Modelling the climate of the last millennium: What causes the differences between simulations?

H. Goosse,[1] T. J. Crowley,[2] E. Zorita,[3] C. M. Ammann,[4] H. Renssen,[5] and E. Driesschaert[1]

Received 6 January 2005; accepted 8 March 2005; published 30 March 2005.

[1] An ensemble of simulations performed with a coarse resolution 3-D climate model driven by various combinations of external forcing is used to investigate possible causes for differences noticed in two recent simulations of the climate of the past millennium using General Circulation Models (GCMs). Our results strongly suggest that differences in sensitivity (equilibrium and transient climate response) could be responsible for temperature changes that differ by more than a factor of two between two models. In addition, the spin-up procedure could explain some differences between the simulations during the first centuries of the second millennium. The choice of the forcing reconstruction is found to play a smaller role for the differences in the simulated climate, in the model configurations analyzed here. Furthermore, at decadal scale, internal climate variability can mask the differences associated with different forcing reconstructions. **Citation:** Goosse, H., T. J. Crowley, E. Zorita, C. M. Ammann, H. Renssen, and E. Driesschaert (2005), Modelling the climate of the last millennium: What causes the differences between simulations?, *Geophys. Res. Lett.*, *32*, L06710, doi:10.1029/2005GL022368.

1. Introduction

[2] Several simulations covering the last 1000 to 2000 years, driven by both natural and anthropogenic forcings, have been recently performed [e.g., *Crowley*, 2000; *Bertrand et al.*, 2002; *Gerber et al.*, 2003; *Bauer et al.*, 2003; *González-Rouco et al.*, 2003; *Jones and Mann*, 2004; *Goosse et al.*, 2005]. As illustrated in the recent review of *Jones and Mann* [2004], the results of those simulations display some general similarities but also large differences. The latter could have four causes. Firstly, the experiments have been performed using different reconstructions of the past radiative forcing. Secondly, as the models are different, their response to an external forcing could be different. This is often measured by the equilibrium climate sensitivity of the model and the related Transient Climate Response (TCR) that, in addition, takes into account the ocean heat uptake efficiency. The climate sensitivity is conventionally defined as the temperature change at steady state in response to a doubling of the CO_2 concentration in the atmosphere while TCR is usually estimated as the transient response of the climate system at the time of CO_2 doubling in experiments in which CO_2 concentration is increased by 1% per year (for more information, see, e.g., *Raper et al.* [2002] and *Gregory et al.* [2004]). Thirdly, because of the lack of data, it is not possible to specify the initial conditions for the model simulations from observations. Using different procedures to start the experiments could thus have an impact on the model results, at least initially. Finally, the most comprehensive models include their own representation of the internal variability of the climate system at many time scales. As a consequence, the two simulations performed with the same model and driven by the same external forcing may differ randomly because of internal model variability.

[3] The goal of the present study is to provide insight into which of these differences might contribute most significantly to the discrepancy in the simulations. To do so, we analyze the variations of the annual mean temperature averaged over the Northern Hemisphere, the variable most frequently discussed when presenting model results. As it is not feasible here to comprehensively analyze all the experiments that have been performed in any great detail, we will focus our attention on the two available 3-D climate model integrations covering the past millennium, using the comprehensive coupled Atmosphere Ocean GCMs ECHO-G [*González-Rouco et al.*, 2003] and NCAR-CSM (Ammann et al.; as described by *Jones and Mann* [2004]). Unfortunately, it is not presently possible to directly compare these GCMs through a large number of 1000-year comparison experiments (e.g. by switching the forcings). Consequently, the comparison is made here using as an intermediary step a series of experiments performed with a more efficient coupled 3-D climate model (ECBILT-CLIO-VECODE, see below) driven by various forcing combinations, including ones that mimic those used in the two GCMs.

[4] The results of those 3-D models have been compared to reconstructions [e.g., *González-Rouco et al.*, 2003; *Goosse et al.*, 2005; *Jones and Mann*, 2004]. Nevertheless, because of the uncertainties in the past radiative forcing and in the reconstructions themselves, it is not easy to interpret any difference between model results and observations. Therefore, we use here a different approach and deliberately restrict our analysis to model results as our goal is to understand the difference between models, not to identify the simulation that has the best agreement with paleoclimatic reconstruction.

2. Model Description and Experimental Design

[5] The simulations performed by the AOGCMs have been previously published (see above) and will not been

[1]Institut d'Astronomie et de Géophysique G. Lemaître, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.
[2]Department of Earth and Ocean Sciences, Duke University, Durham, North Carolina, USA.
[3]Institute for Coastal Research, GKSS Research Center, Germany.
[4]Climate and Global Dynamics Division - Paleoclimatology, National Center for Atmospheric Research, Boulder, Colorado, USA.
[5]Faculty of Earth and Life Sciences, Vrije Universiteit, Amsterdam, Netherlands.

0094-8276/05/2005GL022368$05.00

Table 1. Description of the Experiments

Symbol of the Group	Starting Date	Forcing	
		Solar	Volcanic
K	1001 AD	*Lean et al.* [1995]/ *Bard et al.* [2000]	*Crowley* [2000]
C	1001 AD	*Crowley* [2000]	*Crowley* [2000]
D	1 AD	*Crowley*[a] 2004	*Crowley*[a] 2004
M	851 AD	*Bard et al.* [2000]	Ammann et al.[b]
Z	1001 AD	*Crowley*[c] (2000)	*Crowley* [2000]

[a]Updated from *Crowley et al.* [2003].
[b]As described by *Jones and Mann* [2004].
[c]Experiments of group Z are driven by the solar irradiance of *Crowley* [2000] but using a different scaling factor compared to the ones of the group C.

described here in detail. Our version of ECBILT-CLIO-VECODE is identical to the one of *Renssen et al.* [2005] and *Goosse et al.* [2005]. The atmospheric component is ECBILT2 [*Opsteegh et al.*, 1998], a T21, 3-level quasi-geostrophic model. The oceanic component is CLIO3 [*Goosse and Fichefet*, 1999] that is made up of an ocean general circulation model coupled to a comprehensive thermodynamic–dynamic sea ice model. ECBILT–CLIO is coupled to VECODE, a dynamic global vegetation model that simulates dynamics of two main terrestrial plant functional types, trees and grasses, as well as desert [*Brovkin et al.*, 2002]. More information about the model and a complete list of references is available at http://www.knmi.nl/onderzk/CKO/ecbilt-papers.html.

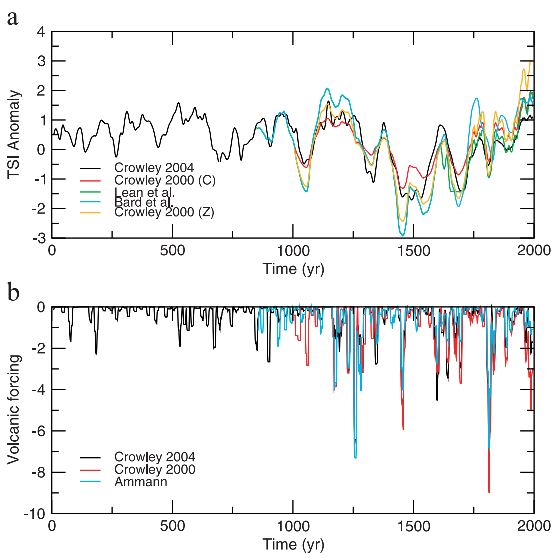

Figure 1. (a) Time variations of solar irradiance at the top of the atmosphere following the reconstructions of *Lean et al.* [1995] (green), *Bard et al.* [2000] (turquoise), Crowley-2004 (black, updated from *Crowley et al.* [2003]) and two scalings derived from *Crowley* [2000] used in the experiments of group C (red) and Z (orange), respectively. (b) Time variations of volcanic forcing scaled as an effective change in solar irradiance for comparison with (a), following the reconstructions of *Crowley* [2000] (red), Crowley-2004 (updated from *Crowley et al.* [2003]) (black) and Ammann (as described by *Jones and Mann* [2004], turquoise). A 10-year running mean has been applied to the time series in order to highlight low frequency variations.

[6] Using this model, we have performed a group of simulations over last 1000 to 2000 years (Table 1). Each set of individual simulation is driven by a reconstruction of the past radiative forcing due to volcanic eruptions and changes in solar irradiance (Figure 1). The forcing due to variations of orbital parameters follows *Berger* [1978] and the observed evolution of greenhouse gases (based on a compilation of ice cores measurements; J. Flueckiger, personal communication, 2004) is imposed over the whole simulated period, in all the experiments. Furthermore, in the simulations of groups K, C, D, M the influence of sulfate aerosols due to anthropogenic activity is taken into account during the period 1850–2000 AD through a modification of surface albedo [*Charlson et al.*, 1991] and the forcing due to change in land-use is applied [*Ramankutty and Foley*, 1999]. For each group, an ensemble of at least 15 simulations is performed. The elements of an ensemble differ only in their initial conditions that were taken from model states in previous experiments covering the last millennium, each state selected being separated from the next one by 150 years.

3. Results

[7] The ensemble mean of decadally averaged surface temperature averaged over the Northern Hemisphere is displayed on Figure 2 for all the groups of simulation. In the present study, we are using the period 1200–1850 AD as a reference period for all the simulations. This reference period was chosen to eliminate possible problems related to initial conditions (e.g., years 1000–1200) and to the strong differences of specified anthropogenic forcing in the 20th century. At first sight, the differences between the simulations appear quite modest, in particular if compared to *Bertrand et al.* [2002], who have studied the response to a wide range of possible forcings for the last millennium. This is to a large extent related to the forcings selected. Compared to *Bertrand et al.* [2002], the reconstructions that present the largest variations of solar irradiance are not included here, as they have not been used by GCMs to simulate the climate of the last millennia.

[8] The difference between the ensemble mean of the groups of simulations is generally smaller than the scatter

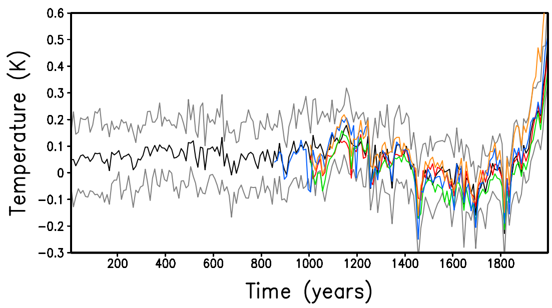

Figure 2. Time evolution of the anomaly of decadal mean temperature averaged over the Northern Hemisphere during the last 2000 years for the ensemble mean of experiments using different external forcings: groups D(black), C(red), K(green), M(blue), Z(orange). The grey lines represent the ensemble mean plus and minus 2 standard deviation for experiments of group D. The reference period is 1200–1850 AD.

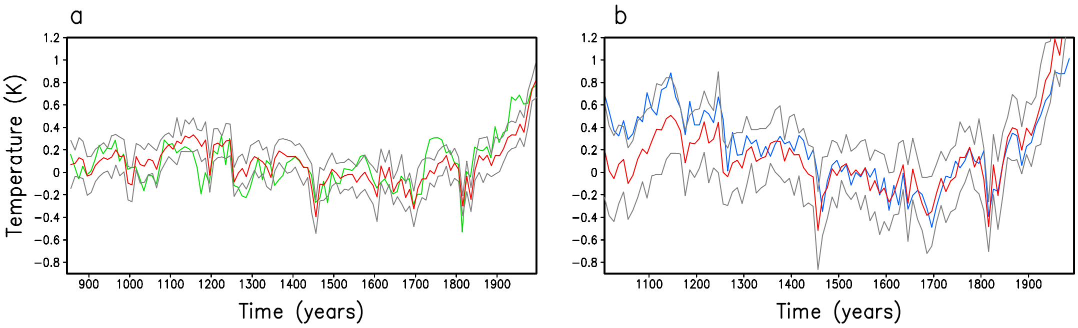

Figure 3. Time evolution of the anomaly of decadal mean temperature averaged over the Northern Hemisphere in (a) experiments of group M (red) compared to the simulations performed by Ammann el al. using CSM (green) and (b) in experiments of group Z (red) compared to *González-Rouco et al.* [2003] (blue). The grey lines represent the ensemble mean plus and minus 2 standard deviation for ECBILT-CLIO-VECODE. In those figures, the results of ECBILT-CLIO-VECODE have been scaled in order to have the same standard deviation as the GCMs over the period 1200–1850.

within an ensemble set for 10-yr averages caused by internal variability. This is illustrated in Figure 2 where all ensemble means of the different simulations remain within the interannual range determined by the long D simulation. The only exception is found for group Z which shows a much larger warming during the last 150 years. This behavior is in direct response to the absence of the negative radiative forcing due to anthropogenic sulfate aerosols and land use changes. The influence of the forcing could thus be masked by internal variability at decadal scale in our model. For longer periods, the systematic impact of the forcing on the climate is still at best only marginally detectable [e.g., *Hegerl et al.*, 2003]. However, the role of the forcing can be recognized when analyzing the variability of the whole simulations. Averaged over the experiments of group C, the standard deviation of the temperature for the reference period 1200–1850 is 0.08 K while values up to 0.11 K are found for group M. For group K, D and Z, the standard deviation is 0.09 K for each group.

[9] Experiments of group Z could be directly compared to the ones performed with ECHO-G because the external forcing, except for orbital forcing, is the same in both cases. Group M is designed to be similar to the experiment performed using CSM but the latter simulation includes the forcing due to volcanic and anthropogenic sulfate aerosols in a much more sophisticated way than ECBILT-CLIO-VECODE. Furthermore, the CSM experiment does not take into account land-use change. As a consequence, the comparison of our results with those of CSM should be qualitative, in particular for the last 150 years.

[10] The experiments performed with the GCMs display larger temperature variability than the corresponding ones using ECBILT-CLIO-VECODE. As a consequence, for the comparison of the time series, the results of experiments Z and M have been scaled to have the same standard deviation over the reference period 1200–1850 AD as ECHO-G and CSM, respectively (Figure 3). This corresponds to a factor 1.6 for group M (0.17 K, the standard deviation of CSM experiment, divided by 0.11 K for group M) and 2.7 for group Z (0.24 K for ECHO-G divided by 0.09 K for group Z). The standard deviation of the ensemble (grey curves) has also been scaled using the standard deviation of decadal mean temperature obtained in long control experiments performed without changes in external forcing, which provides a measure of internal variability in the various models. This leads to a scaling of the standard deviation of group M by a factor of 1.2 (=0.07 K/0.06 K) and for group Z to 2.2 (=0.13 K/0.06 K). This shows that, at decadal scale, the ratio of the standard deviation associated with total variability and the one associated with internal variability is higher in the GCMs than in our model, with values of about 2.5 for CSM (0.17/0.07), 1.8 for ECHO-G (0.24/0.13), and less than 1.8 for ECBILT-CLIO-VECODE (0.11/0.06). Nevertheless, an underestimation of the influence of the forced variability of this magnitude would not change qualitatively our results about the role of internal and forced variability deduced from Figure 2.

[11] Undeniably, the scaling factors reflects in a very crude way the different climate response characteristics of the models with equilibrium sensitivity (1.8K for ECBILT-CLIO, ~2K for CSM and ~3.2 for ECHO-G) and TCR (0.9K for ECBILT-CLIO, ~1.4K for CSM, not available for ECHO-G) providing only elements for the comparison. Unfortunately, not enough information is available to us to estimate the contribution of the various processes. Furthermore, for the last 100 years of the experiments, a scaling factor of 2.7 for group Z appears too large. It is an illustration of the limitations of our scaling as, for instance, a model may respond differently to each particular forcing [e.g., *Gregory et al.*, 2004].

[12] However, this comparison of ECBILT-CLIO-VECODE results with those of the GCMs strongly suggests that the differences between the GCMs' simulations lie to a large extent in the differences between the models transient sensitivities themselves. Indeed, group Z using a forcing similar to ECHO-G has a standard deviation 16% smaller during pre-industrial times than group M using a similar forcing as CSM. Nevertheless, the impact of the differences in the forcing, as isolated by ECBILT-CLIO-VECODE experiments, is overwhelmed by the difference in climate sensitivities (~1.6 times higher in ECHO-G than in CSM). Overall, this leads to a standard deviation that is 1.4 higher is the ECHO-G experiment during pre-industrial times than in the one with CSM.

[13] The temperature simulated by ECHO-G is also higher than the ensemble mean of group Z over the period 1000–1200 AD after scaling. In ECBILT-CLIO-VECODE, the ensemble mean follows qualitatively the forcing, as expected. The solar irradiance is generally high during the first centuries of the second millennium AD, with values

similar to the ones during the 20th century. Nevertheless, because of the forcing due to the increase in the greenhouse gas concentrations, the temperature in ECBILT-CLIO-VECODE is much higher during the second half of the 20th century. On the other hand, in the ECHO-G simulation, the temperature around 1150 AD is nearly as high as the one at the end of the experiment. As the differences between group Z and ECHO-G are systematic over the whole period 1000 AD–1250 AD, it is thus very unlikely that the relatively high temperatures during this period in ECHO-G are due to internal variability alone. A more likely cause of the difference between the results of the two models lies thus in the design of the initial conditions or in the spin up procedure. Testing more precisely this hypothesis is out of the scope of the present study but this is in agreement with the preliminary results of an additional experiment performed with ECHO-G using different initial conditions that shows lower temperatures over the period 1000–1200 AD than the one used here.

4. Conclusions

[14] First of all, our results must be taken with caution, as we are comparing our 15 member ensembles of simulations to single realizations of the more complex models. Furthermore, our conclusions are based on a particular model and are thus dependent on its nominal sensitivity, on its response speed to an external forcing and the way internal variability is reproduced.

[15] The forcing is to a large extent responsible for the general cooling between the 12th century and the 19th century and for the large warming during the 20th century. Nevertheless, the forcings used in the two recent GCM simulations are relatively similar during pre-industrial times. According to our results, using different forcings is thus not the dominant cause of the differences between those simulations since it is associated with changes well within the range of the internal variability of the model at decadal scale. However, we should recall that the forcing could have a larger role in explaining the difference between simulations if significantly stronger or weaker forcing were used in some GCM experiments.

[16] According to our results, the difference in the results of CSM and ECHO-G simulations is mainly due to their different climate sensitivities. Furthermore, the design of the spin-up procedure could likely explain some differences between the simulations during the early part of the millennium (ECHO-G was started in year 1000, CSM in 850 AD). Thus, an accurate assessment of the temperature during the relatively warm period early in the last millennium requires experiments starting significantly earlier than 1000 AD.

[17] **Acknowledgments.** This study is supported by the Federal Science Policy Office (Belgium), contracts EV/10/7D and EV/10/9A and the Action Concertée Incitative Changement Climatique from the French Ministry of Research. H. Goosse is research associate with the Belgian National Fund for Scientific Research. H. Renssen is supported by the Netherlands Organization for Scientific Research. The work by E. Zorita is part of the EU project SO&P.

References

Bard, E., et al. (2000), Solar irradiance during the last 1200 years based on cosmogenic nuclides, *Tellus, Ser. B*, *52*, 985–992.
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*(6), 1276, doi:10.1029/2002GL016639.
Berger, A. L. (1978), Long-term variations of daily insolation and Quaternary climatic changes, *J. Atmos. Sci.*, *35*, 2363–2367.
Bertrand, C., et al. (2002), Climate of the last millennium: A sensitivity study, *Tellus, Ser. A*, *54*, 221–244.
Brovkin, V., J. Bendtsen, M. Claussen, A. Ganopolski, C. Kubatzki, V. Petoukhov, and A. Andreev (2002), Carbon cycle, vegetation, and climate dynamics in the Holocene: Experiments with the CLIMBER-2 model, *Global Biogeochem. Cycles*, *16*(4), 1139, doi:10.1029/2001GB001662.
Charlson, R. J., et al. (1991), Perturbation of the Northern Hemisphere radiative balance by backscattering from anthropogenic sulfate aerosols, *Tellus, Ser. AB*, *43*, 152–163.
Crowley, T. J. (2000), Causes of climate change over the past 1000 years, *Science*, *289*, 270–277.
Crowley, T. J., S. K. Baum, K. Kim, G. C. Hegerl, and W. T. Hyde (2003), Modeling ocean heat content changes during the last millennium, *Geophys. Res. Lett.*, *30*(18), 1932, doi:10.1029/2003GL017801.
Gerber, S., et al. (2003), Constraining temperature variations over the last millennium by comparing simulated and observed atmospheric CO_2, *Clim. Dyn.*, *20*, 281–299.
González-Rouco, F., H. von Storch, and E. Zorita (2003), Deep soil temperature as proxy for surface air-temperature in a coupled model simulation of the last thousand years, *Geophys. Res. Lett.*, *30*(21), 2116, doi:10.1029/2003GL018264.
Goosse, H., and T. Fichefet (1999), Importance of ice-ocean interactions for the global ocean circulation: A model study, *J. Geophys. Res.*, *104*, 23,337–23,355.
Goosse, H., et al. (2005), Natural and forced climate variability during the last millennium: A model-data comparison using ensemble simulations, *Quat. Sci. Rev.*, in press.
Gregory, J. M., W. J. Ingram, M. A. Palmer, G. S. Jones, P. A. Stott, R. B. Thorpe, J. A. Lowe, T. C. Johns, and K. D. Williams (2004), A new method for diagnosing radiative forcing and climate sensitivity, *Geophys. Res. Lett.*, *31*, L03205, doi:10.1029/2003GL018747.
Hegerl, G. C., T. J. Crowley, S. K. Baum, K. Kim, and W. T. Hyde (2003), Detection of volcanic, solar and greenhouse gas signals in paleo-reconstructions of Northern Hemisphere temperature, *Geophys. Res. Lett.*, *30*(5), 1242, doi:10.1029/2002GL016635.
Jones, P. D., and M. E. Mann (2004), Climate over past millennia, *Rev. Geophys.*, *42*, RG2002, doi:10.1029/2003RG000143.
Lean, J., J. Beer, and R. Bradley (1995), Reconstruction of solar irradiance since 1610: Implications for climate change, *Geophys. Res. Lett.*, *22*, 1591–1594.
Opsteegh, J. D., et al. (1998), ECBILT: A dynamic alternative to mixed boundary conditions in ocean models, *Tellus, Ser. A*, *50*, 348–367.
Ramankutty, N., and J. A. Foley (1999), Estimating historical changes in global land cover: Croplands from 1700 to 1992, *Global. Biogeochem. Cycles*, *13*, 997–1027.
Raper, S. C. B., J. M. Gregory, and J. M. Stouffer (2002), The role of climate sensitivity and ocean heat uptake on AOGCM transient temperature response, *J. Clim.*, *15*, 124–130.
Renssen, H., et al. (2005), Simulating the Holocene climate evolution at northern high latitudes using a coupled atmosphere-sea ice-ocean-vegetation model, *Clim. Dyn.*, *24*, 57–69.

C. M. Ammann, Climate and Global Dynamics Division-Paleoclimatology, National Center for Atmospheric Research, Boulder, CO 80307-3000, USA.

T. J. Crowley, Department of Earth and Ocean Sciences, Duke University, Durham, NC 27708, USA.

E. Driesschaert and H. Goosse, Institut d'Astronomie et de Géophysique G. Lemaître, Chemin du Cyclotron, 2, Université Catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium. (hgs@astr.ucl.ac.be)

H. Renssen, Faculty of Earth and Life Sciences, Vrije Universiteit, NL-1081HV Amsterdam, Netherlands.

E. Zorita, Institute for Coastal Research, GKSS Research Center, D-21502, Geesthacht, Germany.