Data assimilation over the last millennium using ensemble techniques

Goosse, H.; Mann, M.E.; Renssen, H.; Timmerman, A.

published in
PAGES News
2008

Data assimilation over the last millennium using ensemble techniques

Hugues Goosse[1], M.E. Mann[2], H. Renssen[3] and A. Timmermann[4]

[1]Université Catholique de Louvain, Institut d'Astronomie et de Géophysique G. Lemaître, Belgium; hgs@astr.ucl.ac.be
[2]Department of Meteorology, and Earth and Environmental Systems Institute (EESI), Pennsylvania State University, USA
[3]Faculty of Earth and Life Sciences, Vrije Universiteit Amsterdam, The Netherlands
[4]IPRC, SOEST, University of Hawaii, Honolulu, USA

Over the last few years, a large number of new proxies, new large-scale reconstructions and model simulations covering the last millennium have become available (e.g., Mann and Jones, 2003; Jansen et al., 2007). They consistently show that at the hemispheric (or quasi-hemispheric) scale, the most recent decades are likely the warmest of at least the past millennium. Moreover, the recent warming can only be simulated in models if the effect of the increase in greenhouse gases concentration in the atmosphere due to human activities is taken into account.

However, our understanding of climate change in the period before this clear anthropogenic influence is still fragmentary. One reason for this is the relatively small magnitude of the signal recorded over this period, compared with, for instance, the changes observed during the Last Glacial Maximum. A second element is the spatial heterogeneity of the signal, with climates in disparate regions evolving very differently over time. Part of this regional variability could be related to changes in atmospheric and oceanic circulation. For instance, it has been proposed that the North Atlantic Oscillation or the El Niño Southern Oscillation may be influenced by solar and volcanic forcing, leading to a spatial response of the system similar to the characteristic patterns associated with those modes of variability (e.g., Robock, 2000; Shindell et al., 2001, 2003; Mann et al., 2005). On the other hand, a significant part of the variability is purely related to the internal dynamics of the climate system and could not be linked to any modification in the external forcing. At the regional scale, particularly in mid and high latitudes, this internal variability is generally the dominant cause of the observed temperature changes before the 20th century (e.g., Goosse et al., 2005; Tett et al., 2007).

As long as a model contains the correct basic atmospheric and oceanic physics, it can simulate the forced variability. In this framework, models can produce patterns similar to those observed when averaged over a relatively large number of events or years. This could then be tested by comparing simulated and observed changes during the years following several major volcanic eruptions, or on average over long periods characterized by high (or low) total solar irradiance.

Models cannot be expected to precisely reproduce the time evolution of the true climate due to the large and potentially chaotic internal variability. Comparing regional details between model and observations for a single simulation is thus unlikely to yield meaningful insights, even if averaged over multiple decades before the anthropogenic warming of the late 20th century. One solution is to perform an ensemble of simulations over the last millennium in order to provide a reasonable estimate of the range of anomalies potentially associated with internal variability. The consistency between model results and proxy records can then be established by verifying that the observed evolution is well within the range obtained from the ensemble of simulations.

An alternative is to use data assimilation to optimally combine information from proxy data and model simulations, and obtain a consistent picture of past climate changes. Such an approach coerces the model towards agreement with the proxy records, taking into account the potential uncertainties in both model results and proxy data. In this approach, a realization of the model's internal variability is selected that is closest to the actual climate state, as inferred from the proxy data. The result of the data assimilation exercise over the last millennium is then a reconstruction of past changes, taking into account all the information available from proxy-data, model results and reconstructions of the changes in external forcings.

Despite the potential merit of this approach, only a few attempts at such data assimilation have thus far been made, perhaps because of the technical challenges involved. Simulations have been performed in which the model state is nudged, using a sophisticated statistical approach, to a pattern reconstructed from available observations (e.g., Jones and Widmann, 2003; van der Schrier and Barkmeijer, 2005). One advantage to such a technique is its computational efficiency. Furthermore, the spatial pattern of climate change is constrained at all locations of the model's domain. On the other hand, this approach requires one to have first statistically reconstructed from proxy data the pattern toward which the model state will be nudged.

Another option is to use ensemble techniques, where the information provided by the ensemble of simulations is combined with the one provided by the proxies at the time considered. In this framework, the simplest method is probably to first perform a large ensemble of simulations (of the order of 100 or more) over a particular year or period. The dis-

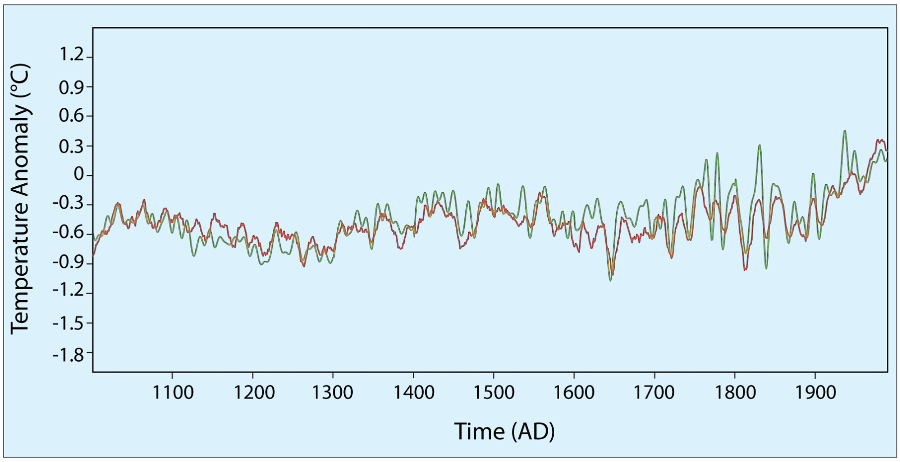

*Figure 1: Anomaly of annual mean temperature averaged over the area 0-30°E, 55-65°N, corresponding to northern Europe, from 1100 to 1900 AD. The **green line** is the average of several proxy records in this area (Mann et al., in prep.) and the **red line** is the result of a simulation with data assimilation using ensemble technique in the LOVECLIM model (www.astr.ucl.ac.be/index.php?page=LOVECLIM%40Description). An 11-year running mean has been applied to both time series. At this scale, the agreement between proxy record and the simulation using data assimilation is very good. For comparison, the range associated with internal variability in this area in the model, as measured by the standard deviation of an ensemble, is larger than 0.4 °C (Goosse et al., 2005), i.e., of the same order of magnitude as the difference between the warmest and coldest period during the pre-industrial period.*

tance between each member of the ensemble and the proxy record is measured by a cost function using reconstructed and simulated variables at the locations where the proxies are available. The best simulation, defined as the one that minimizes this cost function, is then selected as representative for this particular year or period, and used as the initial condition for the subsequent year. The procedure is repeated as many times as required in order to provide a reconstruction for the whole millennium (e.g., Collins, 2003; Goosse et al., 2006; Figure 1).

In this technique, the model is only constrained locally, at the locations where the proxy records are available. This is a clear advantage, compared to the nudging techniques, as the reconstructed spatial pattern is the result of the data assimilation procedure itself and is thus independent of any statistical method used to reconstruct patterns. A downside of the approach is the potential systematic error incurred if the model is not able to reproduce observed teleconnections between different regions.

Preliminary proofs of concept using this ensemble method have demonstrated that it can efficiently yield a plausible large-scale reconstruction if only a small number of proxies are available, and yet can also reconstruct regional detail where the number of available proxy data are sufficiently large (e.g., Goosse et al., 2006).

On the basis of these successful preliminary results, ongoing work is underway to investigate in greater detail the mechanisms that may be responsible for the climate changes of the past millennium. Potential refinements of the approach include a more sophisticated treatment of the uncertainties in both the proxy data and simulation results. This is a challenging issue for a number of reasons. For example, the nature of the uncertainties in proxy-derived climate records are complex, involving complicated physical or biological responses, which may yield frequency-dependent loss of climate information. Further work is necessary to characterize these uncertainties and biases more fully. Furthermore, classical assimilation methods used in meteorology and oceanography cannot be transferred directly to the analysis of the past millennium. Fortunately, focused efforts in this area are now underway, as discussed at one recent workshop on "Data assimilation to study the climate of the past millennium" (www.astr.ucl.ac.be/index.php?page=Wokshop_assim). It is reasonable to expect that significant improvements in the techniques used for data assimilation over the last millennium will be achieved in the years ahead, yielding significantly refined estimates of past changes and a better understanding of the causes of those changes.

References

Collins, M., 2003: Quantitative use of paleo-proxy data in global circulation models, *Geophysical Research Abstracts*, **5**: 1-14.

Goosse, H., Renssen, H., Timmermann, A. and Bradley, R.S., 2005: Internal and forced climate variability during the last millennium: A model-data comparison using ensemble simulations, *Quaternary Science Reviews*, **24**: 1345-1360.

Goosse H., Renssen, H., Timmermann A., Bradley R.S. and Mann, M.E., 2006: Using paleoclimate proxy-data to select optimal realisations in an ensemble of simulations of the climate of the past millennium, *Climate Dynamics*, **27**: 165-184, DOI 10.1007/s00382-006-0128-6.

Jones, J. and Widmann, M., 2003: Reconstructing large-scale variability from paleoclimatic evidence by means of Data Assimilation Through Upscaling and Nudging (DATUN). In: Fischer et al. (Eds) *The KIHZ project: towards a synthesis of Holocene proxy data and Climate models*, Springer, Berlin, 171-193.

van der Schrier, G. and Barkmeijer J., 2005: Bjerknes' hypothesis on the coldness during 1790-1820 AD revisited, *Climate Dynamics*, **25**(5): 537-553.

For full references please consult:
www.pages-igbp.org/products/newsletter/ref2008_2.html

Facilitating proxy-data interpretation of abrupt climate events using model simulations

Ane P. Wiersma[1]*, D.M. Roche[1,2] and H. Renssen[1]

[1]Faculty of Earth and Life Sciences, Vrije Universiteit Amsterdam, The Netherlands; ane.wiersma@falw.vu.nl
[2]IPSL/Laboratoire des Sciences du Climat et de l'Environnement, Gif-sur-Yvette, France
*Present affiliation: Deltares Subsurface and Groundwater Systems, Utrecht, The Netherlands

Climate model results vs. proxy-based climate reconstructions

Comparing climate model results for abrupt climate events with proxy-based climate reconstructions is often hampered by the difference in spatial and temporal characteristics. Proxy-based climate reconstructions present a climate signal over a long period of time at a specific location. Due to bioturbation, sample size, aquifer buffering, diffusion, etc., the climate signal recorded at an arbitrary point in time is often the integrated climate signal over several decades. This produces a time series of the climate signal smoothed by a decadal-scale filter. Modeling results, on the other hand, are often visualized as the spatial distribution of the average temperature or precipitation over a decadal- or centennial-scale time-window (averaging time-window) relative to a control climate.

This way of visualizing climate-model results generally gives a good indication of the geographical distribution of the event and the relative magnitude in differ-

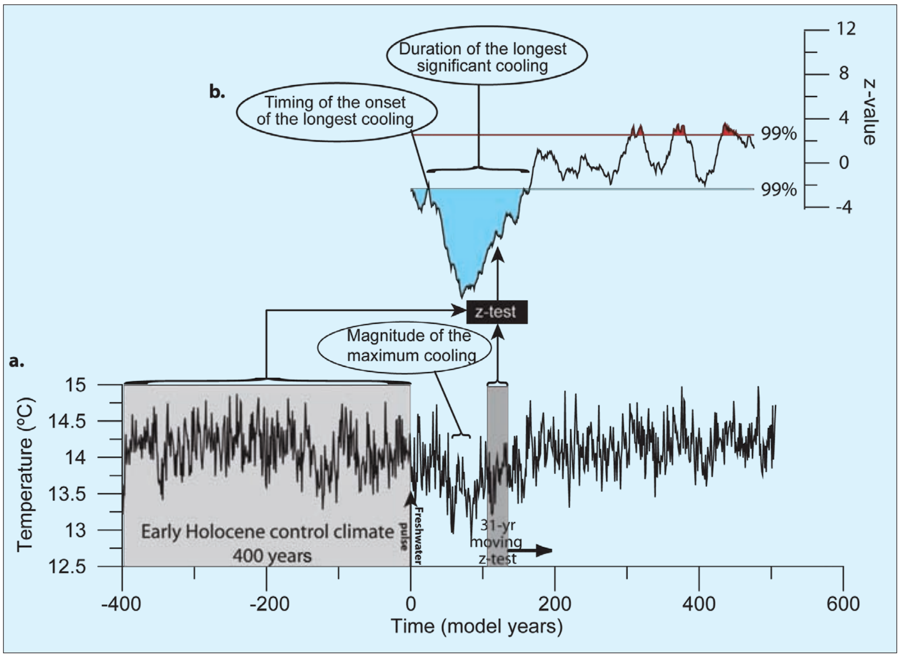

Figure 1: Moving z-test. ***a***) *Surface temperature output from a grid-cell in the climate model. The arrow at t = 0 indicates the introduction of the freshwater pulse. By way of a moving z-test, we assess if the mean of the 31-year moving window is statistically different from the mean of the early Holocene control climate.* ***b***) *the z-test values are separated into significant and non-significant values. Also indicated are the variables plotted in Figure 2: Timing of the longest anomalous cooling, duration of the longest anomalous cooling and the maximum 31-year cooling.*