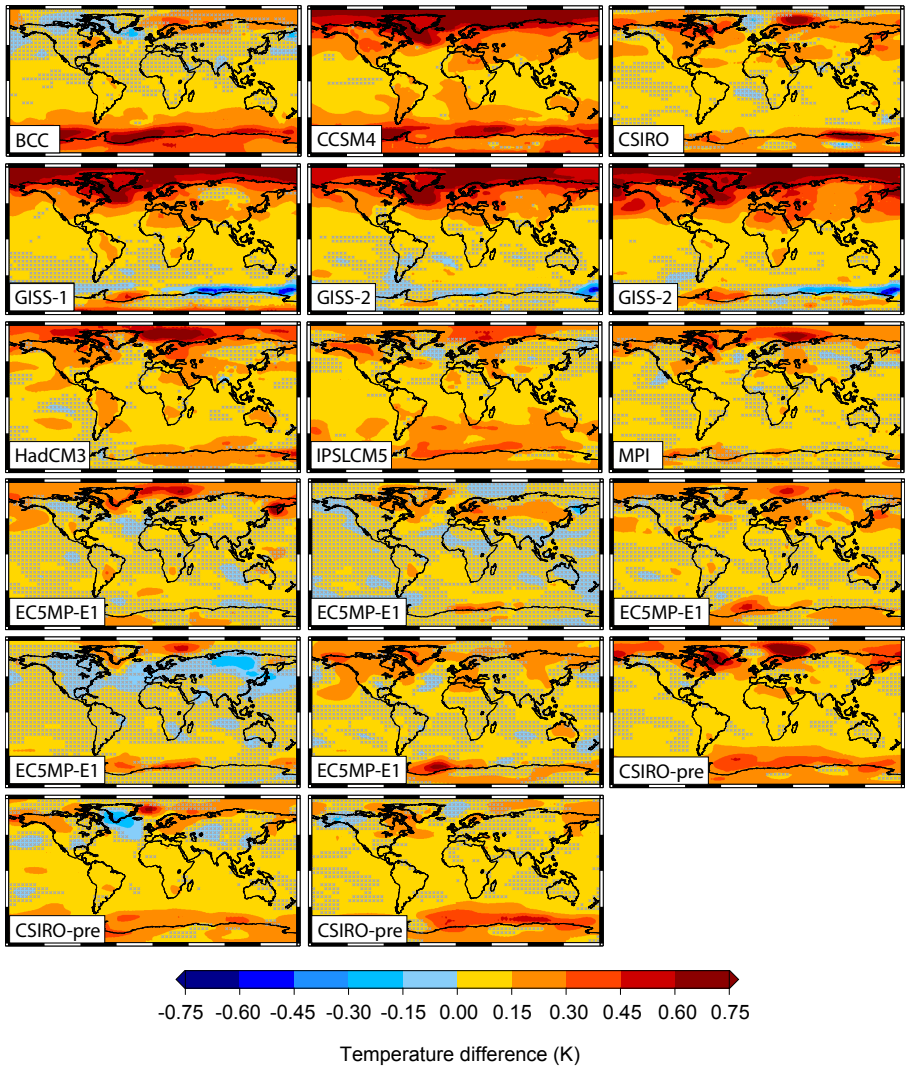

Fig. 6.13: As Figure 6.12 for the ssTSI simulations.

CCSM4 and MPI) than for the STSI (r= 0.69 in the case of CNRM with a member of EC5MP-E2), contrary to the larger TEF changes applied to STSI. There also appear multiple regional features that are dependent on the specific model experiment considered, i.e. cooling in the North Pacific (ECHO-G), in the North Atlantic (CCSM3, IPSLCM5), in North America (BCC) or in Northern Asia (CNRM, MPI, CSM1.4). This causes the lowest inter-model spatial correlation values in the STSI (e.g. r=-0.09 between CNRM and CSM1.4) and in the ssTSI group (e.g. r=-0.25 between BCC and members of GISS).

Many of these regional scale features may well be simulation-dependent and related to initial conditions and internal variability as evidenced by the differences within the members of each EC5MP, the GISS or the CSIRO-pre ensembles. For example, differences arise in the magnitude of warming and cooling over the North Pacific, South America or Africa in EC5MP-E2, over Asia or Antarctica in GISS or in the spread of cooling regions in the EC5MP-E1 and CSIRO-pre members. These diferences within an ensemble lead to intra-model spatial correlation values that range between 0.21 and 0.52 for the EC5MP-E1, between 0.05 and 0.52 for the CSIRO-pre, between 0.60 and 0.81 for the EC5MP-E2 subensemble and between 0.84 abd 0.88 for the GISS. Among the different subensembles, EC5MP-E1 and CSIRO-pre simulate more regional/large-scale widespread cooling, a sign of the lower weight of TSI changes that allows for internal variability to become more prominent. This fact can be observed from the wider range of intra-model correlation values. Therefore, even if widespread warming is simulated in the MCA, the spatial pattern of temperature change is very heterogeneous and can considerably vary across models and even across simulations with the same model.

The spatial pattern observed in the reconstructions by Mann et al. (2009) is not obtained with any of the available model simulations. This is evidenced by the low values of spatial correlation obtained between the reconstructed and the suite of simulated patterns, ranging from -0.21 (GISS) to 0.40 (BCC). The lower values obtained for these simulated-reconstructed pairs than for the inter-model comparison suggest a higher consistency among the simulated than between the simulated and reconstructed patterns.

Mann et al. (2009) is the only spatial reconstruction that offers global scale information about the MCA-LIA transition, and although supported by several studies (Seager et al., 2007; Graham et al., 2011), it is also subjected to important uncertainties (Li and Smerdon, 2012; Smerdon et al., 2011). These uncertainties are mostly associated to the reconstruction methodology and the low proxy replication in the Pacific and North Atlantic basins. However, if this proxy-based reconstruction were to be considered reliable, two possible explanations could be suggested for the aforementioned model-data discrepancies. One is that the spatial pattern of changes for the MCA-LIA was largely influenced by internal

variability and therefore not expected to be reproduced by models. The other is that transient simulations with AOGCMs fail to correctly reproduce some mechanism of response to external forcing, as long as the changes in radiative forcing factors are considered to have contributed importantly to the the MCA-LIA temperature change. One example of the latter may be discussed in relation to the so-called "ocean thermostat" mechanism (Zebiak and Cane, 1987). The complex response of tropical Pacific to radiative forcing still shows important inter-model disagreement in future climate change simulations (Collins et al., 2010). It is thus expected that AOGCMs will struggle to correctly represent potential responses of the thermostat in the past.

6.5 Conclusions

Within this chapter, a variety of model-data comparison exercises has been presented to asses the consistency between temperature simulations and reconstructions over the CE. Overall, reconstructed and simulated temperatures tend to agree on multicentennial timescales. The period spanned between 800 and 1200 CE is characterized, nevertheless, by higher temperatures on the side of the reconstructions that are not supported by the simulations and the reconstructions of external forcing. The latter is specially true if we compare the simulations with the complete suite of avaible hemispheric or global reconstructions. However, given the differences between the subgroups of NH CE reconstructions evidenced in Chapter 5, comparisons between LM simulations and reconstructions have been also evaluated at considering common spatial targets for the NH.

For the first half of the second millennium, particularly for the MCA, the simulations are no longer embedded in the ensemble spread of the reconstructions if we consider the complete hemispheric and land-only target domains discussed in Chapter 5. Interestingly, the NH-C (including land and ocean) group yields the largest disagreement between reconstructions and simulations while the land only groups compare better with the simulated climate. The causes for the model-data discrepancies during the first half of the LM are unclear. The maximum temperatures achieved by the simulations can be related to the maximum in the external forcing used to drive models, a combination of solar forcing and absence of volcanic activity that peaks around 1100 CE for all model experiments. Therefore, the discrepancies between simulations and reconstructions can ultimately be traced to differences between reconstructed temperatures and forcing. It is unlikely that the quality of solar and the dominant volcanic forcing diminished during this time, thus, the differences highlighted could arguably be related to reconstruction and proxy issues. Alternatively, it is also possible that the reconstructions are indicative of variability during the early part of the LM

that is related to internal processes (Schurer et al., 2013; Goosse et al., 2012b) or externally forced processes that are not captured by the models (e.g. solar amplification; Swingedouw et al., 2010; Meehl et al., 2009). Nevertheless, two findings shall be highlighted that arise when different spatial targets are considered: the NH-C and NH-LO are different in the reconstructed climates and model-data comparison differences are more evident for the NH-C ensemble. This result is also evidenced for the SH and GL scales, in spite of the lower availability of reconstructions. Independent of the origin of the discrepancies, it is clear that simulations and reconstructions should be compared over consistent spatial domains and consider differences in reconstruction methods.

The analysis of the temperature differences associated to MCA-LIA transitions at hemispherical scales is consistent with an impact of TEF changes on the reconstructed and simulated temperature changes. Thus, simulations whose TEF evidenced larger forcing differences for such temporal transition tend to show larger temperature differences. Nevertheless, the influence of the climate sensitivity and the internal variability is also observed, specially at analysing the differences observed within an ensemble of simulations from the same model and just differing in the initial conditions. Reconstructions show a larger spread of temperature differences than the simulations, highlighting then a warmer MCA on the side of the reconstructions.

Despite both models and reconstructions agree on depicting a warmer MCA followed by a colder LIA, the spatial distribution of changes during this transition is characterized in a different manner for both sources. Simulations show a land-ocean thermal response with polar amplification at high latitudes, in which internal variability contributes to pronounced inter-model differences. In turn, the available reconstructed spatial pattern (Mann et al., 2009), illustrates a much more regional variability, highlighting a La Niña-like state in the Pacific, that is not reproduced by any of the simulations considered. If we rely on the information provided by multiproxy reconstructions, it is arguable that either the spatial pattern of changes for the MCA-LIA was largely influenced by internal variability or that transient simulations fail to correctly reproduce the potential causal mechanisms of response to external forcing.

The internal variability associated to the MCA spatial distribution of temperatures is also evidenced when the differences between the present and the MCA are analysed. In contrast to the spatially homogeneous warming shown by the simulations, the reconstructed pattern illustrates more regional variability. In turn, for the present-LIA transition a larger consistency is found between simulations and reconstructions, suggesting a larger influence of the external forcing during both periods.

The influence of the external forcing in the temporal variability in simulations and reconstructions at interannual and decadal timescales is assessed by analysing the impact of given forcing events through a superposed epoch analysis. The former illustrates that although at interannual timescales the response to the largest volcanic eruptions are more pronounced in simulations than in reconstructions, at multidecadal timescales, the magnitude of the response to a cluster of volcanic events in both sources is quite similar. In addition, reconstructions tend to need more time to recover pre-eruption temperature values than simulations.

Based on the LMTCR values calculated in Chapters 4 and 5 for the simulations and reconstructions, respectively, a quantitative model-data comparison is also offered at global and hemispheric scales. LMTCR values, as a quantification of the linear temperature response to TEF over the LM, show, for both simulations and reconstructions positive and consisitent values around 0.3 - 0.4 K/Wm^{-2}. Nevertheless, simulations tend to show larger LMTCR values than reconstructions for all the spatial domains considered. The latter is particularly true for the complete domains (NH-C, SH-C, GL-C), including land and oceanic areas, while the continental reconstructions evidence larger consistency with simulations. The level of consistency between simulations and reconstructions is quite similar for the various TEF configurations considered. Alternatively, some reconstructions are sistematically not consistent with the simulated ranges of LMTCR for any, or almost any, of the available TEF configurations. This is the case, for instance, of Mann et al. (2008), Leclercq and Oerlemans (2012), Ammann et al. (2007), (Neukom et al., 2014) and Christiansen and Ljungqvist (2012b) reconstructions. Other reconstructions (e.g. Briffa et al., 2001; Loehle and McCulloch, 2008; Mann et al., 2009) are partially consistent with the simulations, with LMTCR vaues that disagree with the simulated ones for certain TEF configurations and agree for others. Finally, some reconstructions agree with the simulated ranges for all or nearly all the TEF configurations (e.g. Hegerl et al., 2007b; Frank et al., 2007; D'Arrigo et al., 2006).

7

Model-data comparison at continental scales*

Chapters 4 to 6 have focused on analysing the temperature response of simulations and reconstructions, at global and hemispheric scales, during the CE. Both simulations and reconstructions evidence, at these spatial scales, a large influence of the external forcing. Indeed, at multidecadal, and longer, timescales, the temperature from both independent sources of information presents linear responses to the total external forcing. The influence of the internal variability is also noticed at these scales and is evidenced in the differences detected within the members of an ensemble of simulations driven by a common external forcing configuration. The spatial distribution of the temperature differences between some key periods shows a more limited consistency. As discussed in Section 6.4.1, the spatial patterns obtained by the reconstructions are not reproduced in the simulations and additionally, the simulated patterns also present noticeable regional differences. The latter is suggestive of the larger influence of internal variability at sub-hemispherical scales.

Within this context, an analysis of the role played by internal variability and external forcing in the temperature responses at continental scales is presented in this chapter. For this purpose, the collection of PAGES2k reconstructions (Section 2.2.2) will be assessed for consistency with the available simulations. Specifically, Section 7.1 presents and compares the temporal and the spatial variability of the simulations and reconstructions over the continental-scale regions consid-

* The main contents of this chapter are included in:
- PAGES2k-PMIP3 group, 2015: Continental-scale temperature variability in PMIP3 simulations and PAGES 2k regional temperature reconstructions over the past millennium. *Climate of the past.* **Under review**.
- EuroMed2k Consortium, 2015: European summer temperatures since Roman times. *Environmental Research Letters.* **Under review**.

ered. Section 7.2 analyses the regional response in simulations and reconstructions to given forcing events at interannual and decadal timescales. In Section 7.3 the reconstruction for the European region (EuroMed2k Consortium, 2015) is analysed in more detail. Section 7.4 summarizes the main conclusions extracted from the present Chapter.

7.1 Reconstructed and simulated temperatures at continental scales

This section focuses on the PAGES2k reconstructions (Table 2.4) together with the ensemble of available simulations at continental scales. The simulations selected, as in previous chapters, are those from Table 2.1 spanning, at least the 850 - 2000 CE period and considering a minimum set of external forcing factors that includes solar, volcanic and GHG estimations.

Figure 7.1 shows the reconstructed series for the Antarctica, Arctic, Australasia, Asia, Europe, North America and South America regions (PAGES 2k Consortium, 2013) together with the forced simulations for which the same spatial domain has been selected as in the reconstructions (PAGES 2k Consortium, 2013). This Figure complements Figure 1.3 and Figure 5.12 in AR5 (Masson-Delmotte et al., 2013), by illustrating all the available PAGES2k regions with the updated ensemble of simulations.

Some general features can be noticed for the majority of the regions. Most of the reconstructions are characterized by evidencing a cooling trend during the LM, followed by a warming during the industrial period. The simulations also tend to evidence a similar temporal evolution than at hemispherical and global scales (Figures 4.1 and 4.4), with a warmer MCA, followed by a later cooling and the recent industrial warming. The reconstructions are, over most of the LM for all the regions, embedded within the ensemble of simulations, although the former tends to evidence a larger centennial variability than the numerical experiments. Both simulations and reconstructions agree on depicting at multidecadal timescales anomalies for some of the regions (e.g. Europe, Asia) associated to volcanic clusters (Figure 3.1). These peaks, as well as the warming trends evidenced during the last century, are an initial suggestion of the influence of the external forcing in the temperature signal at continental scales.

Nevertheless, each reconstruction shows also features not reproduced by simulations or the external forcing, that may be suggestive of the influence of the internal variability or methodological issues that concern the reconstructions or the simulations. The Antarctica reconstruction shows since 800 CE the long-term cooling trend mentioned above and does not report any warming trend during

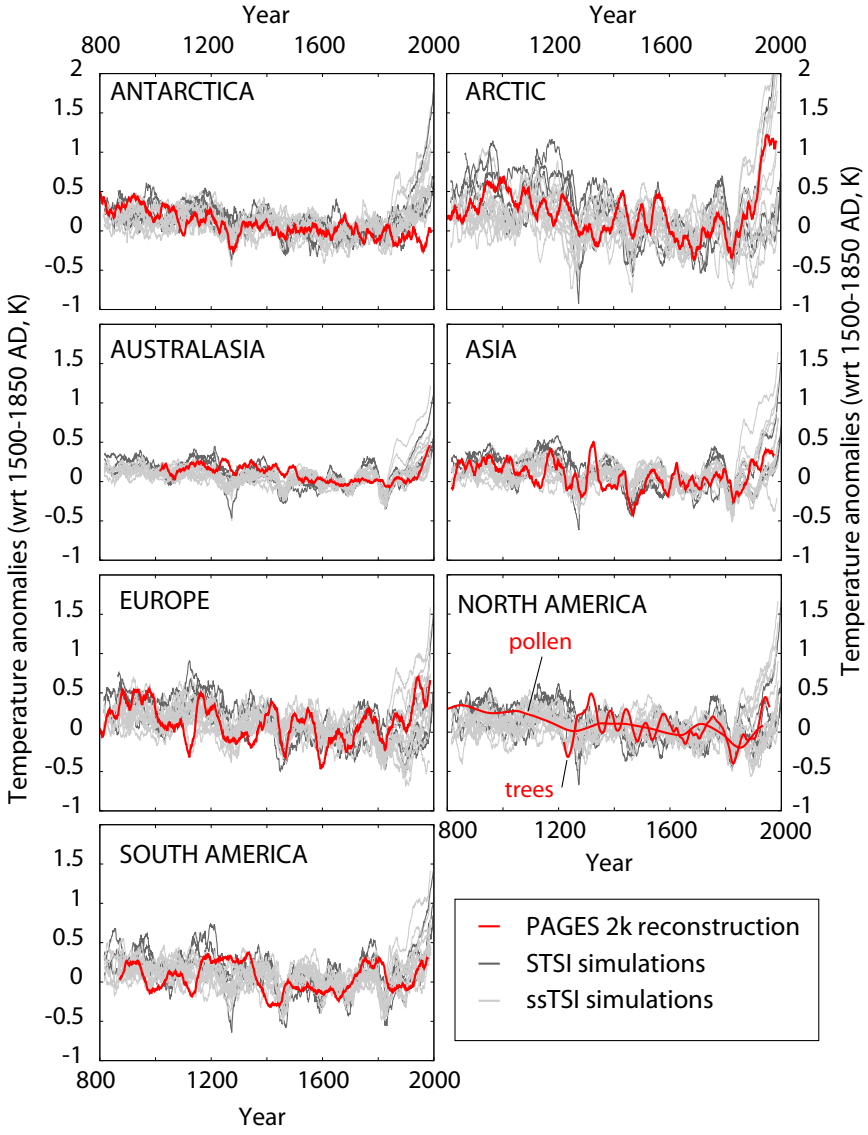

Fig. 7.1: Temperature anomalies from PAGES2k reconstructions (red) and simulations (grey) after being spatially averaged over the same area considered by the reconstruction. The two groups of simulations, according to the level of solar variability forcing applied, STSI (black) and ssTSI (gray), are distinguished. Anomalies are calculated with reference to 1500 - 1850 CE. All series are 31-yr low-pass filter outputs.

the industrial period as already evidenced by Goosse et al. (2012a). The latter implies the largest disagreement with the simulations that show a warming trend following the TEF (Figure 3.4). The Arctic region is characterized in simulations and reconstruction by a similar large low-frequency variability (Kaufman et al., 2009), evidencing the three key periods of the LM. Nevertheless, for the first half of the LM the timing of the multicentennial variations in the Arctic region does not agree on simulations and reconstruction. Australasia is the region with the lowest low frequency variability in simulations and the reconstruction, probably related to the large proportion of oceanic areas included within this region. During the 1200 - 1500 CE period the Australasia reconstruction tends to show slightly warmer temperatures than most simulations, and for the industrial period the reconstructed warming trend is similar to those found in the simulations. For the Asia region a large variability at multidecadal timescales is observed for the reconstruction than simulations. The last century warming trend evidenced by the reconstruction is embedded within the ensemble of simulations. The latter is also observed in the European region, with a strong agreement between simulations and the reconstructions since ca. 1200 CE in the multidecadal variability. In North America simulated low frequency multicentennial trends are in overall agreement with the pollen-based reconstruction, albeit with some discrepancies between the 17th-20th centuries. On the contrary, the best agreement between the tree-ring-based reconstruction and the simulations is obtained for that period due to the influence of the volcanic activity. The South America reconstruction evidences larger centennial variability than simulations and are in better agreement with the simulations during the second half of the LM.

Complementary insights for characterizing the temperature variability in simulations and reconstructions is offered with the spectral analysis of the series, shown in Figure 7.2. Most of the reconstructions, except for the Arctic and Europe, tend to evidence less variance than models at interannual timescales relative to decadal and multicentennial; note that spectra are normalized and spectral densities indicate percentages relative to unit variance. Indeed, all regions except Europe are characterized by a larger accumulation of variance in the models at the 3 - 5 yr timescales due to the EC5MP simulations (Chapter 6; Figure 6.3; Jungclaus et al., 2010), not supported by any of the reconstructions. The Arctic evidences the best model-data agreement, in which the reconstructions spectra is embedded within the simulations for all the frequencies. Europe, Antarctica, Australasia and Asia reconstructions show a general good agreement with the simulations, although there is a tendency to show larger variance than simulations within the 50 - 100 yr band. In turn, South America shows the largest differences between the reconstructed and simulated spectra, except for the 20 - 50 yr time band in which simulations and reconstructions show a similar spectral

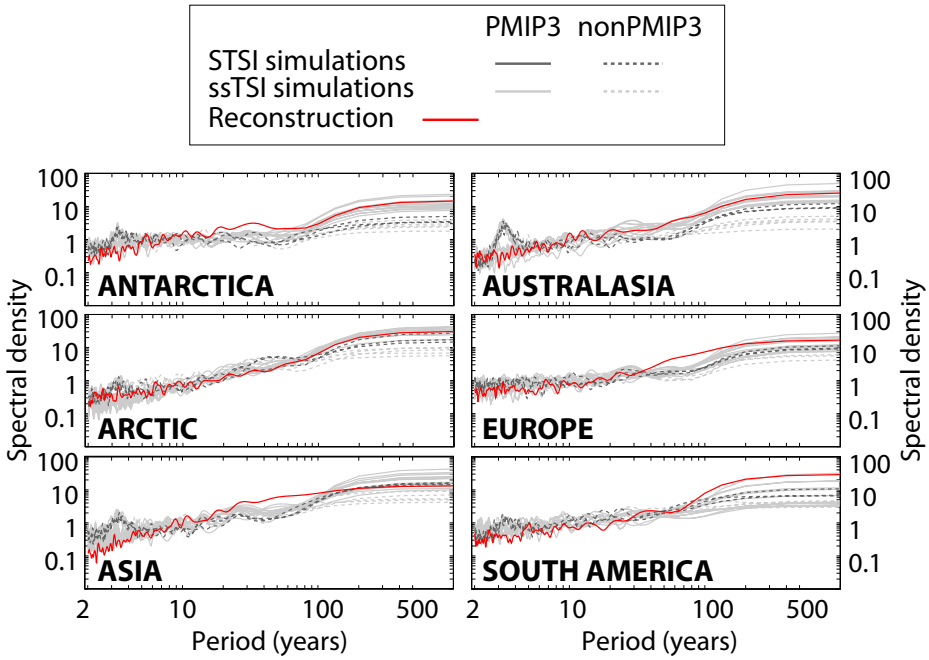

Fig. 7.2: Normalized spectral densities for simulations (grey) and reconstructions (red) for the PAGES2k regions over the 850 - 2000 CE period. North America is excluded from this analysis due to the lack of high frequency variability in the pollen-based reconstruction and the short length of the tree-ring-based one.

density. The reconstruction tends to underestimate (overestimate) the simulated variances at high (low) frequencies.

An additional model-data comparison exercise for exploring the agreement between the simulations and the reconstructions for each region is proposed in Figure 7.3, through the correlation analyses between both sources. Thus, for each simulation, Figure 7.3a shows the correlation indexes obtained between each PAGES2k reconstructions and the simulation averaged over the corresponding region after a 31-yr low-pass filter. The better agreement evidenced previously for some of the regions, such as the Arctic, Europe or Australasia, is quantified. Thus, Figure 7.3a where the largest correlation values are attained for the Arctic region in the PMIP3 simulations. Correlations for Australasia and Europe tend to be also significant and high in the PMIP3 simulations and in the STSI

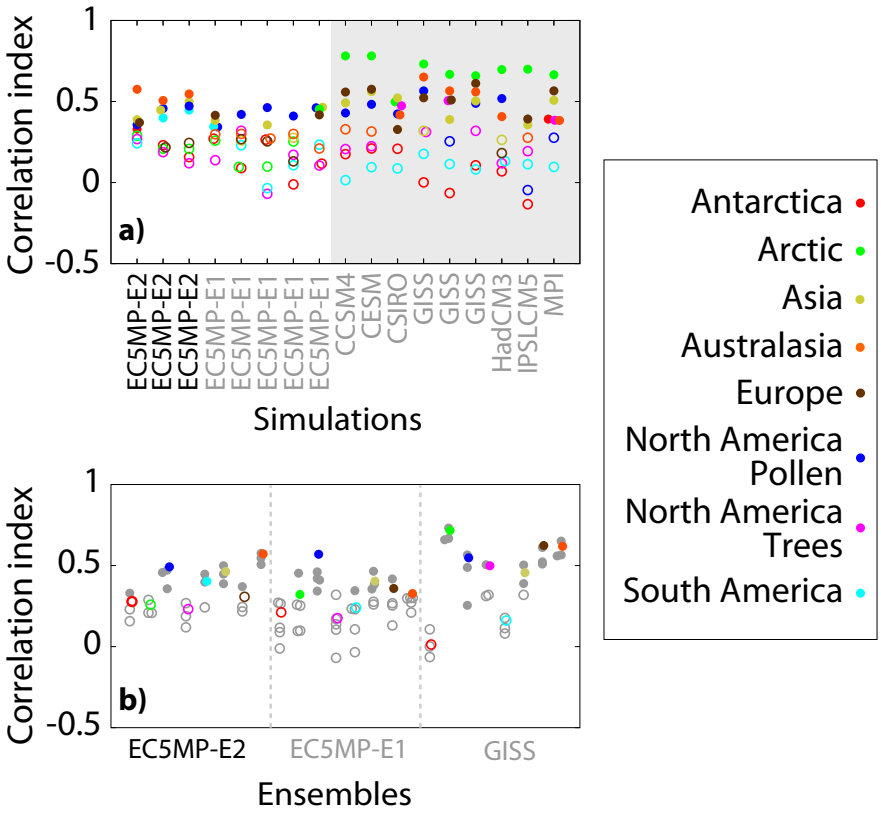

Fig. 7.3: a) Correlations between 31-year moving average PAGES2K reconstructions and simulations. Dots represent the correlation value between the reconstruction (see legend for region-color) and the simulation averaged over the corresponding region. Full (empty) circles stand for significant (non-significant) correlation values. Correlation indexes obtained for nonPMIP3 simulations are plotted over a white background, while the gray one corresponds to the PMIP3 runs. The name of the simulations driven by a high variability solar forcing scenario (STSI) are written in black while those using a low variability solar forcing scenario (ssTSI) are in gray. b) As in a), but only for the available model ensembles; individual ensemble member correlations are shown in gray and the ensemble average in color.

EC5MP simulations. The North America pollen and the Europe reconstructions show significant correlations in most of the simulations. Correlations tend to be non significant for the North America trees, South America and Antarctica reconstructions. Figure 7.3b shows, for the model ensembles considered (EC5MP and GISS), the correlation indexes obtained between the reconstruction and the ensemble mean, as well as with each individual members. The correlation of the ensemble mean with the regional reconstructions is always higher than the average of all individual correlations. Through considering the ensemble average, the internal variability of the simulations in the ensemble is reduced in favor of the response to the external forcing. Therefore, this supports the existence of a relationship between external forcing and reconstructed temperature that will be analysed in Section 7.2.

7.1.1 Inter-regional consistency

Neukom et al. (2014) showed that simulations tend to present a more homogeneous climate within both hemispheres than reconstructions. This homogeneous response has also been evidenced in Chapter 6 during the analysis of the spatial distribution of temperature differences associated to the MCA-LIA transition in simulations and reconstructions. Simulations show, in general, an homogeneous pattern of temperature change while the available climate field reconstruction (Mann et al., 2009) is characterized by multiple regional features that are not accounted for by the models. Within this context, the present section analyses the inter-regional relationships among simulations and reconstructions to assess the level of spatial heterogeneity of both sources of information.

A simple analysis presenting all inter-regional correlation indexes obtained for each simulation and the reconstructions is presented. Thus, Figure 7.4a shows for each simulation and for the ensemble of PAGES2k reconstructions the range of inter-region correlations with box-whiskers plots. Each box plot indicates all the range of the correlations obtained for all possible pairs of compared temperature series defined for the PAGES2k regions within each model experiment. The inter-regional correlation values obtained specifically for the European region with the other spatial regions of the ensemble of simulations is shown in Figure 7.4b as an example. Results evidence that the highest correlations are always attained for the simulations. There are some models that even produce inter-regional correlation ranges that are out and above the reconstructed range (CCSM4, CESM, CSIRO, IPSLCM5 and MPI), thus indicating that inter-regional correlations are significantly larger in these models than in the reconstructions. These models overstate the homogeneous structure of the spatial covariability. In turn, the HadCM3 simulations shows the largest range of correlation values, evidencing a larger spatial heterogeneity than in the other model simulations. The EC5MP and

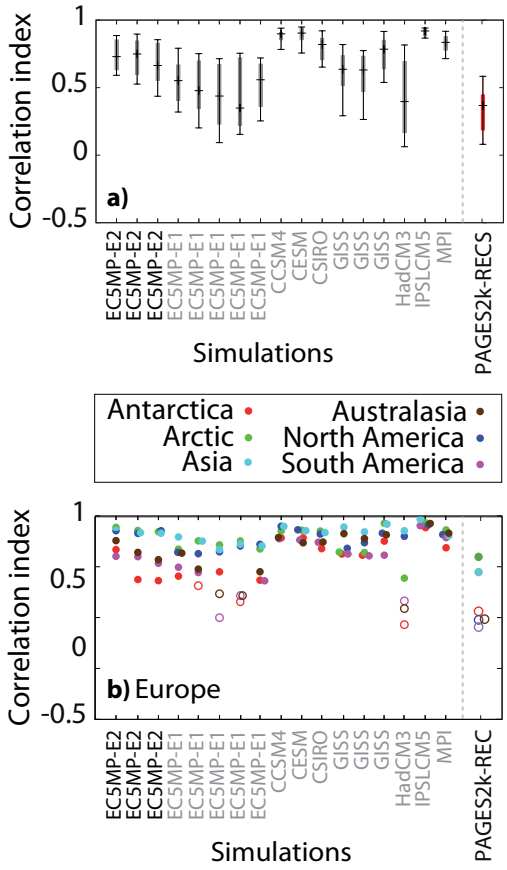

Fig. 7.4: a)Inter-region correlations within each simulation and reconstructions. The range of all possible correlations between two regions obtained is shown for each simulation and for the reconstructions with a box-whiskers plot and depicting the three quartiles and the 10% and 90 % whiskers. b) Inter-region correlations obtained for the European region as an example of the values included in the boxplots of a).

the GISS simulations present high values of inter-regional correlation, although there are some differences observed within the ranges obtained for the various members of each ensemble. These differences are associated to the influence on the internal variability on the different spatial regions. Therefore, internal variability can contribute to shape the level of spatial heterogeneity within the climate of different continents. In general the differences between members of an ensemble are relatively small in comparison to inter-model changes; nevertheless one member of the GISS ensemble shows differences with the other ensemble members that are comparable to using a different model. EC5MP, GISS and HadCM3 overlap with the reconstructed range but the highest values are always obtained within the simulated climate. The latter supports the previous statement that models are more homogenous than reconstructions in their spatial response. Figure 7.4 shows that inter-regional correlations tend to be spatially consistent among models and, at some extent, with the PAGES2k reconstructions. For instance, the European region shows higher correlations in all models and in the reconstructions with Asia and the Arctic, and lower with South America and Antarctica.

7.2 Response to external forcing at interannual and decadal timescales

The influence of the external forcing in the PAGES2k temperature reconstructions and simulations is analysed based on a superposed epoch analysis (see Chapter 6). The response from simulations and reconstructions to the strongest volcanic events and intervals of lower solar activity at interannual and decadal timescales is evaluated. Comparing the intensity of the model response to the imposed forcing and the one found in the reconstructions allows for assessing whether the magnitude of the model response is consistent with the reconstructions at regional scales within the intervals of stronger volcanic and solar signal.

Figures 7.5 and 7.6 show, for each PAGES2k region except North America, the composite of the temperature responses from simulations and reconstructions associated to given events in the volcanic forcing reconstruction of Crowley and Unterman (2012) and Gao et al. (2008), respectively. The events have been selected as in Figure 5.9 of Masson-Delmotte et al. (2013) and described in Section 6.4. As in the case of the espectral analysis, the two reconstructions for North America are not considered due to the lack of high frequency variability (pollen based series) or the short time span(tree-ring based series). The years associated to the strongest volcanic events or to the clusters of volcanic activity differ among the regions and the volcanic forcing estimation considered *(McKay, pers. communication)*. The selection of these periods is then based on the latitudinal information provided by the volcanic forcing from Crowley and Unterman (2012)

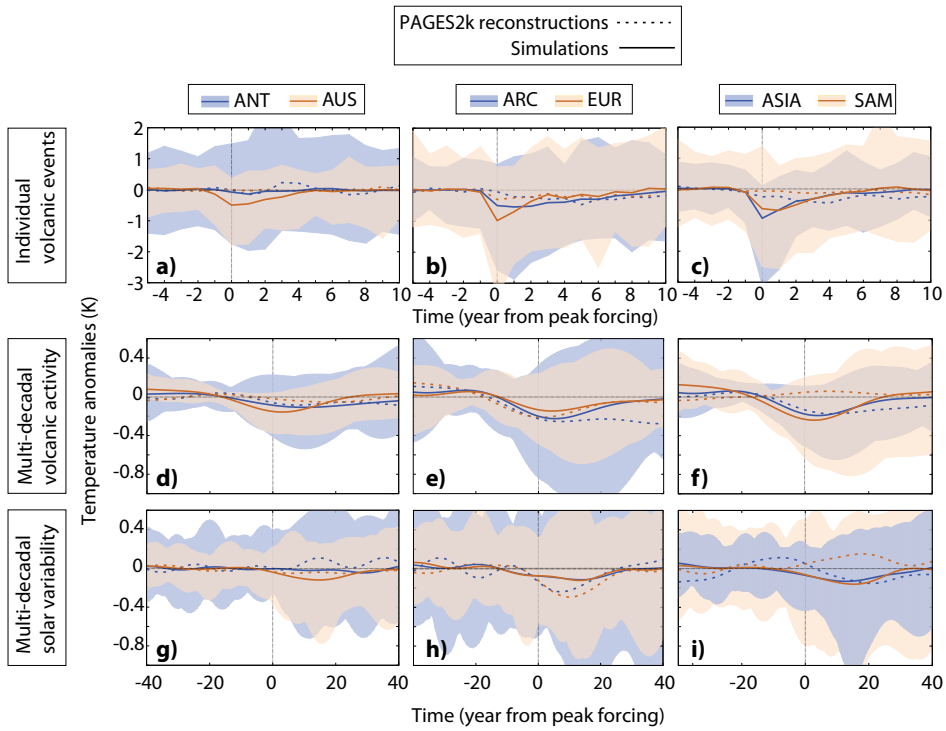

Fig. 7.5: Superposed composites of the regional reconstructed (dashed lines) and simulated (solid lines) temperature responses, considering the Crowley and Unterman (2012) volcanic forcing reconstruction, to: (a-c) 12 of the strongest volcanic events after 850 CE; (d-f) multi-decadal changes in volcanic activity; (g-i) multi-decadal changes associated to periods of low solar variability. Each column shows the responses of two regions: (left) Antarctica (ANT) and Australasia (AUS); (middle) Arctic (ARC) and Europe (EUR); (right) Asia (ASIA) and South America (SAM). The correspondence between the color of the lines/shadings and the region considered in each column is indicated in the legend. Shadings show the complete range of simulated temperature values attained by the various time intervals selected and the composite mean by a solid line. Note that only the simulations driven by the Crowley and Unterman (2012) volcanic forcing (see Table 3.1) are considered.

and Gao et al. (2008). Despite the different years considered, the composites of the regional forcing variations are, for all the cases, very similar to the hemispherical ones shown in Figure 6.9a-c. Thus, the reader is referred to that Figure for analysing the variations in the volcanic and solar forcing that are not included in Figure 7.5. In spite of the very similar regional forcing, the regional temperature response shown in Figures 7.5 and 7.6 is variable both in the simulations and in the reconstructions.

The temperature response associated to the individual volcanic events from the Crowley and Unterman (2012) series (Figure 7.5a-c) is larger in the simulations than in the reconstructions as in the case of the assessment at hemispheric scales (Section 6.4). The largest temperature responses in simulated (reconstructed) temperatures are found in Europe and Asia with composite averages of up to -1 (-0.25) K. Australia, the Arctic and South America show smaller simulated (ca. -0.5 K) temperature changes; average changes in the reconstructions are smaller but stay at levels of -0.1 to -0.2 K during several years. The temperature recovery lasts longer than the forcing decay (Figure 6.9a), on the order of 8 to 10 years, both in the simulations and in the reconstructions. For the Antarctic region, the simulated and temperature response are negligible, as well as the reconstructed response in Australia. Similar results are obtained in Figure 7.6a-c at considering the Gao et al. (2008) volcanic forcing estimations.

Figures 7.5d-f and 7.6d-f show the temperature responses to multidecadal volcanic activity for the Crowley and Unterman (2012) and Gao et al. (2008) reconstructions, respectively. Interestingly, as it was observed in the hemispherical case (Figure 6.9), the simulated and reconstructed temperature responses are in better agreement at these timescales. The simulated response is consistently larger in the experiments using the Gao et al. (2008) boundary conditions (Figure 7.6d-f) for which temperature changes range between -0.1 and -0.2 K, while in the experiments using the Crowley and Unterman (2012) forcing (Figure 7.5d-f), temperature changes range between -0.2 and -0.4 K. Contrary to simulations, reconstructions do not show multidecadal changes associated to volcanic forcing for the South American region. Both reconstructions and simulations driven by the Crowley and Unterman (2012) reconstruction agree in depicting smaller temperature responses for Australia and Antarctica. In those experiments the range of the response is in good agreement with reconstructions for Europe, Asia and the Arctic. In the Gao et al. (2008) experiments (Figure 7.6d-f), reconstructed and simulated changes are in good agreement for Europe and the Arctic but more than double the reconstructed response in Asia. However, for this latter region the composite selection according to Gao et al. (2008) renders a temperature change smaller than the one attained according to Crowley and Unterman (2012); a feature that can be sensitive to the relative small number of events

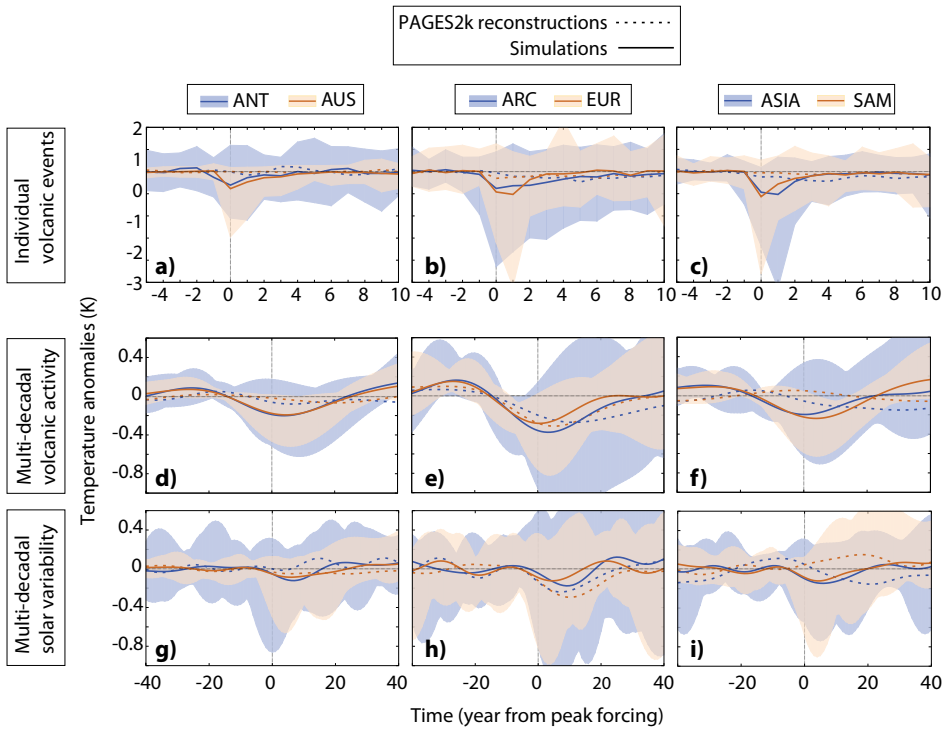

Fig. 7.6: As Figure 7.5, but for the Gao et al. (2008) volcanic forcing reconstruction.

selected for these timescales. The timescale of convergence to mean conditions is of 20 to 40 yrs both in the reconstructions and simulations.

Figures 7.5g-i and 7.6g-i show the temperature responses associated to the years of minima in solar forcing, following the selection of Masson-Delmotte et al. (2013). Due to unique selection of years for analysing the intervals of solar minima, the responses shown by the same reconstruction in Figures 7.5g-i and 7.6g-i are equivalent. Differences are, nevertheless, observed in the temperature responses shown by simulations that are driven by Crowley and Unterman (2012) or Gao et al. (2008) volcanic forcing estimations (Table 3.1). The multidecadal impact of solar forcing is most clear for Europe, the Arctic and Asia in the reconstructions with average changes ranging 0.15 to 0.25 K. Changes in model simulations are between 0.05 and 0.1 K in all regions except for Antarctica where

no changes are perceptible. Noticeably, the reconstructed changes are larger than the simulated ones in Europe and the Arctic. The results of this analysis are not independent and can also suffer from the influence of volcanic eruptions taking place within the selected time intervals as it was evidenced in Chapter 6.

The responses of simulations and reconstructions at the various PAGES2k regions reveal larger inter-regional differences within the reconstructions, while simulations tend to evidence a more similar response. The latter supports the previous statement (Section 7.1.1) of a larger spatial homogeneity in the simulations than the reconstructions.

7.3 A focus over Europe

Two additional model-data comparison exercises are addressed within the present section in order to provide further insights at the continental scale. The case of the European region is selected.

The PAGES2k Europe temperature reconstruction (EuroMed2k Consortium, 2015; PAGES 2k Consortium, 2013) shown in previous analysis within this chapter is based on the Composite Plus Scaling method (CPS; e.g. Jones et al., 2009; Mann et al., 2008). Alternatively, a climate field reconstruction based on the Bayesian Hierarchical Modeling method (BHM; e.g. Tingley and Huybers, 2010) has been also developed for this region over the period 755 - 2003 CE. The mean summer European temperature reconstructions using both methods, BHM and CPS, are shown in Figure 7.7 together with the model simulations classified according to the STSI and ssTSI subgroups. For each ensemble of simulations an uncertainty band as the ones shown in Chapters 5 and 6 for the reconstructions (e.g. Figure 6.1) is calculated. As in previous sections and chapters, this comparison considers all the simulations spanning, at least, the LM and including a minimum set of external forcings, i.e. solar, volcanic and GHG variations (see Table 2.1).

Both reconstructed approaches agree on evidencing a large multidecadal variability, largely influenced by the volcanic activity. Despite both reconstructions are very similar, BHM series tends to evidence warmer temperatures than the CPS, specially during some periods of the MCA. Around 1200 CE, the difference between BHM and CPS approaches reaches a maximum of 0.5 K. Interestingly, the temperature minimum around 1100 CE is shown similarly by the two approaches. From ca. 1300 CE onwards, both reconstructions show almost an identical temperature evolution, depicting a cooling LIA previous to the industrial warming. Both ensembles of simulations, in turn, evidence a quite similar evolution, illustrating the STSI group warmer MCA and industrial periods than ssTSI. The warmer MCA is expected from the larger amplitude of solar forcing applied

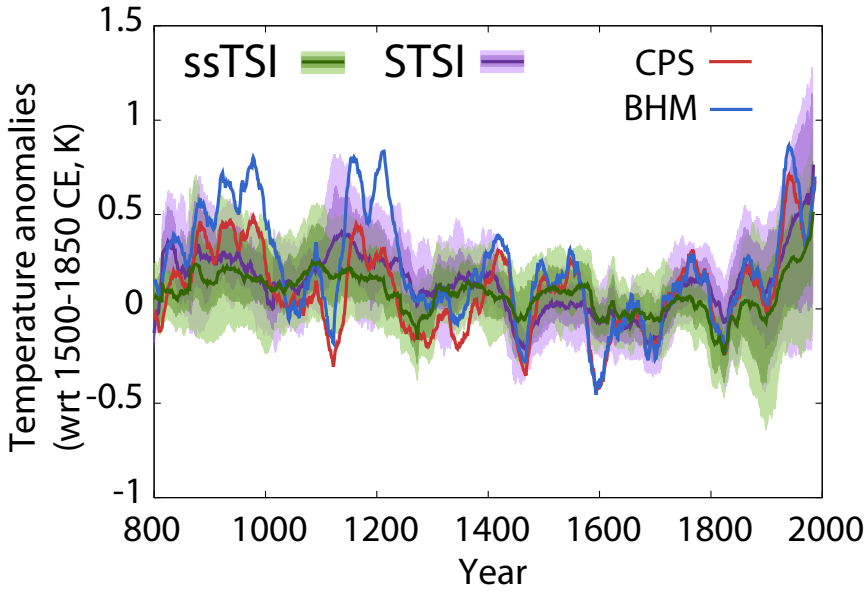

Fig. 7.7: Simulated and reconstructed European summer land temperature anomalies wrt 1500 -1850CE. The two methodological variants of the reconstructions, BHM and CPS, are shown in red and blue llines, respectively. The two ensembles of simulations considered, STSI and ssTSI (Table 2.1) are shown in purple and green colors, respectively. The ensemble mean (solid line) and the two bands accounting for 50% and 80% (shading) of the spread are shown for the model ensemble. All series are 31-yr low-pass filter outputs.

to the simulations. The warmer industrial period is, nevertheless, not related to the solar forcing but to the anthopogenic forcing configurations applied to the simulations included within the STSI group (Chapter 4; Table 2.1).

In general, the simulated multicentennial temperature variability compares well with both reconstructions. The reconstructed cold episodes at muti-decadal timescales mostly agree with simulated temperature minima attributed to solar and volcanic forcing. The reconstructed negative peaks ca. 1150 CE and 1600 CE are, nevertheless, not reproduced by the simulations. The latter may be an indicative of the influence of the internal variability or a regional response to the forcing not reproduced by the models. Another discrepancy is found dur-

ing the MCA, in which reconstructions tend to evidence a warmer temperature than simulations. The different representation of the MCA by simulations and reconstructions is consistent with the discussion for global and hemispheric scales (Chapter 6). As discussed there, if the reconstructions show a reliable range of variability, either model simulations are missing mechanisms of response to external forcing or internal variability has palyed a major role during medieval times and posterior LIA.

Figure 7.8 shows details about the spatial distribution of temperature changes associated to the transitions between key periods over the LM. Similar to Figure 6.11, the differences associated to MCA-LIA, present-LIA and present-MCA transitions are herein presented for the European region based on the simulations and the BHM reconstruction. Herein the definitions of the MCA (950 - 1250 CE) and LIA (1250 - 1700 CE) vary from those used in Chapter 6. The current definitions are based on the specific temporal evolution of the temperature observed over Europe in Figure 7.7, following EuroMed2k Consortium (2015). At these spatial scales a widespread warming is simulated for all key periods with the largest changes taking place in northern Europe. The reconstruction, in turn, evidences more regional differences than the simulations.

While the simulated pattern for both model groups qualitatively matches the reconstruction of the MCA to LIA transition, its amplitude is quite small (Figure 7.8a-c) for both sub-ensembles relative to that in the reconstructions. In addition, simulated temperature differences are not significant over the whole region, in contrast to the reconstruction.

The simulated present-LIA transition (Figure 7.8g-i) evidences a similar behaviour than MCA-LIA at underestimating the magnitude of the temperature differences in comparison to the reconstruction, specially in the ssTSI ensemble. Additionally, the agreement between different simulations is regionally limited to southern and western parts of the target region.

The spatial pattern of the reconstructed present-MCA transition shows significant differences with respect to the other periods. Apart from the significant warming evidenced for the south of Europe, a tendency to warmer MCA than present times is presented for northern and eastern Europe, although not statistically significant. The simulations, in turn, show a general warming phase between the MCA and the present over the whole spatial domain. The discrepancies found between simulations and reconstructions for the distribution of the temperatures during the MCA suggest, as in the case of the hemispherical scales, an influence of internal variability.

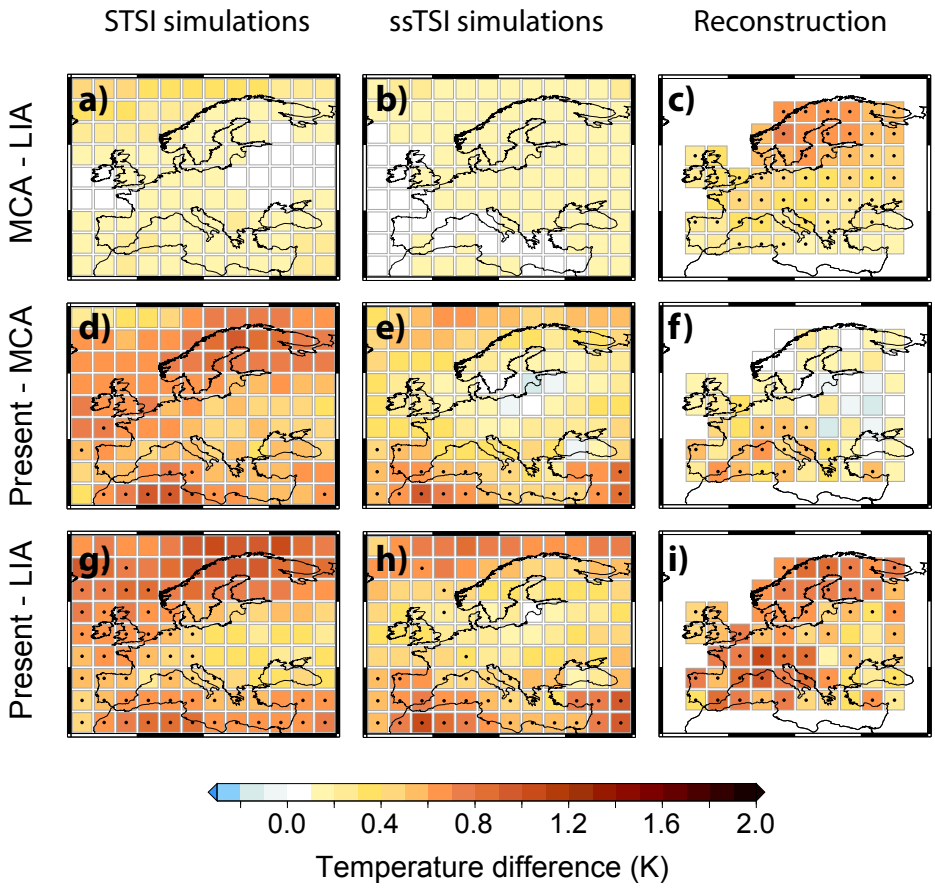

Fig. 7.8: Simulated and reconstructed temperature changes for key periods over the LM. Summer (June-August) mean temperature differences for the: (top, a-c) MCA (950-1250 CE) minus LIA (1250-1700 CE); (middle, d-f) Present (1950-2003 CE) minus MCA; (bottom, g-i) Present minus LIA. Model temperature differences are average temperature changes in the ensemble of model simulations, grouped into those applying STSI and ssTSI forcing in left and middle columns, respectively. Reconstructed temperature differences with the BHM method are shown in the right column. Dotted areas indicate non-significant differences according to a two sided t-test ($\alpha < 0.05$) in reconstructed fields or that <80% of the simulations showed significant changes of the same sign.

7.4 Conclusions

The analyses presented in this chapter have allowed us to extend some of the conclusions extracted previously from the hemispheric ang global model-data comparison exercises to continental scales.

The continental-scale simulations and reconstructions show some common signal in response to forcing changes above multidecadal timescales. This agreement is evidenced by high and significant correlation values between simulations and reconstructions, especially those from the NH. In turn, for the SH regions (Australasia, Antarctica), the agreement between both sources is more limited. The common variability evidenced in simulations and reconstructions highlights the influence of the external forcing also at these spatial scales. Indeed, the correlation increases for the ensemble average in the cases of available ensembles, thus endorsing the previous idea through cancelation of internal variability.

The response to forcing is analysed within each region through a superposed epoch analysis that tends, in general, to support the results found at hemispheric scales in Chapter 6. Thus, at interannual scales, model simulations evidence a larger response to forcing than reconstructions. Nevertheless, the reconstructed response at these timescales, although weaker, is perceptible in several of the PAGES2k regions: larger in Arctic, Europe and Asia and smaller in Australia and Antarctica, in agreement with simulations. At multidecadal timescales, the impact of the volcanic activity is similar for simulations and reconstructions. Specifically, the magnitude of the temperature response in the selected composites is in good agreement over Europe, Arctic and Asia. Interestingly, although model simulations evidence a similar magnitude of response over other areas, reconstructions in SH regions (Antarctica, Australasia and South America) tend to show a much lower response. Solar forcing produces a very similar impact in magnitude over most regions at multidecadal timescales. Nevertheless, the analysis is not independent of the influence of volcanic forcing (Chapter 6). In addition, considering the STSI or ssTSI simulations at continental scales has not lead to any systematic improvement in model-data consistency.

The main differences across the regions are shown by the reconstructions, while simulations tend to show a more homogeneous response across the different regions. This is more evident in some of the PMIP3 experiments, that show very high inter-regional correlation values above 0.8 (e.g. IPSLCM5).

The larger spatial heterogenity found within the reconstructions is consistent with the larger contribution of the internal variability at these spatial scales and may be also influenced by problems in the reconstructions. Alternatively, it may be an indication that AOGCMSs/ESMs are spatially too homogeneous. Further evidence supporting this argument is found in a more detailed analysis of the

summer temperatures for the European region, where the simulated temporal and spatial variability is lower than the reconstructed one.

8

Discussion and conclusions

This Thesis evaluates the relative roles of the external forcing and the internal climate variability embedded in the temperature variability of the CE provided by AOGCM/ESM simulations and reconstructions. Such a fundamental goal involved an exhaustive compilation, analysis and comparison of paleoclimatic temperature evidences from simulations and reconstructions at various spatial scales, ranging from global to continental regions.

A summary of the specific conclusions following each part of the analysis in this work has been provided at the end of the corresponding chapter. The main results of this Thesis are highlighted (in italics) and discussed in Section 8.1 from a broader perspective. Finally, an insight into the open questions for future works is given in section 8.2.

8.1 Main conclusions

The most comprehensive paleoclimatic temperature dataset from AOGCM/ESM simulation and proxy-based reconstructions of the CE has been compiled and analysed. The corresponding set of external forcins is additionally assembled and analysed. The extensive amount of evidences of the evolution of the temperature during the CE represents an update since AR5 and guarantees an outright comparative assessment of common features as well as discrepancies between simulated and reconstructed temperatures. It additionally ensures a deep inspection of the uncertainties that prevent from straight comparisons between simulations and reconstructions like the influence of methodological issues in the reconstructions, or the use of various solar forcing scenarios for the LM, or different generations of AOGCMs.

The Total External Forcing (TEF) estimated for each simulation can be considered a good approximation of the net radiative forcing that drives the simulation. The TEF series aggregate all individual forcing factors used as boundary conditions in each experiment and thus allow for comparing the variety of external forcing configurations considered by the different models. The TEF series are largely influenced by the volcanic activity, not only at interannual scales, as expected, but also a multidecadal timescales. Volcanic forcing is identified as a main driver of the low frequency variability in LM climate. Solar forcing is subject to considerable uncertainties in the ensemble of simulations, with percentage of TSI changes from the LMM to presetn from 0.01% to 0.29%. Nevertheless, these different scenarios do not lead to a clear impact in model-data comparison nor can be clearly derived from the analysis of the available simulations and reconstructions which scenario of solar variability is more likely. Arguably, this can be considered an indication of the comparably larger influence of volcanic forcing.

It is worth to mention that the radiative forcing associated to some of the individual forcing factors, such as the anthropogenic aerosols and the land use, is not available for all simulations. In those cases a corresponding series from some available model estimates is considered as a substitute. The latter may have implications though on the accuracy of TEF estimations during the industrial period.

The hemispheric and global temperature shows linear responses to the TEF at multidecadal and longer timescales. The latter is true for both simulations and reconstructions so that the temperature can be estimated to a good approximation from a linear regression from TEF. Nonetheless, the correlations obtained between TEF and temperature response are higher for the simulations than for the reconstructions. Correlation values from the reconstructions are also positive and significant with a bias to higher values in the case of continental reconstructions compared to those including also the oceans, in agreement with simulations.

The Last Millennium Transient Climate Response (LMTCR) provides, for each simulation and reconstruction, a quantitative assessment of the linear temperature response to the external forcing over the Last Millennium. LMTCR is calculated as the regression coefficient between the temperature and the TEF. It represents therefore a temperature response to forcing ratio, that can be compared to other estimates of the sensitivity of climate to changes in forcing, such as the ECS or TCR. LMTCR values are always lower than ECS and TCR since LMTCR evaluates only the quasi-instantaneous response, in contrast to the delayed responses included also by the other climate sensitivity estimates. Differences between the

equilibrium estimates and the LMTCR could be, indeed, useful to evaluate the amount of non linear response to forcing included in the temperature signal. On average, ECS estimates are around 0.8 K/Wm^2, while the LMTCR values are close to 0.4 K/Wm^2. The latter implies that about half of the equilibrium response could be associated to non-linear processes.

Simulated and reconstructed LMTCR values furnish a quantitative framework in which the model-data consistency can be evaluated. Addittional to the qualitative comparison between the temporal evolution of the multi-model simulations and the ensemble of reconstructions, the LMTCR allows for assessing quantitatively the consistency between both sources of information. LMTCR values obtained from both simulations and reconstructions are, in general, in agreement. LMTCR values are found to be larger for continental only domains than for complete hemisphere domains in spite of the spatial uncertainties affecting reconstructions. Notwithstanding, the reconstructions show a slight bias to lower LMTCR values compared to simulations.

The consistency between simulations and reconstructions of NH temperature during the first half of the LM has been generally overestimated in the literature. The reconstruction ensemble has traditionally been built based on all available NH reconstructions, independently of the existence of various methodological issues that produce biases in the ensemble. One of this is the selection of different target domains. Indeed, reconstructions targeting, for instance, land only or land and ocean, have been treated usually as NH reconstructions, despite their calibration target showing biases to specific regions of the hemisphere. If the ensemble of NH reconstrucions is segregated taking this into account, the spread of the reconstruction ensemble disgregates during the MCA in non-overlapping bands. The latter evidences the relevance of considering methodological biases to estimate the aggregated uncertainty band and in order to compare with model outputs. Consequently, while for the traditional approach simulations are embedded in the reconstruction ensemble, the framework provided herein evidences a larger disagreement during the MCA, specially when the complete NH (land and ocean) is considered.

The internal variability may largely impact the spatial distribution of the temperatures during the MCA. Despite the robust hemispheric and global long term responses to external forcing identified in this work, internal variability may be a big player in the MCA-LIA transition. AOGCM/ESM simulations indicate a large scale pattern of response to forcing with temperature increases that are larger over continental regions and the Arctic. Individual simulations show considera-

ble variability at continental and regional scales as an indication of the influence of initial conditions and thus internal variability. While errors in climate reconstructions cannot be ruled out, if the large scale features of reconstructed climate were to be considered reliable, this would imply that models do not provide an adequate representation of mechanisms of response to external forcing or that internal variability played a major role in shaping the changes from the MCA to the LIA. The latter is supported by other model based and reconstructions studies. The spatial patterns of temperature evidenced from reconstructions for the MCA-LIA transition at global or regional scales are not well reproduced by the numerical experiments.

The consistency between simulations and reconstructions at regional scales is more limited than at hemispheric and global scales, specially for regions in the SH, due to the larger influence of the internal variability. Influence of external forcing is noticed in most of the regional reconstructions analysed, although the intensity of this influence varies depending on the region. The NH reconstructions show a more noticeable relation to external forcing and thus a larger consistency with the simulations. In contrast a lower response is found for the SH regions. Interestingly, the regional dependence of the response to forcing is lower in the simulations that are characterized by a more spatially homogeneous behaviour than reconstructions.

8.2 Outlook

The recent publication of a new ensemble of CESM simulations (see Chapter 2; Otto-Bliesner et al., 2015) offers an unique opportunity of updating, as well as analysing in more detail, some of the results shown within this Thesis. Overall, the inclusion of the CESM ensemble of simulations in future analysis appears as the natural continuation of the present work, at the time that it offers an exceptional framework in which analysing some other scientific questions not reported in this Thesis.

This ensemble includes 10 full-forcing experiments that, together with the 37 forced transient runs already analyzed herein, offers the most comprehensive collection of the high complexity LM model simulations. Such a numerous ensemble, in which only the initial conditions have been modified, allows to better assess the role of the internal variability in the temperature response. This assessment has a direct impact on a better evaluation of the uncertainties related to the model simulations, not only at temporal but also at spatial scales. In addition,

various ensembles of single-forcing simulations are also provided (Otto-Bliesner et al., 2015) which offers the possibility of further analysing the impact of each single forcing factor. Within the present work the influence in the temperature of the TEF has been widely assessed in simulations and reconstructions, but not the impact of each individual factor. Albeit the volcanic forcing has evidenced a clear influence at multidecadal and longer timescales, it is an interesting issue to quantify this effect and specially to analyse the influence of the other forcings during different periods of the LM.

This work has evidenced the larger influence of internal variability when analysing continental regions. Nevertheless, the influence of the external forcing has also been noticed, specially in some regions of the NH. A quantification of this forced response a regional scales would be a desirable objective. Indeed, a natural extension of the present work is to analyse whether exists a linear relationship between simulations and reconstructions at these spatial scales and if so to calculate the corresponding ratio of response by the LMTCR estimates. To this end it would also be useful to develop a better characterization of the regional TEF. The hemispheric TEF presented herein does not show clear improvement for characterizing the hemispheric tempertaure response, nevertheless, it should be noted that it was rather an approximation than a proper hemispheric TEF, since only the volcanic forcing was adapted to this spatial scale. A regional TEF based on the real radiative influence of each forcing factor becomes crucial for a better characterization of the externally forced response at continental scales.

Finally, the analysis of the influence of internal variability and the forced response during the CE has been focused herein solely based on the temperature response. The influence of the external forcing in other climatic variables, such as the sea level pressure and related hydroclimate variables (i.e. precipitation, drought), has been already attempted by several studies (e.g. Swingedouw et al., 2010; Bond et al., 2001) focussing on a single model. A comprehensive analysis of the response to forcing by the collection of simulations as well as the corresponding existing reconstructions constitutes an interesting challenge for future works.

References

Ammann, C. M., M. G. Genton, and B. Li, 2010: Technical note: correcting for signal attenuation from noisy proxy data in climate reconstructions. *Clim. of the Past*, **6**, 273–279.

Ammann, C. M., F. Joos, D. Schimel, B. Otto-Bliesner, and R. Tomas, 2007: Solar influence on climate during the past millennium: Results from transient simulations with the NCAR Climate System Model. *Proceedings of the National Academy of Sciences*, **104**, 3713.

Ammann, C. M., G. A. Meehl, W. M. Washington, and C. Zender, 2003: A monthly and latitudinally varying volcanic forcing dataset in simulations of 20th century climate. *Geophys. Res. Lett.*, **30**, 1657, doi:10.1029/2003GL016875.

Ammann, C. M. and E. Wahl, 2007: The importance of the geophysical context in statistical evaluations of climate reconstruction procedures. *Clim. Change*, **85**, 71–88, doi:10.1007/s10584-007-9276-x.

Baker, A., C. Bradley, S. J. Phipps, M. Fischer, I. J. Fairchild, L. Fuller, C. Spötl, and C. Azcurra, 2012: Millennial-length forward models and pseudoproxies of stalagmite δ^{18}O: an example from NW Scotland. *Clim. of the Past*, **8**, 1153–1167, doi:10.5194/cp–8–1153–2012.

Bard, E., G. Raisbeck, F. Yiou, and J. Jouzel, 2000: Solar irradiance during the last 1200 years based on cosmogenic nuclides. *Tellus B*, **52**, 985–992.

Battle, M., M. Bender, T. Sowers, P. P. Tans, J. H. Butler, J. W. Elkins, J. T. Ellis, T. Conway, N. Zhang, P. Lang, and A. D. Clarket, 1996: Atmospheric gas concentrations over the past century measured in air from firn at the South Pole. *Nature*, **383**, 231–235.

Bauer, E., M. Claussen, V. Brovkin, and A. Huenerbein, 2003: Assessing climate forcings of the Earth system for the past millennium. *Geophys. Res. Lett*, **30**,

1276, 1–4.

Bauer, E., A. Ganopolski, and M. Montoya, 2004: Simulation of the cold climate event 8200 years ago by meltwater outburst from Lake Agassiz. *Paleoceanography*, **19**, PA3014, doi: 10.1029/2004PA001030.

Berger, A. and M. F. Loutre, 1991: Insolation values for the climate of the last 10 million years. *Quaternary Science Reviews*, **10**, 297–317.

Berger, A. L., 1978: Long-term variations of daily insolation and quaternary climatic changes. *J. Atmos. Sci.*, **35**, 2362–2367.

Bloomfield, P., 1976: Fourier analysis of time series: an introduction. *New York: John Wiley and Sons.*.

Blunier, T., J. Chappellaz, J. Schwander, B. Stauffer, and D. Raynaud, 1995: Variations in atmospheric methane concentration during the Holocene epoch. *Nature*, **374**, 46–49.

Bond, G., B. Kromer, J. Beer, R. Muscheler, M. N. Evans, W. Showers, S. Hoffmann, R. Lotti-Bond, I. Hajdas, and G. Bonani, 2001: Persistent solar influence on north atlantic climate during the holocene. *Science*, **294**, 2130–2136.

Bothe, O., J. Jungclaus, D. Zanchettin, and E. Zorita, 2013a: Climate of the last millennium: ensemble consistency of simulations and reconstructions. *Clim. of the Past*, **9**, 1089–1110.

Bothe, O., J. H. Jungclaus, and D. Zanchettin, 2013b: Consistency of the multimodel CMIP5/PMIP3-past1000 ensemble. *Clim. of the Past*, **9**, 2471–2487.

Boucher, O. and M. Pham, 2002: History of sulfate aerosol radiative forcings. *Geophys. Res. Lett.*, **29**, 1308.

Braconnot, P., S. P. Harrison, M. Kageyama, P. J. Bartlein, V. Masson-Delmotte, A. abe Ouchi, B. Otto-Bliesner, and Y. Zhao, 2012: Evaluation of climate models using paleoclimate data. *Nature Climate Change*, **2**, 417–424, doi: 10.1038/nclimate1456.

Bradley, R. S. and P. D. Jones, 1993: 'Little Ice Age' summer temperature variations: their nature and relevance to recent global warming trends. *The Holocene*, **3**, 367–376.

Bretagnon, P. and G. Francou, 1988: Planetary theories in rectangular and spherical variables - VSOP87 solutions. *Astron. Astrophys.*, **202**, 309–315.

Briffa, K., P. Jones, F. Schweingruber, and T. Osborn, 1998: Influence of volcanic eruptions n Northern Hemisphere summer temperature over the past 600 years. *Nature*, **393**, 450–455.

Briffa, K., T. Osborn, F. Schweingruber, I. Harris, P. Jones, S. Shiyatov, and E. Vaganov, 2001: Low-frequency temperature variations from a northern tree ring density network. *J. Geophys. Res.*, **106**, 2929–2941.

Briffa, K. R., 2000: Annual climate variability in the holocene: interpreting the message of ancient trees. *Quaternary Science Reviews*, **19**, 87–105.

Briffa, K. R., T. S. Bartholin, D. Eckstein, P. D. Jones, W. Karlén, F. H. Schweingruber, and P. Zetterberg, 1990: A 1,400-year tree-ring record of summer temperatures in Fennoscandia. *Nature*, **346**, 434–439.

Brohan, P., J. J. Kennedy, I. Harris, S. F. B. Tett, and P. D. Jones, 2006: Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850. *J. Geophys. Res.*, **111**, D12106, doi:10.1029/2005JD006548.

Buerger, G. and U. Cubasch, 2005: Are multiproxy climate reconstructions robust? *Geophys. Res. Lett.*, **32**, L23711, doi:10.1029/2005GL024155.

Buerger, G., I. Fast, and U. Cubasch, 2006: Climate reconstruction by regression - 32 variations on a theme. *Tellus*, **58A**, 227–235.

Byrne, M. and P. O'Gorman, 2013: Land-Ocean Warming Contrast over a Wide Range of Climates: Convective Quasi-Equilibrium Theory and Idealized Simulations. *J. Climatol.*, **26**, 4000–4016.

Cane, M., M. Khodri, P. Braconnot, S. Joussaume, M. Kageyama, D. Paillard, A. Clement, H. Gildor, S. Tett, and E. Zorita, 2006: Progress in Paleoclimate Modeling. *J. Climate*, **19**, 5031–5057.

Christiansen, B., 2011: Reconstructing the NH mean temperature: can underestimation of trends and variability be avoided? *J. Climate*, **24**, 674–692, doi: 10.1175/2010JCLI3646.1.

— 2012: Reply to Comments on 'Reconstructing the NH mean temperature:Can underestimation of trends and variability be avoided?' by M. Tingley and B. Li. *J. Climate*, **25**, 3447–3452, doi: 10.1175/JCLI–D–11–00162.1.

Christiansen, B. and F. C. Ljungqvist, 2011: Reconstruction of the Extra-tropical NH mean temperature over the last millennium with a method that preserves low-frequency variability. *J. Climate*, **24**, 6013–6034, doi: 10.1175/2011JCLI4145.1.

— 2012a: Reply to Comments on ' Reconstruction of the Extratropical NH mean temperature over the last millennium with a method that preserves low-frequency variability ' by A. Moberg. *J. Climate*, **25**, 7998–8003, doi: 10.1175/JCLI–D–11–00642.1.

— 2012b: The extra-tropical Northern Hemisphere temperature in the last two millennia: reconstructions of low-frequency variability. *Clim. of the Past*, **8**, 765–786, doi: 10.5194/cp–8–765–2012.

Christiansen, B., T. Schmith, and P. Thejll, 2009: A surrogate ensemble study of climate reconstruction methods: stochasticity and robustness. *J. Climate*, **22**, 951–976 doi: 10.1175/2008JCLI2301.1.

Collins, M., S. An, W. Cai, A. Ganachaud, E. Guilyardi, F. Jin, M. Jochum, M. Lengaigne, S. Power, A. Timmermann, G. Vecchi, and A. Wittenberg, 2010: The impact of global warming on the tropical Pacific Ocean and El Niño. *Nature Geoscience*, **3**, 391–397.

Collins, M., R. Knutti, J. Arblaster, J.-L. Dufresne, T. Fichefet, P. Friedlingstein, X. Gao, W. Gutowski, T. Johns, G. Krinner, et al., 2013: Long-term climate change: projections, commitments and irreversibility. Climate Change 2013: The Physical Science Basis. *Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate change*.

Crowley, T., 2000: Causes of climate change over the past 1000 years. *Science*, **289**, 270–277.

Crowley, T., G. Zielinski, B. Vinther, R. Udisti, K. Kreutz, J. Cole-Dai, and E. Castellano, 2008: Volcanism and the Little Ice Age. *PAGES Newsletter*, **16**, 22–23.

Crowley, T. J., S. K. Baum, K.-Y. Kim, G. C. Hegerl, and W. T. Hyde, 2003: Modeling ocean heat content changes during the last millennium. *Geophys. Res. Lett.*, **30**, 1932, doi: 10.1029/2003GL017801.

Crowley, T. J. and M. B. Unterman, 2012: Technical details concerning development of a 1200-yr proxy index for global volcanism. *Earth Syst. Sci. Data Discuss*, **5**, 1–18, doi:10.5194/essdd–5–1–2012.

— 2013: Technical details concerning development of a 1200 yr proxy index for global volcanism. *Earth System Science Data*, **5**, 187–197.

D'Arrigo, R., R. Wilson, and G. Jacoby, 2006: On the long-term context for late twentieth century warming. *J. Geophys. Res.*, **111**, D03103, doi:10.1029/2005JD006352.

Delworth, T. and F. Zeng, 2012: Multicentennial variability of the Atlantic Meridional Overturning Circulation and its climatic influence in a 4000 year simulation of the GFDL CM2.1 climate model. *Geophys. Res. Lett.*, **39**, L13702, 6 pp., doi: 10.1029/2012GL052107.

Diaz, H., R. Trigo, M. Hughes, M. Mann, E. Xoplaki, and D. Barriopedro, 2011: Spatial and temporal characteristics of climate in Medieval Times revisited. *Bull. of the Amer. Met. Soc.*, **92**, 1487, doi:10.1175/BAMS–D–10–05003.1.

Dufresne, J.-L., M.-A. Foujols, S. Denvil, A. Caubel, O. Marti, O. Aumont, Y. Balkanski, S. Bekki, H. Bellenger, R. Benshila, et al., 2013: Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. *Clim. Dyn.*, **40**, 2123–2165.

Emile-Geay, J. and M. Tingley, 2014: Inferring climate variability from nonlinear proxies. application to paleo-enso studies. *Paleoceanography*.

Esper, J., E. Cook, and F. Schweingruber, 2002: Low-frequency signals in long tree-ring chronologies for reconstructing past temperature variability. *Science*, **295**, 2250–2253.

Esper, J. and D. Frank, 2009: The IPCC on a heterogeneous Medieval warm period. *Climatic Change*, **94**, 267–273.

Esper, J., D. C. Frank, R. J. Wilson, and K. R. Briffa, 2005: Effect of scaling and regression on reconstructed temperature amplitude for the past millennium. *Geophys. Res. Lett.*, **32**, 7, doi: 10.1029/2004GL021236.

Etheridge, D., L. Steele, R. Langenfelds, R. Francey, J. Barnola, and V. Morgan, 1996: Natural and anthropogenic changes in atmospheric CO2 over the last 1000 years from air in Antarctic ice and firn. *J. Geophys. Res.*, **101**, 4115–4128.

Etheridge, D. M., L. P. Steele, R. J. Francey, and R. L. Langenfelds, 1998: Atmospheric methane between 1000 A.D. and present: Evidence of anthropogenic emissions and climatic variability. *J. Geophys. Res.*, **103**, 15979–15993.

EuroMed2k Consortium, 2015: European summer temperatures since Roman times. *Environmental Research Letters*, **under review**.

Evans, M. N., B. K. Reichert, A. Kaplan, K. J. Anchukaitis, E. A. Vaganov, M. K. Hughes, and M. A. Cane, 2006: A forward modeling approach to paleoclimatic interpretation of tree-ring data. *Journal of Geophysical Research - Biogeosciences*, **111**, 148–227.

Fernández-Donado, L., J. González-Rouco, C. Raible, C. M. Ammann, D. Barriopedro, E. García-Bustamante, J. H. Jungclaus, S. J. Lorenz, J. Luterbacher, S. J. Phipps, J. Servonnat, D. Swingedouw, S. F. B. Tett, S. Wagner, P. Yiou, and E. Zorita, 2013: Large-scale temperature response to external forcing in simulations and reconstructions of the last millennium. *Clim. of the Past*, **9**, 393–421, doi:10.5194/cp–9–393–2013.

Feulner, G., 2011: Are the most recent estimates for Maunder Minimum solar irradiance in agreement with temperature reconstructions? *Geophys. Res. Lett.*, **38**, L16706, doi:10.1029/2011GL048529.

Fischer, E., J. Luterbacher, E. Zorita, S. Tett, C. Casty, and H. Wanner, 2007: European climate response to tropical volcanic eruptions over the last half millennium. *Geophys. Res. Lett.*, **34**, L05707, doi:10.1029/2006GL027992.

Flato, G., J. Marotzke, B. Abiodun, P. Braconnot, S. C. Chou, W. Collins, P. Cox, F. Driouech, S. Emori, V. Eyring, et al., 2013: Evaluation of climate models. Climate Change 2013: The Physical Science Basis. *Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, 741–866.

Flückiger, J., E. Monnin, B. Stauffer, J. Schwander, T. f. Stocker, J. Chappellaz, D. Raynaud, and J. M. Barnola, 2002: High resolution Holocene N_2O ice core record and its relationship with CH_4 and CO_2. *Global Biogeochem. Cy.*, **16**, 1010, doi:10.1029/2001GB001417.

Folland, C., T. Karl, J. Christy, R. Clarke, G. Gruza, J. Jouzel, M. Mann, J. Oerlemans, M. Salinger, and S. Wang, 2001: Observed climate variability and change. *Climate Change 2001: The Scientific Basis. Contribution of Working*

Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change, University PressClimate change, 101–181.

Folland, C. K., T. Karl, and K. Y. Vinnikov, 1990: Observed climate variations and change. *Scientific Assessment of Climate Change. Contribution of Working Group I to the First Assessment Report of the Intergovernmental Panel on Climate Change, University Press*, **7**, 195–238.

Forster, P., V. Ramaswamy, P. Artaxo, T. Berntsen, R. Betts, D. Fahey, J. Haywood, J. Lean, D. Lowe, G. Myhre, J. Nganga, R. Prinn, G. Raga, M. Schulz, and R. Van Dorland, 2007: Changes in atmospheric constituents and in radiative forcing. *Climate Change 2007: The Physical Science Basis.Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, **30**, 129–234.

Foukal, P., G. North, and T. Wigley, 2004: CLIMATE: A Stellar View on Solar Variations and Climate. *Science*, **306**, 68–69.

Frank, D., J. Esper, and E. Cook, 2007: Adjustment for proxy number and coherence in a large-scale temperature reconstruction. *Geophys. Res. Lett.*, **34**, 16709, doi:10.1029/2007GL030571.

Frank, D., J. Esper, E. Zorita, and R. Wilson, 2010a: A noodle, hockey stick, and spaghetti plate: a perspective on high-resolution paleoclimatology. *Wiley Interdisciplinary Reviews: Climate Change*, **1**, 507–516.

Frank, D. C., J. Esper, C. C. Raible, U. Büntgen, V. Trouet, B. Stocker, and F. Joos, 2010b: Ensemble reconstruction constraints on the global carbon cycle sensitivity to climate. *Nature*, **463**, 527–530, doi:10.1038/nature08769.

Gagen, M., E. Zorita, D. McCarroll, G. H. Young, H. Grudd, R. Jalkanen, N. J. Loader, I. Robertson, and A. Kirchhefer, 2011: Cloud response to summer temperatures in Fennoscandia over the last thousand years. *Geophys. Res. Lett.*, **38**.

Gao, C. C., A. Robock, and C. M. Ammann, 2008: Volcanic forcing of climate over the past 1500 years: an improved ice core-based index for climate models. *J. Geophys. Res.*, **113**, D23111, doi: 10.1029/2008JD010239.

Goldewijk, K. K., 2001: Estimating global land use change over the past 300 years: The HYDE database. *Global Biogeochem Cycles*, **15**, 417–433.

Goldewijk, K. K., A. Beusen, G. van drecht, and M. de Vos, 2011: The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12000 years. *Global Ecol. Biogeogr.*, **20**, 73–86, doi: 10.1111/j.1466–8238.2010.00587.x.

González-Rouco, F., L. Fernández-Donado, C. Raible, D. Barriopedro, J. Luterbacher, J. Jungclaus, D. Swingedouw, J. Servonnat, E. Zorita, S. Wagner, and C. M. Ammann, 2011: Medieval Climate Anomaly to Little Ice Age transition

as simulated by current climate models. *Medieval Climate Anomaly. Pages News*, **19**, 7–8.

González-Rouco, F., H. Von Storch, and E. Zorita, 2003a: Deep soil temperature as proxy for surface air-temperature in a coupled model simulation of the last thousand years. *Geophys. Res. Lett.*, **30**, 2116, doi:10.1029/2003GL018264.

González-Rouco, J., H. Beltrami, E. Zorita, and M. Stevens, 2009: Borehole climatology: a discussion based on contributions from climate modeling, Clim. *Clim. of the Past*, **5**, 97–127.

González-Rouco, J., H. Beltrami, E. Zorita, and H. Von Storch, 2006: Simulation and inversion of borehole temperature profiles in surrogate climates: Spatial distribution and surface coupling. *Geophys. Res. Lett.*, **33**, L01703, doi:10.1029/2005GL024693.

González-Rouco, J., E. Zorita, U. Cubasch, H. von Storch, I. Fisher-Bruns, F. Valero, J. Montavez, U. Schlese, and S. Legutke, 2003b: Simulating the climate since 1000 AD with the AOGCM ECHO-G. *In: Solar variability as an input to the Earth's environment. International Solar Cycle Studies (ISCS) Symposium, 23-28 June 2003, Tatranská Lomnica, Slovak Republic. Ed.: A. Wilson. ESA SP-535, Noordwijk: ESA Publications Division, ISBN 92*, volume 9092, 329–338.

Goosse, H., M. Braida, X. Crosta, A. Mairesse, V. Masson-Delmotte, P. Mathiot, R. Neukom, H. Oerter, G. Philippon, H. Renssen, et al., 2012a: Antarctic temperature changes during the last millennium: evaluation of simulations and reconstructions. *Quaternary Science Reviews*, **55**, 75–90.

Goosse, H., E. Crespin, S. Dubinkina, M. F. Loutre, M. E. Mann, H. Renssen, Y. Sallaz-Damaz, and D. Shindell, 2012b: The role of forcing and internal dynamics in explaining the 'Medieval Climate Anomaly'. *Clim. Dyn.*, **in press**, doi:10.1007/s00382–012–1297–0.

Goosse, H., T. Crowley, E. Zorita, C. M. Ammann, H. Renssen, N. Riedwyl, A. Timmermann, E. Xoplaki, and H. Wanner, 2005: Modelling the climate of the last millennium: What causes the differences between simulations. *Geophys. Res. Lett.*, **32**, L06710, doi:10.1029/2005GL022368.

Goosse, H., J. Guiot, M. E. Mann, S. Dubinkina, and Y. Sallaz-Damaz, 2012c: The medieval climate anomaly in Europe: Comparison of the summer and annual mean signals in two reconstructions and in simulations with data assimilatio. *Global and Planetary Change*, **84-85**, 35–47, doi:10.1016/j.gloplacha.2011.07.002.

Goosse, H., H. Renssen, A. Timmermann, R. Bradley, and M. Mann, 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.

Graham, N., C. M. Ammann, D. Fleitmann, K. M. Cobb, and J. Luterbacher, 2011: Support for global climate reorganization during the Medieval Climate Anomaly. *Clim. Dyn.*, **37**, 1217–1245.

Gray, J. L., J. beer, M. Geller, J. D. Haigh, M. Lockwood, K. Matthes, U. Cubasch, D. Fleitmann, G. Harrison, L. Hood, J. Luterbacher, G. A. Meehl, D. shindell, B. V. Geel, and W. White, 2010: Solar influence on climate. *Rev. Geophys.*, **48, RG4001**, doi:10.1029/2009RG000282.

Groveman, B. S. and H. E. Landsberg, 1979: Simulated northern hemisphere temperature departures 1579-1880. *Geophys. Res. Lett.*, **6**, 767–769.

Guillot, D., B. Rajaratnam, J. Emile-Geay, et al., 2015: Statistical paleoclimate reconstructions via Markov random fields. *The Annals of Applied Statistics*, **9**, 324–352.

Haigh, J. D., A. R. Winning, R. Toumi, and J. W. Harder, 2010: An influence of solar spectral variations on radiative forcing of climate. *Nature*, **467**, 696–699.

Hansen, J. and M. Sato, 2004: Greenhouse gas growth rates. *Proceed. Nat. Acad. of Sci.*, **101**, 16109–16114.

Hansen, J., M. Sato, L. Nazarenko, R. Ruedy, A. Lacis, D. Koch, I. Tegen, T. Hall, D. Shindell, B. Santer, P. Stone, T. Novakov, L. Thomason, R. Wang, Y. Wang, D. Jacob, S. Hollandsworth, L. Bishop, J. Logan, A. Thompson, R. Stolarski, J. Lean, R. Willson, S. Levitus, J. Antonov, N. Rayner, D. Parker, and J. Christy, 2002: Climate forcings in Goddard Institute for Space Studies SI2000 simulations. *J. Geophys. Res.*, **107**.

Hartmann, D., A. Klein Tank, M. Rusicucci, L. Alexander, B. Broenniman, Y. Charabi, F. Dentener, E. Dlugokencky, D. Easterling, A. Kaplan, et al., 2013: Observations: atmosphere and surface. Climate Change 2013: The Physical Science Basis. *Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*.

Hegerl, G., T. Crowley, M. Allen, W. Hyde, H. Pollack, J. Smerdon, and E. Zorita, 2007a: Detection of human influence on a new, validated 1500-year temperature reconstruction. *J. Climate*, **20**, 650–666.

Hegerl, G., T. Crowley, W. Hyde, and D. Frame, 2006: Climate sensitivity constrained by temperature reconstructions over the past seven centuries. *Nature*, **440**, 1029–1032.

Hegerl, G. C., J. Luterbacher, J. F. González-Rouco, S. Tett, T. Crowley, and E. Xoplaki, 2011: Influence of human and natural forcing on European seasonal temperatures. *Nature Geoscience*, **4**, 99–103, doi: 10.1038 / NGEO1057.

Hegerl, G. C., F. W. Zwiers, N. P. Braconnot, P. Gillett, Y. Luo, N. Orsini, J. A. M.and Nicholls, J. E. Penner, and P. A. Stott, 2007b: Understanding and attributing climate change. *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..

Hind, A., A. Moberg, and R. Sundberg, 2012: Statistical framework for evaluation of climate model simulations by use of climate proxy data from the last millennium–Part 2: A pseudo-proxy study addressing the amplitude of solar forcing. *Clim. of the Past*, **8**, 1355–1365.

Hofer, D., C. Raible, and T. Stocker, 2011: Variations of the Atlantic meridional overturning circulation in control and transient simulations of the last millennium. *Clim. of the Past*, **7**, 133–150, doi:10.5194/cp–7–133–2011.

Houghton, J., 2005: Global warming. *Rep. Progress Physics*, **68**, 1343–1403.

Houghton, J., Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, and C. A. Johnson, 2001: *Climate Change, 2001: The scientific basis. Contribution of Working Group I to Third Assessment Report of the Intergovernmental Panel on Climate Change.*. Cambridge University Press.

Hoyt, D. V. and K. H. Schatten, 1993: A discussion of plausible solar irradiance variations 1700-1992. *J. Geophys. Res.*, **98**, 18895–18906, doi:10.1029/93JA01944.

Huang, S., 2004: Merging information from different resources for new insights into climate change in the past and future. *Geophys. Res. Lett.*, **31**, L13205, doi:10.1029/2004GL019781.

Huang, S., H. Pollack, and P. Shen, 2000: Temperature trends over the past five centuries reconstructed from borehole temperatures. *Nature*, **403**, 756–758.

Hurrell, J. W., M. M. Holland, P. R. Gent, S. Ghan, J. E. Kay, P. Kushner, J.-F. Lamarque, W. G. Large, D. Lawrence, K. Lindsay, et al., 2013: The Community Earth System Model: a framework for collaborative research. *Bulletin of the American Meteorological Society*, **94**, 1339–1360.

Hurtt, G., L. P. Chini, S. Frolking, R. Betts, J. Feddema, G. Fischer, J. Fisk, K. Hibbard, R. Houghton, A. Janetos, et al., 2011: Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. *Climatic Change*, **109**, 117–161.

Jacoby, G. C. and R. D'Arrigo, 1989: Reconstructed Northern Hemisphere annual temperature since 1671 based on high-latitude tree-ring data from North America. *Climatic Change*, **14**, 39–59.

Jansen, E., J. Overpeck, K. R. Briffa, J. C. Duplessy, V. Masson-Delmontte, D. Olago, B. Otto-Bliesner, W. R. Peltier, S. Rahmstorf, R. Ramesh, D. Raynaud, D. Rind, O. Solomina, R. Villalba, and D. Zhang, 2007: Paleoclimate: Climate Change 2007: The Physical Science Basis. *Contribution of Working*

Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, University Press.

Johns, T. C., J. M. Gregory, W. J. Ingram, C. E. Johnson, A. Jones, J. A. Lowe, J. F. B. Mitchell, D. L. Roberts, D. M. H. Sexton, D. S. Stevenson, S. F. B. Tett, and M. J. Woodage, 2003: Anthropogenic climate change for 1860 to 2100 simulated with the HadCM3 model under updated emissions scenarios. *Clim. Dyn.*, **20**, 583–612.

Jones, J. M. and M. Widmann, 2003: Instrument- and tree-ring-based estimates for the Antarctic Oscillation. *J. Climate*, **16**, 3511–3524.

Jones, P., K. Briffa, T. Barnett, and S. Tett, 1998: High-resolution palaeoclimatic records for the last millennium: interpretation, integration and comparison with General Circulation Model control-run temperatures. *The Holocene*, **8**, 455–471.

Jones, P. and M. Mann, 2004: Climate over past millennia. *Rev. Geophys.*, **42**, 1–42.

Jones, P., T. Osborn, and K. Briffa, 2001: The evolution of climate over the last millennium. *Science*, **292**, 662–667.

Jones, P. D., K. R. Briffa, T. J. Osborn, J. M. Lough, T. D. van Ommen, B. M. Vinther, J. Luterbacher, E. R. Wahl, F. W. Zwiers, M. E. Mann, G. A. Schmidt, C. M. Ammann, B. M. Buckley, K. M. Cobb, J. Esper, H. Goosse, N. Graham, E. Jansen, T. Kiefer, C. Kull, M. Kttel, E. Mosley-Thompson, J. T. Overpeck, N. Riedwyl, M. Schulz, A. W. Tudhope, R. Villalba, H. Wanner, E. Wolff, and E. Xoplaki, 2009: High-resolution palaeoclimatology of the last millennium: a review of current status and future prospects. *The Holocene*, **19**, 3–49.

Joos, F., I. Prentice, S. Sitch, R. Meyer, G. Hooss, G. Plattner, S. Gerber, and K. Hasselmann, 2001: Global warming feedbacks on terrestrial carbon uptake under the Intergovernmental Panel on Climate Change (IPCC) emission scenarios. *Global Biogeochemical Cycles*, **15**, 891–908.

Joos, F. and R. Spahni, 2008: Rates of change in natural and anthropogenic radiative forcing over the past 20000 years. *Proceed. Nat. Acad. of Sci.*, **105**, 1425–1430.

Juckes, M., M. Allen, K. Briffa, J. Esper, G. Hegerl, A. Moberg, T. Osborn, and S. Weber, 2007: Millennial temperature reconstruction intercomparison and evaluation. *Clim. of the Past*, **3**, 591–609.

Jungclaus, J., S. Lorenz, C. Timmreck, C. Reick, V. Brovkin, K. Six, J. Segschneider, M. Giorgetta, T. Crowley, J. Pongratz, et al., 2010: Climate and carbon-cycle variability over the last millennium. *Clim. of the Past*, **6**, 723–737.

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

Kaplan, J. O., K. M. Krumhardt, E. C. Ellis, W. F. Ruddiman, C. Lemmen, and K. K. Goldewijk, 2011: Holocene carbon emissions as a result of anthropogenic land cover change. *The Holocene*, **21**, 775–791, doi: 10.1177/0959683610386983.

Kaufman, D., K. PAGES 2k Consortium: Anchukaitis, U. Bentgen, J. Emile-Geay, M. N. Evans, H. Goosse, J. Luterbacher, J. S. Smerdon, M. Tingley, L. von Gunten, et al., 2014: A Community-Driven Framework for Climate Reconstructions. *EOS, Transactions American Geophysical Union*, **95**, 361–362.

Kaufman, D. P., D. S. amd Schneider, N. P. McKay, C. M. Ammann, R. S. Bradley, K. R. Briffa, G. H. Miller, B. L. Otto-Bliesner, J. T. Overpeck, B. M. Vinther, and A. L. 2k Project Members, 2009: Recent warming reverses long-term arctic cooling. *Science*, **325**, 1236–1239.

Kirtman, B., S. Power, J. Adedoyin, G. Boer, R. Bojariu, I. Camilloni, F. Doblas-Reyes, A. Fiore, M. Kimoto, G. Meehl, et al., 2013: Near-term climate change: projections and predictability. climate change 2013: The physical science basis. *Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate change*, 953–1028.

Knutti, R., F. Joos, S. A. Müller, G. K. Plattner, and T. F. Stocker, 2005: Probabilistic climate change projections for co_2 stabilization profiles. *Geophys. Res. Lett.*, **32**, L20707, doi:10.1029/2005GL023294.

Krivova, N., L. Balmaceda, and S. Solanki, 2007: Reconstruction of solar total irradiance since 1700 from the surface magnetic flux. *Astronomy and Astrophysics*, **467**, 335–346.

Krivova, N. A. and S. K. Solanki, 2008: Models of solar irradiance variations: Current status. *J. Astrophys. Astr.*, **29**, 151–158.

Lâiné, A., M. Kageyama, P. Braconnot, and R. Alkama, 2009: Impact of Greenhouse Gas Concentration Changes on Surface Energetics in IPSL-CM4: Regional Warming Patterns, Land-xsSea Warming Ratios, and Glacial-Interglacial Differences. *J. Climatol.*, **22**, 4621–4635, doi: 10.1175/2009/JCLI2771.1.

Lamarque, J.-F., T. C. Bond, V. Eyring, C. Granier, A. Heil, Z. Klimont, D. Lee, C. Liousse, A. Mieville, B. Owen, et al., 2010: Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application. *Atmospheric Chemistry and Physics*, **10**, 7017–7039.

Lamb, H. H., 1965: The early medieval warm epoch and its sequel. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **1**, 13–37.

Landrum, L., B. L. Otto-Bliesner, E. R. Wahl, A. Conley, P. J. Lawrence, N. Rosenbloom, and H. Teng, 2013: Last millennium climate and its variability in ccsm4. *J. Climatol.*, **26**, 1085–111, doi: 10.1175/JCLI–D–11–00326.1.

Laskar, J., P. Robutel, F. Joutel, M. Gastineauu, A. C. M. Correia, and B. Levrard, 2004: A long term numerical solution for the insolation quantities of the earth. *Astronomy and Astrophysics*, **428**, 261–285, 10.1051/0004–6361:20041335.

Lawrimore, J. H., M. J. Menne, b. E. Gleason, C. N. Williams, D. B. Wuertz, R. S. Vose, and J. Rennie, 2011: An overview of the global historical climatology network monthly mean temperature data set, version 3. *J. Geophys. Res.*, **116**, D19121,doi:10.1029/2011JD016187.

Lean, J., J. Beer, and R. Bradley, 1995: Reconstruction of solar irradiance since 1610: Implications for climate change. *Geophys. Res. Lett.*, **22**, 3195–3198.

Lean, J., Y. Wang, and N. Sheeley Jr, 2002: The effect of increasing solar activity on the Sun's total and open magnetic flux during multiple cycles-Implications for solar forcing of climate. *Geophys. Res. Lett.*, **29**, 2224, doi:10.1029/2002GL015880.

Leclercq, P. and J. Oerlemans, 2012: Global and hemispheric temperature reconstruction from glacier length fluctuations. *Clim. Dyn.*, **38**, 1065–1079, doi: 10.1007/s00382–011–1145–7.

Lee, T. C., F. W. Zwiers, and M. Tsao, 2008: Evaluation of proxy-based millennial reconstruction methods. *Clim. Dyn.*, **31**, 263–281, doi:10.1007/s00382–007–0351–9.

Lefohn, A. S., J. D. Husar, and R. B. Husar, 1999: Estimating historical anthropogenic global sulfur emission patterns for the period 1850–1990. *Atmospheric Environment*, **33**, 3435–3444.

Lehner, F., F. Joos, C. C. Raible, J. Mignot, A. Born, K. M. Keller, and T. F. Stocker, 2015: Climate and carbon cycle dynamics in a cesm simulation from 850–2100 ce. *Earth System Dynamics Discussions*, **6**, 351–406.

Lehner, F., C. Raible, and T. Stocker, 2012: Testing the robustness of a precipitation proxy-based north atlantic oscillation reconstruction. *Quaternary Science Reviews*, **45**, 85–94.

Li, B., D. W. Nychka, and C. M. Ammann, 2010: The value of multiproxy reconstruction of past climate. *J. Amer. Stat. Assoc.*, **105**, 883–895,doi:10.1198/jasa.2010.ap09379.

Li, B. and J. Smerdon, 2012: Defining spatial comparison metrics for evaluation of paleoclimatic field reconstructions of the common era. *Environmetrics*, doi: 10.1002/env.2142.

Linán, I. D., U. Büntgen, F. González-Rouco, E. Zorita, J. Montávez, J. Gómez-Navarro, M. Brunet, I. Heinrich, G. Helle, and E. Gutiérrez, 2012: Estimating 750 years of temperature variations and uncertainties in the Pyrenees by tree ring reconstructions and climate simulations. *Clim. of the Past*, **8**, 919–933.

Ljungqvist, F., 2010: A new reconstruction of temperature variability in the extra-tropical northern hemisphere during the last two millennia. *Geografiska Annaler: Series A, Physical Geography*, **92**, 339–351.

Ljungqvist, F. C., P. J. Krusic, G. Brattstroem, and H. S. Sundqvist, 2012: Norhtern hemisphere temperature patterns in the last 12 centuries. *Clim. of the Past*, **8**, 227–249, doi:10.5194/cp–8–227–2012.

Loehle, C., 2007: A 2000-year global temperature reconstruction based on non-treering proxies. *Energy and Environment*, **18**, 1049–1058.

Loehle, C. and J. McCulloch, 2008: Correction to: A 2000-year global temperature reconstruction based on non-tree ring proxies. *Energy & Environment*, **19**, 93–100.

Lorenz, E., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141.

Luterbacher, J., D. Dietrich, E. Xoplaki, M. Grosjean, and H. Wanner, 2004: European seasonal and annual temperature variability, trends, and extremes since 1500. *Science*, **303**, 1499–1503.

Luterbacher, J., E. Xoplaki, D. Dietrich, R. Rickli, J. Jacobeit, C. Beck, D. Gyalistras, C. Schmutz, and H. Wanner, 2002: Reconstruction of sea level pressure fields over the Eastern North Atlantic and Europe back to 1500. *Clim. Dyn.*, **18**, 545–561.

MacFarling Meure, C., D. Etheridge, C. Trudinger, P. Steele, R. Langenfelds, T. Van Ommen, A. Smith, and J. Elkins, 2006: Law Dome CO2, CH4 and N2O ice core records extended to 2000 years BP. *Geophys. Res. Lett.*, **33**, L14810, doi:10.1029/2006GL026152.

Mann, M. E., 2007: Climate over the past two millennia. *Annual Review of Earth and Planetary Sciences*, **35**, 111–136.

Mann, M. E., R. S. Bradley, and M. K. Hughes, 1999: Northern hemisphere temperatures during the past millennium: Inferences, uncertainties, and limitations. *Geophys. Res. Lett.*, **26**, 759–762.

Mann, M. E. and M. K. Hughes, 2002: Tree-ring chronologies and climate variability. *Science*, **296**, 848–849.

Mann, M. E. and P. D. Jones, 2003: Global surface temperatures over the past two millennia. *Geophys. Res. Lett.*, **30**, 1820, 1–4.

Mann, M. E. and S. Rutherford, 2002: Climate reconstruction using pseudoproxies. *Geophys. Res. Lett.*, **29**, 139–1.

Mann, M. E., S. Rutherford, E. Wahl, and C. M. Ammann, 2005: Testing the fidelity of methods used in proxy-based reconstructions of past climate. *J. Climate*, **18**, 4097–4107.

Mann, M. E. and G. A. Schmidt, 2003: Ground vs. surface air temperature trends: Implications for borehole surface temperature reconstructions. *Geophys. Res. Lett.*, **30**.

Mann, M. E., Z. Zhang, M. K. Hughes, R. S. Bradley, S. K. Miller, S. Rutherford, and F. Ni, 2008: Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia. *Proceed. Nat. Acad. of Sci.*, **105**, 13252–13257.

Mann, M. E., Z. Zhang, S. Rutherford, R. S. Bradley, M. K. Hughes, D. Shindell, C. Ammann, G. Faluvegi, and F. Ni, 2009: Global signatures and dynamical origins of the little ice age and medieval climate anomaly. *Science*, **326**, 1256–1260.

Marland, G., T. Boden, and R. Andres, 2003: Global, regional, and national emissions. *Trends: a compendium of data on global change. Carbon Dioxide Information Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, TN*.

Masson-Delmotte, V., M. Schulz, A. Abe-Ouchi, J. Beer, A. Ganopolski, J. González Rouco, E. Jansen, K. Lambeck, J. Luterbacher, T. Naish, T. Osborn, B. Otto-Bliesner, T. Quinn, R. Ramesh, M. Rojas, X. Shao, and A. Timmermann, 2013: Information from Paleoclimate Archives. Climate Change 2013: The Physical Science Basis. *Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*.

McIntyre, S. and R. McKitrick, 2003: Corrections to the Mann et. al.(1998) proxy data base and northern hemispheric average temperature series. *Energy & Environment*, **14**, 751–771.

— 2005: Hockey sticks, principal components, and spurious significance. *Geophys. Res. Lett.*, **32**.

McKay, N. P. and D. S. Kaufman, 2014: An extended Arctic proxy temperature database for the past 2,000 years. *Scientific data*, **1**.

Meehl, G. A., J. M. Arblaster, K. Matthes, F. Sassi, and H. van Loon, 2009: Amplifying the pacific climate system response to a small 11-year solar cycle forcing. *Science*, **325**, 1114–1118.

Moberg, A., 2013: Comments on "Reconstruction of the extra-tropical NH mean temperture over the last millennium with a method that preserves low-frequency variability". *J. Climate*, **25**, 7991–7997, doi: 10.1175/JCLI–D–11–00404.1.

Moberg, A., D. Sonechkin, K. Holmgren, N. Datsenko, and W. Karlén, 2005: Highly variable Northern Hemisphere temperatures reconstructed from low- and high-resolution proxy data. *Nature*, **433**, 613–617.

Myhre, G., E. Highwood, K. Shine, and F. Stordal, 1998: New estimates of radiative forcing due to well mixed greenhouse gases. *Geophys. Res. Lett.*, **25**, 2715–2718.

Myhre, G., D. Shindell, F. Bréon, W. Collins, J. Fuglestvedt, J. Huang, D. Koch, J. Lamarque, D. Lee, B. Mendoza, et al., 2013: Anthropogenic and natural

radiative forcing. climate change 2013: The physical science basis. *Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, 658–740.

Neukom, R., J. Gergis, D. J. Karoly, H. Wanner, M. Curran, J. Elbert, F. González-Rouco, B. K. Linsley, A. D. Moy, I. Mundo, C. C. Raible, E. J. Steig, T. van Ommen, T. Vance, R. Villalba, J. Zinke, and D. Frank, 2014: Inter-hemispheric temperature variability over the past millennium. *Nature Climate Change*, **4**, 362–367, doi:10.1038/NCLIMATE2174.

Nicholls, N., G. Gruza, J. Jouzel, T. Karl, L. Ogallo, D. Parker, J. Christy, J. Eischeid, P. Groisman, M. Hulme, P. Jones, and R. Knight, 1995: Observed climate variability and change. *The Science of Climate Change. Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change, University Press*, 133–192.

North, G., F. Biondi, P. Bloomfield, J. R. Christy, K. Cuffey, R. E. Dickinson, E. R. M. Druffel, D. Nychka, B. Otto-Bliesner, N. Roberts, K. Turekian, J. M. Wallace, and I. Kraucunas, 2006: Surface temperature reconstructions for the last 2,000 years. *Washington, D.C.:National Research Council. The National Academies Press, 92 pp*.

Ohlwein, C. and E. Wahl, 2012: Review of probabilistic pollen-climate transfer methods. *Quaternary Science Reviews*, **31**, 17–29, doi:10.1016/j.quascirev.2011.11.002.

Oreskes, N., 2003: The role of quantitative models in science. *Models in ecosystem science*, 13.

Oreskes, N., K. Shrader-Frechette, K. Belitz, et al., 1994: Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, **263**, 641–646.

Ortega, P., F. Lehner, D. Swingedouw, V. Masson-Delmotte, C. C. Raible, M. Casado, and P. Yiou, 2015: A model-tested North Atlantic Oscillation reconstruction for the past millennium. *Nature*, **523**, 71–74.

Ottera, M., O. H.and Bentsen, H. Drange, and S. Lingling, 2010: External forcing as a metronome for atlantic multidecadal variability. *Nature Geosci.*, **4**, 99–103.

Otto-Bliesner, B. L., E. C. Brady, J. Fasullo, A. Jahn, L. Landrum, S. Stevenson, N. Rosenbloom, A. Mai, and G. Strand, 2015: Climate Variability and Change since 850 CE: An Ensemble Approach with the Community Earth System Model (CESM). *Bull. of the Amer. Met. Soc.*.

PAGES 2k Consortium, 2013: Continental-scale temperature variability during the past two millennia. *Nature Geosci.*, **6**, 339–346 doi: 10.1038/NGEO1797.

PAGES2k-PMIP3 group, 2015: Continental-scale temperature variability in PMIP3 simulations and PAGES 2k regional temperature reconstructions

over the past millennium. *Clim. of the Past Discuss.*, **11**, 2483–2555 doi: 10.5194/cpd–11–2483–2015.

Pauling, A., J. Luterbacher, and H. Wanner, 2003: Evaluation of proxies for european and north atlantic temperature field reconstructions. *Geophys. Res. Lett.*, **30**.

Peixoto, J. and A. Oort, 1984: Physics of climate. *Reviews of Modern Physics*, **56**, 365–429.

Phipps, S., L. Rotstayn, H. Gordon, J. Roberts, A. Hirst, and W. Budd, 2012: The csiro mk3l climate system model version 1.0–part 2: Response to external forcings. *Geoscientific Model Development*, **5**, 649–682.

Phipps, S. J., H. V. McGregor, J. Gergis, A. J. Gallant, R. Neukom, S. Stevenson, D. Ackerley, J. R. Brown, M. J. Fischer, and T. D. Van Ommen, 2013: Paleoclimate data–model comparison and the role of climate forcings over the past 1500 years. *J. Climate*, **26**, 6915–6936.

Plummer, C., M. Curran, T. van Ommen, S. Rasmussen, A. Moy, T. Vance, H. Clausen, B. Vinther, and P. Mayewski, 2012: An independently dated 2000-yr volcanic record from law dome, east antarctica, including a new perspective on the dating of the c. 1450s eruption of kuwae, vanuatu. *Clim. of the Past*, **8**, 1929–1940, doi:10.5194/cp–8–1929–2012.

Pollack, H. and J. Smerdon, 2004: Borehole climate reconstructions: Spatial structure and hemispheric averages. *J. Geophys. Res.*, **109**, D11106, doi:10.1029/2003JD004163.

Pongratz, J., T. Raddatz, C. Reick, M. Esch, and M. Claussen, 2009: Radiative forcing from anthropogenic land cover change since a. d. 800. *Geophys. Res. Lett.*, **36**, L02709, doi: 10.1029/2008GL036394.

Pongratz, J., C. Reick, T. Raddatz, and M. Claussen, 2008: A reconstruction of global agricultural areas and land cover for the last millennium. *Global Biogeochemical Cycles*, **22**, GB3018, 16 pp.

Pongratz, J., C. H. Reick, T. Raddatz, and M. Claussen, 2010: Biogeophysical versus biogeochemical climate response to historical anthropogenic land cover change. *Geophys. Res. Lett.*, **37**, L08702, doi: 10.1029/2010GL043010.

Rahmstorf, S., H. Von Storch, E. Zorita, J. Jones, F. González-Rouco, and S. Tett, 2006: Testing climate reconstructions. *Science*, **312**, 1872.

Ramankutty, N. and J. Foley, 1999: Estimating historical changes in global land cover: Croplands from 1700 to 1992. *Global Biogeochemical Cycles*, **13**, 997–1027.

Randall, D. A., R. Wood, S. Bony, R. Colman, T. Fichefet, J. Fyfe, V. Kattsov, A. Pitman, J. Shukla, J. Srinivasan, R. Stouffer, A. Sumi, and K. Taylor, 2007: *Climate Models and Their Evaluation. 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 [S. Solomon and D. Qin and M. Manning and Z. Chen and M. Marquis and K. B. Averyt and M. Tignor and H. L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Robock, A., 2000: Volcanic eruptions and climate. *Reviews of Geophysics*, **38**, 191–219.

Rutherford, S., M. E. Mann, T. J. Osborn, K. R. Briffa, P. Jones, R. S. Bradley, and M. K. Hughes, 2005: Proxy-Based Northern Hemisphere Surface Temperature Reconstructions: Sensitivity to Method, Predictor Network, Target Season, and Target Domain. *J. Climate*, **18**, 2308–2329, doi:10.1175/JCLI3351.1.

Schmidt, G., J. Jungclaus, C. Ammann, E. Bard, P. Braconnot, T. Crowley, G. Delaygue, F. Joos, N. Krivova, R. Muscheler, et al., 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.

Schmidt, G. A., J. H. Jungclaus, C. M. Ammann, E. Bard, P. Braconnot, T. J. Crowley, G. Delaygue, F. Joos, N. A. Krivova, R. Muscheler, B. L. Otto-Bliesner, J.Pongratz, D. T. Shindell, S. K. Solanki, F. Steinhilber, and L. E. A. Vieira, 2012: Climate forcing reconstructions for use in pmip simulations of the last millennium. *Geosci. Model Dev.*, **5**, 185–191.

Schmidt, G. A., M. Kelley, L. Nazarenko, R. Ruedy, G. L. Russell, I. Aleinov, M. Bauer, S. E. Bauer, M. K. Bhat, R. Bleck, et al., 2014: Configuration and assessment of the GISS ModelE2 contributions to the CMIP5 archive. *Journal of Advances in Modeling Earth Systems*, **6**, 141–184, doi:10.1002/2013MS000265.

Schneider, S. H., W. W. Kellogg, and V. Ramanathan, 1980: Carbon-dioxide and climate. *Science*, **210**, 6–7.

Schurer, A. P., G. C. Hegerl, M. E. Mann, S. F. B. Tett, and S. J. Phipps, 2013: Separating forced from chaotic climate variability over the past millennium. *J. Climatol.*, **26**, 6954–6973, doi: 10.1175/JCLI–D–12–00826.1.

Seager, R., N. Graham, C. Herweijer, A. Gordon, Y. Kushnir, and E. Cook, 2007: Blueprints for Medieval hydroclimate. *Quaternary Science Reviews*, **26**, 2322–2336.

Servonnat, J., P. Yiou, M. Khodri, D. Swingedouw, and S. Denvil, 2010: Influence of solar variability, CO 2 and orbital forcing during the last millennium in the IPSLCM4 model. *Clim. of the Past*, **6**, 445–460.

Shapiro, A. I., W. Schmutz, E. Rozanov, M. Schoell, M. Haberreiter, A. V. Shapiro, and S. Nyeki, 2011: A new approach to long-term reconstruction of the solar irradiance leads to large historical solar forcing. *Astronomy & Astrophysics*, **529**, A67, doi:10.1051/0004–6361/201016173.

Shi, F., B. Yang, A. Mairesse, L. von Gunten, J. Li, A. Bräuning, F. Yang, and X. Xiao, 2013: Northern hemisphere temperature reconstruction during the

last millennium using multiple annual proxies. *Clim Res*, **56**, 231–244, doi: 10.3354/cr01156.

Shindell, D., G. Schmidt, M. Mann, D. Rind, and A. Waple, 2001: Solar forcing of regional climate change during the Maunder Minimum. *Nature*, **294**, 2149–2152.

Smerdon, J. E., 2012: Climate models as a test bed for climate reconstruction methods: pseudoproxy experiments. *WIREs Clim. Change*, **3**, 63–77, doi: 10.1002/wcc.149.

Smerdon, J. E., S. Coats, and T. R. Ault, 2015: Model-dependent spatial skill in pseudoproxy experiments testing climate field reconstruction methods for the Common Era. *Clim. Dyn.*, 1–22.

Smerdon, J. E., A. Kaplan, D. Chang, and M. N. Evans, 2010: A Pseudoproxy Evaluation of the CCA and RegEM Methods for Reconstructing Climate Fields of the Last Millennium. *J. Climate*, **23**, 4856–4880, doi: 10.1175/2010JCLI3328.1.

Smerdon, J. E., A. Kaplan, E. Zorita, J. F. González-Rouco, and M. N. Evans, 2011: Spatial performance of four climate field reconstruction methods targeting the common era. *Geophys. Res. Lett.*, **38**, L11705, doi:10.1029/2011GL047372.

Soden, B. and I. Held, 2006: An assessment of climate feedbacks in coupled ocean-atmosphere models. *J. Climate*, **19**, 3354–3360.

Solanki, S. K. and N. Krivova, 2004: Solar irradiance variations: from current measurements to long-term estimates. *Solar Physics*, **224**, 197–208.

Solanki, S. K., I. G. Usoskin, B. Kromer, M. Schuessler, and J. Beer, 2004: Unusual activity of the sun during recent decades compared to the previous 11000 years. *Nature*, **431**, 1084–1087.

Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. M. (eds.), 2007: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Steinhilber, F., J. Beer, and C. Frohlich, 2009: Total solar irradiance during the holocene. *Geophys. Res. Lett.*, **36**, L19704, doi:10.1029/2009GL040142.

Stocker, T., D. Qin, G.-K. Plattner, L. Alexander, S. Allen, N. Bindoff, F.-M. Bréon, J. Church, U. Cubasch, S. Emori, et al., 2013: Technical summary. Climate Change 2013: The Physical Science Basis. *Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, 33–115.

Sundberg, R., A. Moberg, and A. Hind, 2012: Statistical framework for evaluation of climate model simulations byuse of climate proxy data from the last millennium: Part 1: Theory. *Clim. of the Past*, **8**, 1339–1353.

Sutton, R., B. Dong, and J. Gregory, 2007: Land/sea warming ratio in response to climate change: IPCC AR4 model results and comparison with observations. *Geophys. Res. Lett.*, **34**, L02701, doi:10.1029/2006GL028164.

Swingedouw, D., L. Terray, C. Cassou, A. Voldoire, D. Salas-Mélia, and J. Servonnat, 2010: Natural forcing of climate during the last millennium: fingerprint of solar variability. *Clim. Dyn.*, **36**, published online, 1–16.

Taylor, K., R. Stouffer, and G. Meehl, 2012: An overview of cmip5 and the experiment design. *Bull. of the Amer. Met. Soc.*, **93**, 485–498.

Tett, S., R. Betts, T. Crowley, J. Gregory, T. Johns, A. Jones, T. Osborn, E. Öström, D. Roberts, and M. Woodage, 2007: The impact of natural and anthropogenic forcings on climate and hydrology since 1550. *Clim. Dyn.*, **28**, 3–34.

Timmreck, C., S. J. Lorenz, T. J. Crowley, S. Kinne, T. J. Raddatz, M. A. Thomas, and J. H. Jungclaus, 2009: Limited temperature response to the very large AD 1258 volcanic eruption. *Geophys. Res. Lett.*, **36**.

Tingley, M. P., P. F. Craigmile, M. Haran, B. Li, E. Mannshardt, and B. Rajaratnam, 2012: Piecing together the past: statistical insights into paleoclimatic reconstructions. *Quaternary Science Reviews*, **35**, 1–22.

Tingley, M. P. and P. Huybers, 2010: A bayesian algorithm for reconstructing climate anomalies in space and time. Part I: development and applications to paleoclimate reconstruction problems. *J. Climate*, **7**, 2759–2781, doi:10.1175/2009JCLI3015.1.

Tingley, M. P. and B. Li, 2012: Comments on 'reconstructing the nh mean temperature: Can underestimation of trends and variability be avoided?'. *J. Climate*, **25**, 3441–3446, doi:10.1175/2009JCLI3016.1.

Trenberth, K. E., J. T. Fasullo, and J. Kiehl, 2009: Earth's global energy budget. *Bull. of the Amer. Met. Soc.*, **33**, 311–323, DOI:10.1175/2008BAMS2634.1.

Trouet, V., J. Esper, N. Graham, A. Baker, J. Scourse, and D. Frank, 2009: Persistent positive North Atlantic Oscillation mode dominated the medieval climate anomaly. *Science*, **324**, 78–80.

Trouet, V., J. D. Scourse, and C. C. Raible, 2012: North atlantic storminess and atlantic meridional overturning circulation during the last millennium: reconciling contradictory proxy records of nao variability. *Global and Planetary Change*, **84-85**, 48–55, doi:10.1016/j.gloplacha.2011.10.003.

Usoskin, I. G., S. K. Solanki, and G. A. Kovaltsov, 2007: Grand minima of solar activity: new observational constraints. *Astron. Astrophys.*, **471**, 301–309.

Vaughan, D. G., J. C. Comiso, I. Allison, J. Carrasco, G. Kaser, R. Kwok, P. Mote, T. Murray, F. Paul, J. Ren, et al., 2013: Observations: cryosphere. climate change 2013: The physical science basis. *Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate change*, 317–382.

Vieira, L. and S. Solanki, 2010: Evolution of the solar magnetic flux on time scales of years to millennia. *Astronomy & Astrophysics*, **509**, A100, doi: 10.1051/0004–6361/200913276.

von Storch, H., 1995: Inconsistencies at the interface of climate impact studies and global climate research. *Meteorol. Zeitschrift*, **4**, 71–80.

— 1999: On the use of inflation in statistical downscaling. *J. Climate*, **12**, 3505–3506.

— 2004: *A discourse about quasi-realistic climate models and their applications in paleoclimate studies*. H. Fischer, T. Kumke, G. Lohmann, G. Flser, H. Miller, H von Storch and J. F. W. Negendank (eds.) The Climate in Historical Times. towards a Synthesis of Holocene Proxy Data and Climate Models, Springer Verlag, Berlin, Heidelberg, New York, 43-56.

— 2010: Climate models and modeling: an editorial essay. *Wiley Interdisciplinary Reviews: Climate Change*, **1**, 305–310.

Von Storch, H. and E. Zorita, 2005: Comment on Hockey sticks, principal components, and spurious significance by S. McIntyre and R. McKitrick. *Geophys. Res. Lett.*, **32**.

Von Storch, H., E. Zorita, J. Jones, Y. Dimitriev, F. Gonzalez-Rouco, and S. Tett, 2004: Reconstructing past climate from noisy data. *Science*, **306**, 679–682.

Von Storch, H., E. Zorita, J. Jones, F. González-Rouco, and S. Tett, 2006a: Response to "Testing climate reconstructions" by S. Rahmstorf. *Science*, **312**, 1872–1873.

Von Storch, H., E. Zorita, J. Jones, F. González-Rouco, and S. Tett, 2006b: Response to the comment by wahl, ritson and ammann reconstruction of century scale temperature variations,. *Science*, **312**, 529.

Wagner, S., , M. Widmann, J. Jones, T. Haberzettl, A. Lücke, C. Mayr, C. Ohlendorf, F. Schäbitz, and B. Zolitschka, 2007: Transient simulations, empirical reconstructions and forcing mechanisms for the mid-holocene hydrological climate in southern patagonia. *Clim. Dyn.*, **29**, 333–355, doi:10.1007/s00382–007–0229–x.

Wahl, E. R., D. M. Ritson, and C. M. Ammann, 2006: Comment on" Reconstructing past climate from noisy data". *Science*, **312**, 529–529.

Wang, Y., J. Lean, and N. Sheeley, 2005: Modeling the Sun's magnetic field and irradiance since 1713. *The Astrophysical Journal*, **625**, 522–538.

Werner, J., J. Luterbacher, and J. Smerdon, 2013: A Pseudoproxy Evaluation of Bayesian Hierarchical Modelling and Canonical Correlation Analysis for Climate Field Reconstructions over Europe. *J. Climate*, **26**, 851–867.

Wigley, T. M. L., C. M. Ammann, B. D. Santer, and S. C. B. Raper, 2005: Effect of climate sensitivity on the response to volcanic forcing. *J. Geophys. Res.*, **110**, D09107, doi:10.1029/2004JD005557.

Wilson, M. F. and A. Henderson-Sellers, 1985: A global archive of land cover and soils data for use in general circulation climate models. *J. Climatol*, **5**, 119–143.

Yiou, P., J. Servonnat, M. Yoshimori, D. Swingedouw, M. Khodri, and A. Abe-Ouchi, 2012: Stability of weather regimes during the last millennium from climate simulations. *Geophysical Research Letters*, **39**, L08703, doi:10.1029/2012GL051310.

Zebiak, S. and M. Cane, 1987: A model el nino-southern oscillation. *Monthly Weather Review*, **115**, 2262–2278.

Zhang, D., R. Blender, X. Zhu, and K. Fraedrich, 2010: Temperature variability in china in an ensemble simulation for the last 1200 years. *Theor. Appl. Climatol.*, **103**, 387–399. doi:10.1007/s00704–010–0305–8.

Zhou, T., B. Li, W. Man, L. Zhang, and J. Zhang, 2011: A comparison of the medieval warm period, little ice age and 20th century warming simulated by the fgoals climate system model. *Chinese Science Bulletin*, **56**, 3028–3041d, doi: 10.1007/s11434–011–4641–6.

Zorita, E., J. González-Rouco, H. Von Storch, J. Montávez, and F. Valero, 2005: Natural and anthropogenic modes of surface temperature variations in the last thousand years. *Geophys. Res. Lett.*, **32**, 755–762.