Climate and carbon-cycle variability over the last millennium

J. H. Jungclaus¹, S. J. Lorenz¹, C. Timmreck¹, C. H. Reich¹, V. Brovkin¹, K. Six¹, J. Segschneider¹, M. A. Giorgetta¹, T. J. Crowley², J. Pongratz¹,*, N. A. Krivova³, L. E. Vieira³**, S. K. Solanki³, ⁴, D. Klocke¹, M. Botzet¹, M. Esch¹, V. Gayler¹, H. Haak¹, T. J. Raddatz¹, E. Roeckner¹, R. Schnur¹, H. Widmann¹, M. Claussen¹, ⁴, B. Stevens¹, ⁵, and J. Marotzke¹

¹Max Planck Institute for Meteorology, Bundesstrasse 53, 20146 Hamburg, Germany
²School of Geosciences, University of Edinburgh, West Mains Road, Edinburgh EH9 3JW, UK
³Max Planck Institute for Solar System Research, Max Planck Strasse 2, 37191 Katlenburg-Lindau, Germany
⁴KlimaCampus, University of Hamburg, Bundesstrasse 55, 20146 Hamburg, Germany
Climate and carbon-cycle variability over the last millennium

J. H. Jungclaus et al.

Department of Atmospheric and Oceanic Sciences, University of California Los Angeles, Los Angeles, CA 90095, USA

*now at: Department of Global Ecology, Carnegie Institution of Washington, 260 Panama St., Stanford, CA 94305, USA

**now at: Laboratoire de Physique et Chemie de l’Environnement et de l’Espace, CNRS & University of Orleans, 3A Av. De la Recherche Scientifique, 45071 Orléans, France

Received: 4 May 2010 – Accepted: 12 May 2010 – Published: 26 May 2010

Correspondence to: J. H. Jungclaus (johann.jungclaus@zmaw.de)

Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

A long-standing task in climate research has been to distinguish between anthropogenic climate change and natural climate variability. A prerequisite for fulfilling this task is the understanding of the relative roles of external drivers and internal variability of climate and the carbon cycle. Here, we present the first ensemble simulations over the last 1200 years with a comprehensive Earth system model including a fully interactive carbon cycle. Applying up-to-date reconstructions of external forcing including the recent low-amplitude estimates of solar variations, the ensemble simulations reproduce temperature evolutions consistent with the range of reconstructions. The 20th-century warming trend stands out against all pre-industrial trends within the ensemble. Volcanic eruptions are necessary to explain variations in pre-industrial climate such as the Little Ice Age; yet only the strongest, repeated eruptions lead to cooling trends that stand out against the internal variability across all ensemble members. The simulated atmospheric CO$_2$ concentrations exhibit a stable carbon cycle over the pre-industrial era with multi-centennial variations somewhat smaller than in the observational records. Early land-cover changes have modulated atmospheric CO$_2$ concentrations only slightly. We provide a model-based quantification of the sensitivity (termed $\gamma$) of the global carbon cycle to temperature for a variety of climate and forcing conditions. The magnitude of $\gamma$ agrees with a recent statistical assessment based on reconstruction data. We diagnose a distinct dependence of $\gamma$ on the forcing strength and time-scales involved, thus providing an explanation for the systematic difference in the observational estimates for different segments of the last millennium.

1 Introduction

Northern Hemisphere (NH) temperature reconstructions (Jansen et al., 2007; Mann et al., 2008, 2009) for the last millennium differ substantially among each other, raising questions (Trouet et al., 2009; Mann et al., 2009) about the spatial and temporal extent
of climatic epochs such as the Medieval Warm Period (MWP, ca. 900–1300 AD) and the Little Ice Age (LIA, ca. 1500–1850 AD). Climate models’ ability to reproduce the observed records can help understand the mechanisms behind the climate variability of the last millennium and provides context for current and future climate change. Alternatively, discrepancies between models and observations can help identify gaps in our understanding, or possible inaccuracies in the observational record. A model system that simulates both the carbon-cycle and the spatially resolved internal climate variability allows us to exploit the fullness of the observational record in order to quantify both the forced response of the system and the carbon-cycle climate feedback over a variety of periods and boundary conditions. However, to assess internal variability, it is essential to perform ensemble integrations in a manner that begins to spatially resolve the dynamics of the climate system. Here, we present the first such simulations for the last millennium, with a comprehensive Earth System Model (ESM) including a fully interactive carbon cycle. This is a significant advance over previous efforts, which have been restricted to Energy Balance Models (Crowley, 2000) and ESMs of Intermediate Complexity (e.g., Gerber et al., 2003; Goosse et al., 2005), or to single realisations of coupled models without a carbon cycle (González-Rouco et al., 2003; Ammann et al., 2007).

In this paper we concentrate on the analysis of the relative role of internal variability and external forcing for shaping NH temperature and on the evolution of the carbon cycle over the last 1200 years. In Sect. 2, we describe the experimental design, the Earth system model, and the external forcing datasets. Readers who are not interested in the technical details may skip Sects. 2.1 and 2.2. Section 3 presents the results starting with an analysis of NH temperatures in relation with reconstructions, followed by an assessment of trends in forced and unforced simulations. In Sect. 3.4 we discuss the simulated CO₂ evolution in the forced experiments and (Sect. 3.5) assess the sensitivity of the carbon cycle to temperature changes.
2 Model and experimental design

The experimental strategy is briefly described as follows. After a multi-century spin-up phase in which the carbon cycle was brought into equilibrium, we ran a 3000-year unforced control experiment under 800 AD orbital conditions and pre-industrial greenhouse gas concentrations. Starting from different ocean initial conditions we performed a five-member ensemble (E1) with the standard external forcing set spanning the time 800 AD to 2005 AD. In order to account for uncertainty in the solar forcing we ran a second three-member ensemble (E2) with an alternative reconstruction of solar irradiance (see below). In addition, we have performed sensitivity experiments with just one external forcing at a time.

2.1 The Earth System Model

We have used the Max Planck Institute for Meteorology (MPI-M) ESM consisting of the general circulation models for the atmosphere ECHAM5 (Roeckner et al., 2003) and for the ocean MPIOM (Marsland et al., 2003). ECHAM5 is run at T31 resolution (3.75°) with 19 vertical levels, resolving the atmosphere up to 10 hPa. MPIOM applies a conformal mapping grid with a horizontal resolution ranging from 22 km to 350 km. This grid set-up is a low-resolution version of the model used for the scenario simulations (Jungclaus et al., 2006) for the Intergovernmental Panel of Climate Change (IPCC) and the Coupled Carbon Cycle Climate Modelling Intercomparison Project (C4MIP, Friedlingstein et al., 2006) simulations. Ocean and atmosphere are coupled daily without flux corrections using the OASIS3 coupler (Valcke et al., 2003). The carbon cycle model comprises the ocean biogeochemistry module HAMOCC5 (Wetzel et al., 2006) and the land surface scheme JSBACH (Raddatz et al., 2007). Three-dimensional transport of carbon within the ocean and the atmosphere as well as the exchange between atmosphere and land biosphere are calculated each time step so that the daily and seasonal cycle of the atmospheric CO$_2$ concentrations are resolved. The carbon flux between the ocean and the atmosphere is determined daily at each coupling time step.
2.2 External forcing

2.2.1 Solar forcing

Variations in Total Solar Irradiance (TSI) over the last centuries have recently been argued to be much smaller than previously thought (Wang et al., 2005; Krivova et al., 2007). Therefore, the TSI forcing used as our standard forcing exhibits a total increase of 0.1% (∼1.3 W m\(^{-2}\)) from the Maunder Minimum (1647–1715 AD) to today, which is in agreement with other recent evaluations (Steinhilber et al., 2009; Tapping et al., 2007). The data set has daily sampling from 1627 AD onward (Krivova et al., 2007; Balmaceda et al., 2007). In this period, the 11-year activity cycle is included since solar irradiance was reconstructed from historical records of sunspot numbers. Prior to the Maunder Minimum, no direct estimates of solar variability, such as sunspot observations are available. The TSI time series for the period 800 AD to the Maunder Minimum is reconstructed from estimates of the solar open magnetic flux based on cosmogenic isotope \(^{14}\)C concentrations in tree rings (Solanki et al., 2004; Krivova and Solanki, 2008; Usoskin et al., 2007). In order to derive a consistent time series applied as solar forcing for the model, an 11-years cycle has been artificially superimposed prior to the Maunder Minimum. The amplitude was estimated from the analysis of the relationship between the cycle amplitude and overall solar activity, as represented by the 11-year running mean of the TSI between 1700 AD and present. To account for uncertainty in the amplitude in solar forcing and to enable comparison with earlier model studies (González-Rouco et al., 2003; Ammann et al., 2007), we ran a second ensemble (E2) of three simulations with considerably stronger Maunder-Minimum-reduction. The TSI series used for the E2 ensemble is based on another compilation of \(^{14}\)C and \(^{10}\)Be production rates (Bard et al., 2000) from which an annual data set was derived (Ammann et al., 2007). The E2 solar forcing exhibits roughly 3.5 W m\(^{-2}\) (0.25%) change from the Maunder Minimum to the modern period (1950–2000 AD). Both TSI forcing series are shifted by a constant in order to start with a solar constant of 1367 W m\(^{-2}\) as it is used in the control run for the coupled model.
2.2.2 Volcanic forcing

The volcanic effect on radiation is calculated online in the model using time series of aerosol optical depth (AOD) at 0.55 µm and of the effective radius ($R_{\text{eff}}$) (Crowley et al., 2008). The time resolution of the series is ten days and the data are provided at four equal area latitude bands. AOD estimates are based on a correlation between sulphate in Antarctic ice cores and satellite AOD data from the 1991 eruptions of Mt. Pinatubo and Cerro Hudson (Sato et al., 1993). $R_{\text{eff}}$ growth and decay is based on satellite observations of the Pinatubo eruption in 1991, and to particle sizes of very large eruptions by microphysical simulations (Pinto et al., 1989). In the model AOD is distributed between 20–86 hPa over three vertical levels, with a maximum at 50 hPa. Extinction, single scattering albedo and the asymmetry factor are calculated online for the six solar bands (0.185–4 µm) of the ECHAM5 radiation scheme and extinction for the 16 long-wave wavelength bands (3.3–100 µm) from the time dependent AOD, $R_{\text{eff}}$ and normalized optical parameters. The normalized parameters are calculated for all wavelength bands dependent on the assumed aerosol size distribution with an effective radius between 0.02 and 1 µm and a constant standard deviation of 1.8.

Sensitivity experiments for the model response to the Pinatubo eruption yield an average global temperature change (0.4 K) comparable to observations. For the largest eruption of the last millennium, the 1258 AD eruption, a NH summer temperature anomaly over land of 1.2 K is found in agreement with reconstructions (Timmreck et al., 2009).

2.2.3 Land cover changes

Anthropogenic land cover change is considered by applying the reconstruction of global agricultural areas and land cover (Pongratz et al., 2008). The global maps with a spatial resolution of 0.5° and an annual timescale contain 14 vegetation types and discriminate between the agricultural categories cropland, and C3 and C4 pastures. The reconstruction merges published maps of agriculture from AD 1700 to 1992 and
a population-based approach to quantify agriculture from AD 800 to 1700. This approach captures the general expansion of agriculture but also abandonment of agricultural areas and subsequent re-growth of natural vegetation due to regionally confined historical events, such as warfare and epidemics (Pongratz et al., 2009).

2.2.4 Orbital forcing

The ECHAM5 model contains a representation of periodic changes of the Earth’s orbit around the Sun on all frequencies, including short term variations. It uses the Variation Seculaires des Orbites Planetaires (VSOP) analytical solution (Bretagnon and Francou, 1988) and determines today’s orbit for an interval −4000 to 8000 years with respect to the year 2000 accurately. Additionally, the nutation, a small wobble of the Earth’s rotational axis with a period of 18.6 years is accounted for in the model.

2.2.5 Greenhouse gas forcing

The CO₂ concentration is calculated interactively within the model. The concentrations of the next two major greenhouse gases, methane (CH₄) and nitrous oxide (N₂O) are prescribed (MacFarling Meure et al., 2006). The data sets are applied to the model using annual resolution, where a simple spline function is used for smoothing the data. For all simulations, present day ozone climatology (Fortuin and Kelder, 1998) is prescribed.

2.2.6 Aerosol forcing

The climatology of the background aerosol distribution (Tanre et al., 1984) distinguishes time independent spatial distributions of tropospheric and stratospheric background aerosols. The aerosols are described by a maximum AOD at 0.55 µm, normalized horizontal distributions, normalized vertical integrals, and a troposphere-stratosphere discrimination factor. The AOD at 0.55 µm is rescaled to the spectral intervals of the
ECHAM radiation scheme. The natural and anthropogenic sulphate dry mass are provided as monthly means in µg SO$_4^{2-}$ m$^{-3}$ and converted online to optical parameters. The data were interpolated linearly between 1750 (i.e. natural pre-industrial distribution) and 1850 and the 1850 to 1980 sulphate fields are based on historical reconstructions (Lefohn et al., 1999). The sulphur emission scenarios are based on model simulations (Boucher and Pham, 2002).

3 Evolution of Northern Hemisphere temperature and global CO$_2$ concentration over the last 1200 years

3.1 20th century temperature and CO$_2$ evolution

As a first step we demonstrate the ability of the model to reproduce important aspects of the recent period of global climate change. The simulated NH temperature evolution over the 20th century agrees well with the instrumental record (Fig. 1a). Both observed and modelled time series exhibit a warming trend of about 0.6°C over the 20th century that is superimposed by pronounced multidecadal variability. The two ensembles with different solar forcing do not differ significantly from each other, indicating that greenhouse-gas forcing and internal variability dominate over solar forcing over the last century. The global CO$_2$ increase in the 20th century (Fig. 1b) shows somewhat less upward trend than the (local) Mauna-Loa record. However, the deviations between simulated and observed CO$_2$ concentration at the beginning of the 21st century are well in the range of state-of-the-art climate carbon models, such as those carried out in the framework of C4MIP (Friedlingstein et al., 2006; Raddatz et al., 2007). The CO$_2$ increase from land-cover changes is moderate compared to contribution from fossil-fuel emissions. Over the last millennium, land-cover changes contribute roughly 20 ppm (Pongratz et al., 2009).
3.2 Northern Hemisphere temperature changes over the last 1200 years

Over time-intervals from decades to centuries, simulated NH temperatures (Fig. 2a) from all ensemble members differ significantly from the range of internal variability defined by the control experiment. Strong volcanic eruptions, particularly the cumulative effect of several volcanoes around the most severe eruptions in 1258 AD, 1453 AD, and 1815 AD, leave a long-lasting imprint on NH climate. Note that the generally somewhat cooler mean states in the ensembles are a result of the absence of volcanic aerosol forcing in the control run. Modulation by changing solar irradiance is more pronounced in the E2 ensemble where we identify the largest pre-industrial temperature anomalies in the 15th century during a superposition of the 1453 AD Kuwae eruption and the Spörer-Minimum (1450–1550 AD) in TSI. Overall, the simulations show the warmest pre-industrial NH temperatures around 1050–1250 AD and in the late 18th century while cold anomalies prevail during the 13th, 15th, 17th and early 19th centuries. In contrast to the E1 experiments, the E2 ensemble exhibits a notable MWP that is associated with the peak in the solar forcing in the 12th century (Fig. 2c). Most temperature reconstructions, however, indicate a MWP centred on the turn of the millennium (Jansen et al., 2007; Mann et al., 2008, 2009). From the 15th to the mid-18th century, the two ensembles do not overlap and the ensemble mean of the E2 simulations agrees somewhat better with the consensus of reconstructions. However, there is considerable variation within the ensembles. The ensemble spreads (Fig. 2b) are of similar magnitude as the control run’s range of internal variability, indicating that they provide an adequate representation of internal variability even though the ensemble size is small. On the other hand, of the two ensembles, E1 has an overall higher (0.23 K) ensemble spread than E2 (0.18 K) showing that E2 temperatures are constrained more strongly by the external forcing (see Sect. 3.3). It is interesting to note that the E1 ensemble exhibits a large spread in the individual realisations almost continuously between 1500 and 1700 AD. At the time where most of the reconstructions show the coldest part of the LIA (1600–1650 AD), a “cold” E1 realization gives a
negative temperature anomaly nearly as strong as the much more strongly forced E2 simulations. Therefore, multi-century climate swings such as the MWP-LIA transition may not require particular strong solar forcing but can be attributed, at least in parts, to internal variability on centennial time-scales.

Simulations with weak solar forcing are also consistent with reconstructions of the transition between MWP and LIA. We follow a recently proposed methodology (Frank et al., 2010) wherein the warmest 30 year period during the MWP epoch is compared with the coldest 30 year period during the time of the LIA (Fig. 3). We find that the strength of the MWP-LIA transition as measured in this way depends somewhat on its seasonal and regional representation. Results are qualitatively similar also for 100-year periods (not shown). For the 30-year segments, the low-solar-amplitude ensemble (E1) is more consistent with Frank et al.’s best estimate of 0.38 K of the strength of this transition. Sensitivity experiments forced solely with varying solar irradiance yield a sensitivity of 0.1 K global temperature change per W m\(^{-2}\) (TSI). Recent assessments over the last century (Camp and Tung, 2007; Lean and Rind, 2008) arrive at 0.2 and 0.1 K per W m\(^{-2}\), respectively. Therefore it seems unlikely that too large a model climate sensitivity is compensating for a weak forcing. Moreover, with a larger sensitivity the model would then agree less well with the 20th century record (Fig. 1). We thus conclude that solar variations with much stronger amplitude than the most recent reconstructions are not necessary to explain climate variations such as the MWP-LIA transition, and that differences among reconstructions may partly be explained as a result of different spatio-temporal sampling among the records.

### 3.3 NH temperature trends

The long control experiment and the ensembles of forced experiments allow us to analyze whether decadal to centennial warming and cooling trends in NH-mean temperature fall outside the range of internal variability. To identify periods during the last millennium where the NH temperatures in the forced simulations leave the range of internal variability, we apply detection methods previously developed for the discrimination...
between anthropogenic climate change and natural variability (Santer et al., 1995; Baehr et al., 2008). In a first step, NH temperatures from the control run are sampled for a specific segment length (Santer et al., 1995) estimating linear trends for 2900 years. For a given length of the estimation period, the probability density function (pdf) derived from the linear trends represents the variability of the unforced system. Following Baehr et al. (2008), the procedure is repeated for segment lengths between 1 and 100 years to derive upper and lower confidence limits for each segment length. Thereafter, linear trends derived throughout the experiments are compared with the confidence limits for the respective segment lengths (see Appendix C). For certain periods, trends over segment lengths between one and 100 years are significantly different from those based on internal variability, with a random probability of occurrence of less than 0.05. In Fig. 4 these periods are indicated by colour shading where the respective depth of the shading indicates the number of ensemble members showing the same behaviour at the same time. The darkest shades of red or blue, respectively, indicate that all ensemble members show a statistically significant deviation from the control experiment. Only the 20th century global warming and the strongest volcanic eruptions leave an imprint of, respectively warming and cooling trends throughout all eight experiments and over all timescales from a few years to a century. Clearly, the stronger solar variability in the E2 ensemble leads to several more significant warming and cooling trends across the entire ensemble. On the other hand, light shadings in Fig. 4a indicate periods (including parts of the MWP and LIA) in E1 where internal variability plays a larger role.

3.4 Evolution of atmospheric CO₂ over the last 1200 years

Pre-industrial CO₂ variations during the last millennium are small in comparison with glacial-cycle changes, but ice-core records indicate a long-term decline from medieval times to the LIA with considerable difference among various sites. Most notable in the reconstructions (MacFarling Meure et al., 2006; Siegenthaler et al., 2005) is a period of relatively low CO₂ concentration that lasted from the late 16th to the mid
18th century, during which time the Law Dome data set indicates a pronounced drop by about 8 ppm. It has been hypothesized that land cover changes in the aftermath of warfare and epidemics (Ruddiman, 2007) have caused natural vegetation re-growth and CO₂ sequestration on decadal to centennial time scales. Other studies (Joos et al., 1999; Trudinger et al., 2002) deny an active role of anthropogenic CO₂ in shaping the LIA climate but argue that the CO₂ drop around 1600 AD was a response of the terrestrial biosphere and the oceans to the LIA cooling. Our experiments allow us to explore the two competing hypotheses.

Both model ensembles exhibit relatively stable CO₂ concentrations (Fig. 5a) over the pre-industrial era. During the 20th century the full forcing runs reproduce the observed increase in atmospheric CO₂ similar to other carbon-cycle models (Fig. 1b). The E1 realisations fall below the ±1 ppm 5–95 percentile variability of the control run only after the very strong 1258 AD volcanic eruption, and rise above the control run variability only from the early 18th century on. The E2 realisations simulate CO₂ significantly high during the MWP and significantly low during the LIA, akin to the NH temperatures, but none of the simulations reaches the amplitude suggested by the Law Dome data.

The simulated land-to-atmosphere carbon fluxes (Fig. 6) indicate that the land biosphere was a source of CO₂ throughout most of the millennium, with increasing strength after 1700 AD owing to increasing land-cover changes (Pongratz et al., 2009). The role of the land biosphere as a source was punctuated by brief episodes where it acted as a sink, following volcanic eruptions; carbon uptake then occurred mostly in tropics as a consequence of reduced heterotrophic respiration on land in response to surface cooling (Jones and Cox, 2001). The land biosphere carbon source throughout most of the millennium was partly offset by a consistent ocean uptake, mirroring in strength and timing the land source. After 1950, CO₂ fertilization turned the land biosphere into a sink, so since 1950 both ocean and land have acted as massive carbon sinks, again signalling the exceptional role of that period.

Solar radiation changes, land-cover changes, and volcanic eruptions have competing impacts on the carbon budget and on atmospheric CO₂ concentration as demonstrated...
by experiments where just one forcing component was applied (Fig. 5b). Solar forcing modulate atmospheric CO$_2$ as can be most clearly seen in the experiment with the stronger amplitude TSI reconstruction. However, while the E2 ensemble gives generally lower CO$_2$ concentrations throughout the LIA (Fig. 5a), the pronounced drop in the early 17th century can not be associated with solar forcing because both TSI (Fig. 2c) and CO$_2$ (Fig. 5b, dark blue line) records show a positive anomaly around 1610 AD. Before 1700 AD, land cover changes modulate the CO$_2$ record only by a few ppm, slightly exceeding the range of internal variability. While the effect of abandonment of agricultural areas due to warfare and epidemics is discernable in regional and global emissions (Pongratz et al., 2009), it is not sufficient to explain the apparent CO$_2$ decrease during the LIA. On the contrary, after 1500 AD, atmospheric CO$_2$ in the land-cover-change-only experiment is almost always higher than the mean of the control experiment. Similar to the solar modulation, cooling by strong and/or cumulative volcanic activity causes changes in the CO$_2$ record of several ppm. The mid 13th century eruption leaves a long-lasting imprint (Brovkin et al., 2010). The atmospheric carbon loss through volcanic activity in the 18th to 19th century (Fig. 5b) is overcompensated by a massive effect from land-cover-changes. In fact, the strong (and, in comparison with the reconstructions too early) rise in CO$_2$ levels in the full-forcing experiments (Fig. 5a) can almost exclusively attributed to the increase of emissions from agriculture (Fig. 5b). This emission increase is one possible reason for the discrepancies between our simulations and the CO$_2$ reconstructions. Other explanations include underestimation of the MWP-LIA cooling or the temperature-carbon cycle feedback.

The processes controlling carbon fluxes between the atmosphere, biosphere, and the oceans are temperature dependent and, on glacial timescales, the sensitivity of the global carbon cycle to temperature $\gamma$, is roughly linear with a slope of about 8 ppm K$^{-1}$ (Woodwell et al., 1998). While empirical estimates based on last-millennium data have reported values up to 40 ppm K$^{-1}$ (Schefter et al., 2006; Cox and Jones, 2008), a recent assessment (Frank et al., 2010) quantified the median $\gamma$ as 7.7 ppm K$^{-1}$, with a likely range of 1.7–21.4 ppm K$^{-1}$. For our simulations $\gamma$ falls well within this range (Fig. 7a),
but, strikingly, $\gamma$ is much larger for the forced as compared to the unforced simulations. The regression slope reads 5.6 ppm K$^{-1}$ for the E2 ensemble, but it is considerably smaller for the control experiment (3.2 ppm K$^{-1}$). An analysis in the frequency domain (Fig. 7b) reveals increasing sensitivity on longer (centennial to millennial) time-scale. Moreover, running regressions over the pre-industrial last millennium (Fig. 7c, d) reveal that $\gamma$ is not time-invariant but varies on multidecadal to centennial time-scales. Both control and forced simulation indicate the strongest response of atmospheric CO$_2$ variations at a time lag of roughly ten years, but the amplitude in the forced run is much higher reaching at regression slopes of up to 12 ppm K$^{-1}$. Such centennial-scale variations in the sensitivity are apparently also present in the observational record. Frank et al. (2010) found considerable variations for the first (4.3 ppm K$^{-1}$) and second (16.1 ppm K$^{-1}$) half of the pre-industrial period. The stronger response in the forced simulations may reflect non-linearities in the system, or the different spatio-temporal structure of the temperature patterns in the forced simulations. The mechanisms behind the carbon-cycle response to external forcing have been investigated in a separate study focusing on the impact of a strong volcanic eruption. Brovkin et al. (2010) analyzed the time period around the eruption of the 1258 unknown volcano in the same experiments. They conclude that the CO$_2$ decrease in the atmosphere is explained mainly by reduced heterotrophic respiration on land in response to surface cooling corroborating findings by Jones and Cox (2001). Furthermore, the magnitude of the atmospheric response is determined by the land carbon storage while its duration is set-up by the marine carbon cycle. In particular the stronger sensitivity at low frequencies (Fig. 6b) suggests that these slow processes associated with carbon storage in the biosphere and oceans determine the feedback strength. Therefore, the slowly varying solar irradiance changes and the cumulative effect of volcanoes that lead to multi-centennial climate variations provide time-scales in which the carbon-cycle response can fully develop.
4 Discussion and conclusion

For the first time, ensemble experiments using a comprehensive ESM including a fully-interactive carbon-cycle have been analysed with respect to climate variability and climate-carbon-cycle sensitivity. The last millennium with its high-resolution proxies of surface temperatures, volcanic and solar forcing and concomitant CO₂ changes challenges our understanding of the climate system. While many of the features of the observed record appear compatible with our simulations and serve to highlight the peculiarity of the present epoch, some mysteries remain. In particular, the magnitude and rate of CO₂ change during the LIA and the timing of the MWP prove difficult to reconcile with our best estimates of the climate forcing and response over the last millennium.

The experiments presented here are among the first ESM simulations that comply with the protocols of the Paleo Modelling Intersomparison Project Phase 3 (PMIP-3, http://pmip3.lsce.ipsl.fr) and the upcoming Paleo Carbon Model Intercomparison Project (PCMIP). Analysing the role of external forcings and internal variability and the climate-carbon cycle feedbacks in a multi-model framework is a promising way to improve climate models to be used in future international assessments of climate change.

Appendix A

A1 Temperature reconstruction data (Figs. 2a and 3)

We used reconstruction data featured in Fig. 6.10 of the Intergovernmental Panel on Climate Change Fourth Assessment Report (Jansen et al., 2007). Data time series and the shaded representation of overlap (consensus) (main text Fig. 2a) were obtained from http://www.cru.uea.ac.uk/datapages/ipccar4.htm. Acronyms used in Fig. 3 of the main text are identical with those in the IPCC AR4 (Jansen et al., 2007) Fig. 6.10: JBB1998 (Jones et al., 1998), MBH1999 (Mann et al., 1999), ECS2002 (Esper et al., 2002).
A2 CO₂ reconstructions (Fig. 5)

The grey shading in Fig. 5 of the main text represents a compilation of atmospheric CO₂ values together with their uncertainties into a “certainty” map by taking into account uncertainties in measured CO₂ values and dating. It is based on data from the Antarctic ice cores Law Dome (MacFarling Meure et al., 2006a, b; Etheridge et al., 2006; Levchenko et al., 1996, 1997), EPICA Dome C (Siegenthaler et al., 2005; Munnin et al., 2004a, b), Dronning Maud Land (Siegenthaler et al., 2005a, b) and South Pole (Siegenthaler et al., 2005b). Each measurement value is associated with a Gaussian distribution over the combined space of CO₂ values and time. The Gaussians are constructed such that they are centered at the particular measurement values, and the uncertainties in value and time are used as their standard deviations. Considering one data point, in this way, a value far off in CO₂ value and/or time gets a smaller weight than a value close-by, and this reflects decreasing certainty with more distance. The certainty map is constructed from the individual Gaussians by summing up their contributions for a grid of points in the CO₂/time plane and normalizing along a line of constant time to 1; this normalization reflects the fact that CO₂ must have a value at every time. In this way, points close to 1 (dark) are more certain than points close to zero (light).

Appendix B

Calculation of radiative forcing by the external drivers

All radiative forcings (Fig. 2c) are calculated off line with the ECHAM5 isolated radiative transfer code following the Wetherald and Manabe (1998) approach for calculating...
radiative feedbacks. The radiative forcing at the top of the atmosphere is defined as the change in radiative fluxes at the top of the atmosphere due to the change in one single variable $x$. All other variables are taken from one reference year and do not change from year to year. The radiative fluxes are calculated every six hours and are averaged globally and annually. This method allows accurate forcing separation. The forcings defined here are all instantaneous and do not allow any atmospheric adjustment. For CO$_2$ concentration, the forcing is calculated in bins of CO$_2$ changes with respect to the average CO$_2$ concentrations of the control run (280.08 ppm). Land-cover-change related radiative forcing reflects only the effect of changing surface albedo and is calculated relative to the period from 800 to 850 from the experiment with land-cover-changes as the only forcing.

Appendix C

Estimation of trends in the forced simulations (Fig. 4)

For a given length of the estimation period, the probability density function (pdf) derived from the linear trends represents the variability of the unforced system. The procedure is repeated for segment lengths between 1 and 100 years to derive upper and lower confidence limits. Black lines in Fig. C1 indicate the 5th and 95th percentile range. Thereafter, for specific times in the forced experiments, linear trends are compared with the confidence limits for the respective segment lengths. As an example, Fig. C1 displays trend estimates for forced experiments starting in the year 1400 AD. Trends (red and blue symbols) are calculated and displayed as a function of segment lengths in comparison with the range obtained from the control run pdfs (black lines). In these particular cases, trends between approximately 15 and 40 years (experiment E1_5), and 20 and 55 years (experiment E2_3) can be said to be significantly different from that expected on the basis of internal variability, with a random probability of occurrence of less then 0.05. To construct Fig. 4 of the main text, this analysis was then carried out...
for starting dates throughout the millennium simulations with a 10-year cadence (800, 810, …, 1900) for all ensemble members.

Appendix D

Co-variability of global mean temperature and CO₂

In constructing the scatter plot and regression lines (Fig. 7a of the main text), 3000 years of data from the control run and the first 800 years of the three ensemble members of the forced simulations were analyzed. Annually and globally averaged data was analyzed in the frequency domain using 200 year windows with a 50 year overlap, resulting in 15 analysis windows for the forced simulations and 19 windows for the control simulations. The co-spectral data for the ensemble average of the 15, respectively 19, analysis periods was analyzed. Low pass filtering with a sharp spectral window was performed on detrended and tapered time-series. The filter eliminated wave numbers larger than 3. Tapering was performed over 5% of the time-series, but results were not sensitive to the length of the tapering period. The correlation analysis was performed for a five year lag of the CO₂ time-series, but similar results are obtained for halving or doubling the lag.

To construct the regressions the data were randomly sampled with a mean stride of 25±5 years. The sample size was estimated following the methods of Zwiers and von Storch (1995). Regression slopes were calculated for different realizations. The realizations whose slopes matched the ensemble average were chosen for presentation in Fig. 7. Regression slopes varied by ±0.07 for the control and 0.06 ppm for the forced simulations. A similar analysis was performed for the 5-member ensemble of simulations with weak solar variability, these produced results intermediate between those of the control and strongly forced simulations on which our presentation focuses.
Acknowledgements. This work has been carried out as part of the MPI-M Integrated Project Millennium and contributes to the Cluster of Excellence CLISAP at Hamburg University. Funding was provided by the ENIGMA project of the Max Planck Society (SJL), the DFG CAWSES Priority Program (NAK, LEV), the DFG Collaborative Research Center 574 (CT), and the WCU program of the Korean Ministry of Education, Science and Technology (SKS). The model simulations were carried out on the supercomputing system of the German Climate Computation Centre (DKRZ) in Hamburg. Data processing and storage were provided by the Model & Data group at MPI-M.

The service charges for this open access publication have been covered by the Max Planck Society.

References


Brovkin, V., Lorenz, S. J., Jungclaus, J. H., Raddatz, T., Timmreck, C., Reick, C., Segschneider,


Goosse, H., Renssen, H., Timmermann, A., and Bradley, R. S.: Internal and forced climate variability during the last millennium: A model-data comparison using ensemble simulations,


Levchenko, V., Francey, R. J., Etheridge, D. M., Tuniz, C., Head, J., Morgan, I., Lawson, E.,


core high resolution Holocene and transition CO₂ data, Technical report, IGBP PAGES/World Data Center for Paleoclimatology, OAA/NGDC Paleoclimatology Program, Boulder CO, USA, 2004b.


1033
Fig. 1. NH temperature and CO$_2$ concentration in the 20th century. (a) 20th century Northern Hemisphere (land and ocean) 2m air temperatures (11-year running means) simulated in ensemble E1 (solid coloured lines) and E2 (dashed coloured lines) in comparison with the HadCRUT3v dataset (obtained from the Climatic Research Unit, http://www.cru.uea.ac.uk/cru/data/temperature), (b) 20th century global CO$_2$ concentration (yearly data) simulated in ensemble E1 (red) and E2 (blue) in comparison with the Mauna Loa data set (black) (obtained from the Carbon Dioxide Information Analysis Center, http://cdiac.ornl.gov). The green line is the respective curve for the land-cover-change-only experiment.
Fig. 2.

a) Evolution of simulated temperature over the last 1200 years, and the prescribed forcings: (a) Northern Hemisphere 2m land temperature anomalies w.r.t. the 1961-1990 mean for ensembles E1 (red) and E2 (blue) in comparison with the range of reconstructions (gray scale, redrawn from Jansen et al. (2007), see Appendix A). Black horizontal lines indicate the control experiment mean and its 5th - 95th percentile range. Green horizontal bars indicate periods where the ensembles do not overlap. Time series are smoothed by a 31-yr running mean. Crosses at the right axis denote the ensemble means (annual average) at the end of the simulation (2005).

b) Ensemble spread for the ensembles E1 (red), and E2 (blue).

c) Radiative forcing at the top of the atmosphere. Anomalies from solar irradiance and CO$_2$ variations are calculated w.r.t. their pre-industrial control mean (1367 Wm$^{-2}$ and 280.02 ppm, respectively). The radiative forcing from volcanic aerosol injections and land-cover changes are calculated from the single forcing experiments (see Appendix B). All time series are smoothed by a 31-yr running mean. Crosses at the right axis denote the respective values at the end of the simulation (2005).
**Fig. 2.** Evolution of simulated temperature over the last 1200 years, and the prescribed forcings: (a) Northern Hemisphere 2 m land temperature anomalies w.r.t. the 1961–1990 mean for ensembles E1 (red) and E2 (blue) in comparison with the range of reconstructions (gray scale, redrawn from Jansen et al. (2007), see Appendix A). Black horizontal lines indicate the control experiment mean and its 5th–95th percentile range. Green horizontal bars indicate periods where the ensembles do not overlap. Time series are smoothed by a 31-yr running mean. Crosses at the right axis denote the ensemble means (annual average) at the end of the simulation (2005), (b) ensemble spread for the ensembles E1 (red), and E2 (blue), (c) Radiative forcing at the top of the atmosphere. Anomalies from solar irradiance and CO₂ variations are calculated w.r.t. their pre-industrial control mean (1367 W m⁻² and 280.02 ppm, respectively). The radiative forcing from volcanic aerosol injections and land-cover-changes are calculated from the single forcing experiments (see Appendix B). All time series are smoothed by a 31-yr running mean. Crosses at the right axis denote the respective values at the end of the simulation (2005).
Fig. 3. Northern Hemisphere temperature difference between the coldest 30 year period during the LIA (1550–1750) and the warmest 30 year period during the MWP (900–1300) from the ensembles E1 (red) and E2 (blue), respectively, and the reconstruction data available from Jansen et al. (2007) (green). As an estimate for internal variability, black symbols denote the 5th–95th percentile range of the respective 30-year means from the 3000-year control experiment. Symbols denote different choices of regional and temporal averaging as motivated by the available reconstructions. Acronyms refer to those used in Jansen et al. (2007) (see Appendix A).
**Fig. 4.** Analysis of trends in the full-forcing experiments in the ensembles E1 (a) and E2 (b). The analysis identifies periods during which warming (shades of red) and cooling (shades of blue) trends (starting dates at the x-axis) exceed the 5th–95th percentile range of the respective trends in the control experiment. The intensity of the colour denotes the number of ensemble members showing the same behaviour so that, for example, the darkest shade of red indicates that all ensemble members show warming trends significantly different from the control run.
Fig. 5. CO$_2$ concentrations (31-year running mean) from (a) ensembles E1 (red) and E2 (blue) in comparison with a compilation of ice core reconstructions (grey shading, see Appendix A). Black horizontal lines denote the control experiment mean and its 5th–95th percentile range, and also indicated is the CO$_2$ evolution, (b) the respective CO$_2$ concentrations from the experiments forced by one single component, i.e. standard solar forcing (red), strong solar forcing (blue), land-cover change (green), and volcanic aerosols (light-blue).
**Fig. 6.** Carbon fluxes (Gigatons Carbon per year) into the atmosphere for the ensembles E1 (red) and E2 (blue) from the ocean (solid lines) and the land biosphere (dashed lines).
Fig. 7.

(a) Co-variability of globally and annually averaged 2m temperature anomalies with globally and annually averaged CO$_2$ anomalies lagged by five years for the E2 ensemble (blue open circles) and control (filled grey circles) experiments. Data points are taken by randomly sampling the low-pass (50-year) filtered data with a mean sample stride of 25 years. Correlations are significant at greater than the 99% level given an equivalent sample size of 68 and 140 for the strongly forced and controlled experiments, respectively (see Appendix D),

(b) Ratio of power spectra as a function of wave-number for both the strongly forced (blue open circles) and control (filled grey circles) simulations. In both panels the forced experiments were only analyzed for the period between 800 and 1600, during which time the anthropogenic influence on the carbon cycle was negligible,

(c) and (d) running regressions (slopes in ppm/K) between globally and annually averaged 2m temperature anomalies and globally and annually averaged CO$_2$ anomalies for different time lags (positive lags mean that temperature is leading) for the control experiment (right) and one of the E2 experiments (left). Running regressions were performed for 200-yr chunks based on the 31-yr running mean time series.
**Fig. 7.** Climate carbon-cycle sensitivity: (a) Co-variability of globally and annually averaged 2 m temperature anomalies with globally and annually averaged CO₂ anomalies lagged by five years for the E2 ensemble (blue open circles) and control (filled grey circles) experiments. Data points are taken by randomly sampling the low-pass (50-year) filtered data with a mean sample stride of 25 years. Correlations are significant at greater than the 99% level given an equivalent sample size of 68 and 140 for the strongly forced and controlled experiments, respectively (see Appendix D), (b) Ratio of power spectra as a function of wave-number for both the strongly forced (blue open circles) and control (filled grey circles) simulations. In both panels the forced experiments were only analyzed for the period between 800 and 1600, during which time the anthropogenic influence on the carbon cycle was negligible, (c), and (d) running regressions (slopes in ppm/K) between globally and annually averaged 2 m temperature anomalies and globally and annually averaged CO₂ anomalies for different time lags (positive lags mean that temperature is leading) for the control experiment (right) and one of the E2 experiments (left). Running regressions were performed for 200-yr chunks based on the 31-yr running mean time series.
Fig. C1. Detection of significant trends in the time series for a particular start date (1400 AD): black lines indicate the 5th–95th percentile of the various trend lengths derived from the 3000-year control run. Coloured lines are trends from NH temperature time series from two forced runs (red: expt. E1_5 (weak solar), blue: expt. E2_3 (strong solar)).