The MARGO synthesis of sea-surface temperature reconstructions for the Last Glacial Maximum as a benchmark for PMIP3 data-model comparison

- perspectives and limitations discussion document -
6.12.2010

MARGO core group: Michal Kucera, Alan Mix, André Paul, Antoni Rosell-Melé, Ralph Schneider, Claire Waelbroeck, Mara Weinelt

a) The available LGM SST data

The sea-surface temperature (SST) of the Last Glacial Maximum (LGM) oceans has long served as the prime target for data-model comparison. The early recognition of this target led to the availability of large amounts of data, which facilitated the reconstruction of spatially continuous LGM SST patterns. The latest synthesis of LGM SST data has been presented by the MARGO (Multiproxy Approach for the Reconstruction of the Glacial Ocean surface) group in 2009. This synthesis faced three major challenges:

- alignment of individual SST reconstructions to a common chronological framework
- harmonization of seasonal and vertical- attribution of SST reconstructions by different methods
- treatment of spatial divergence among individual SST reconstructions

All three challenges add different levels of uncertainty to the final synthesis. Unlike proxy-calibration errors, the uncertainty associated with these challenges remains difficult to quantify objectively. The increasing availability of SST reconstructions by multiple proxies in the same regions indicates that the prime challenge lies in the correct seasonal and vertical attribution of SST reconstructed by different methods in different oceanic settings. The second major challenge is the existence of significant regional structure in the pattern of LGM SST anomalies, which necessarily leads to large uncertainties when SST reconstructions are extrapolated into unsampled regions.

For these reasons the MARGO synthesis marks a step away from a spatially continuous and unambiguously resolved paleo-SST field. Instead, it provides all individual SST reconstructions by a semi-quantitative measure of reliability due to proxy-calibration and sample age attribution. In terms of SST attribution, each reconstructed value has been assigned to the most likely season it represents and the vertical attribution has been set to the 10-m level in the water column. Both of these attributions are simplifications and both are likely to contribute to the observed spatial divergence among individual SST reconstructions. In order to attempt to combine the individual sources of uncertainty, the MARGO synthesis includes a procedure that assigns seasonally resolved paleo-SST values and associated uncertainties to each 5°x5° grid cell of the global ocean that contains actual data. For regions with sufficient data density, like the tropical ocean, a procedure has been suggested that propagates the SST values and uncertainties associated with individual grid cells into unsampled regions, allowing estimates of the total thermal state of the surface ocean in individual latitudinal zones or basins.

The current MARGO synthesis thus contains two types of products:

- Raw data: Locality, proxy, SST value assumed to represent 10-m depth, seasonal attribution, uncertainty
- Grid averages: 5°x5° cell, SST value assumed to represent 10-m depth for each season where data are available, uncertainty

b) Perspectives for PMIP3 data-model comparison

Considering the development both in the data and the modeling community, there is substantial potential to produce a more meaningful, process-oriented and quantitative insight from LGM SST reconstructions. Importantly, the progress requires working together on new ways to analyse models and interpret data. The main ideas are summarized in the following four theses.

- The data-model comparison should **concentrate on localities and regions with SST reconstructions**

The current coverage of the LGM ocean by actual data even in the latest MARGO synthesis is so low (20% of 5°x5° grid cells) that any attempt of convert the data into a spatially continuous reconstruction will necessary add a considerable further uncertainty to the data. Thus, unless spatially continuous coverage exists, the data community should not attempt to produce continuous maps for modelers and the modelers should not ask for these or generate these from discontinuous data. If we aimed at generating a spatially continuous picture of the past, we would indeed need a major progress in data assimilation, but this is not the purpose of PMIP3.

- The comparison should account for **ambiguity of seasonal and vertical SST attribution**

At present, all MARGO SST values are unambiguously assigned to season and 10-m depth. This value is associated with an estimate of uncertainty, which among other aspects also includes the uncertainty in the attribution of the SST. This is a rather indirect and non-transparent method. Instead, the data community could attempt to state for each SST reconstruction during which season and at which depth the signal was likely formed, or more precisely, assign to it a probability density function in the joint distribution of season and depth. This may not be realistic for PMIP3, but it is in our view the way to go in the future. Climate models have the advantage of producing SST at all seasons and depths, which should facilitate a comparison with ambiguously attributed paleodata.

- Data-model **comparisons must become quantitative and intelligent**

At present, PMIP has not gone much beyond a comparison of zonal averages or qualitative statements on the similarity between reconstructions and models. What we need is an objective way of comparing how well a model represents a given set of reconstructions and most importantly, a way to diagnose what the difference really means – a different pattern or a regionally shifted pattern? For example, Guiot et al. (Data-model comparison using fuzzy logic in paleoclimatology, Climate Dynamics, Volume 15, Number 8, 569-581, DOI:10.1007/s003820050301) advocated an approach that would tolerate a model bias such as a shifted pattern in a paleo- or even control simulation. Such step is of course not trivial and the question "different pattern or a regionally shifted pattern" requires that we can be sure about the patterns in a climate model - which we often cannot, and this has to be quantified as the "model error" or "structural error". Nevertheless, without a progress in this field, PMIP3 will hardly represent a significant improvement on previous work.

- **Model divergence should help identify key regions and parameters** for testing model performance

Rather than blindly collecting and generating new data to increase the coverage of the LGM time-slice, much could be gained from a joint effort in which the modeling community finds a way to use the parameter sensitivity of individual models and the divergence between different models to identify regions and oceanographic parameters which are likely to provide the best resolution in testing model output. Armed with such knowledge, the data community could focus on the data synthesis, analysis of existing records or generation of new data on these regions (and proxies). For example, a synthesis of proxies for the state of the deep-sea during the LGM appears possible (see the "`MARGO Deep"' initiative at www.glacialoceanatlas.org), but such efforts would benefit hugely from explicit justification derived from model simulations.

c) Consequences for specific activities

A major consequence of the above concept is that PMIP3 has to clearly acknowledge that a choice is to be made between more realistic simulations of past climates (in the sense of "data assimilation", with the focus on the best [and "spatially continous"] depiction of past climate) against more stringent tests of climate model performance (for the purpose of "parameter optimization", with the goal to improve predictions of future climate). The former is a broader goal for the scientific community; the latter is the agenda for PMIP3. In terms of data-model comparison, it appears to us that PMIP3 does not need continuous and unambiguously attributed data; it needs clever ways of making the most of the existing discontinuous and ambiguous data and methods to justify data choice by analyzing model results.

The following operational steps could be implemented in a joint effort between the data and modeling communities within PMIP3:

(1) For the LGM, there are still orders of magnitude less data than for present day. However, generating new data blindly can no longer be justified. We should only strive **to generate and add new data** to the synthesis in such regions where we have reasons to believe that it will really make a difference to the diagnosis of the models.

(2) If the analysis of the present day MARGO synthesis and/or of model divergence indeed shows that new data are necessary, it would mean that a mechanism is needed to **update** the MARGO benchmark regularly. The best way would be via a community web-service.

(3) Similarly, the present marine synthesis could be **expanded towards new oceanographic parameters**, such as the deep ocean, but all such attempts should be informed by considerations of optimum diagnostic value.

(4) All paleo-data, including MARGO SST has to be associated with **explicit measures of uncertainty and ambiguity** in seasonal and spatial attribution. Error estimates need to be made less subjective and more quantitative. (Note that the MARGO error is only defined within a constant factor, because it is proportional to "mean reliability index", which we deliberately scaled such that its minimum value is one.) Thus studies are needed that assess the quantitative influence of, for example, the quality of the age model, or the (non-) stationarity through time and in space of the calibration error.

(5) The **model uncertainty** needs to be estimated, too, possibly by developing larger multi- and single-model ensembles. This will allow a data-model comparison to be carried out as an overlay between two probability density surfaces.

(6) Known structural problems of model simulations have to be circumvented and potential **structural problems have to be identified**. This is essential to determine which model-data differences arise because of a-priori model limitations and which differences represent genuine failure of the model to generate an objectively existing oceanographic feature or to place it in the correct position.

MARGO resources:

Official site with all data: http://margo.pangaea.de/
André Paul's site with summaries: http://www.geo.uni-bremen.de/~apau/margo/margo_EN.html
Glacial Ocean Atlas: http://www.glacialoceanatlas.org

MARGO references:

Otto-Bliesner, B.L., Schneider, R., Brady, E.C., Kucera, M., Abe-Ouchi, A., Bard, E., Braconnot, P., Crucifix, M., Hewitt, C., Kageyama, M., Marti, O., Paul, A., Rosell-Melé, A., Waelbroeck, C., Weber, S.L., Weinelt, M., Yu, Y. 2009. A comparison of PMIP2 model simulations and the MARGO proxy reconstruction for tropical sea surface temperatures at last glacial maximum. *Climate Dynamics*, 32(6): 799-815.

MARGO Project Members, 2009. Constraints on the magnitude and patterns of ocean cooling at the Last Glacial Maximum. *Nature Geoscience*, 2: 127-132.

de Vernal, A., Rosell-Melé, A., Kucera, M., Hillaire-Marcel, C., Eynaud, F., Weinelt, M., Dokken, T., Kageyama, M., 2006. Comparing proxies for the reconstruction of LGM sea-surface conditions in the northern North Atlantic. *Quaternary Science Reviews,* 25(21-22): 2820-2834.

Kageyama, M., Laîné, A., Abe-Ouchi, A., Braconnot, P., Cortijo, E., Crucifix, M., de Vernal, A., Guiot, J., Hewitt, C.D., Kitoh, A., Kucera, M., Marti, O., Ohgaito, R., Otto-Bliesner, B., Peltier, W.R., Rosell-Melé, A., Vettoretti, G., Weber, S.L., MARGO Project Members, 2006. Last Glacial maximum temperatures over the North Atlantic, Europe and western Siberia: a comparison between PMIP models, MARGO sea-surface temperatures and pollen-based reconstructions. *Quaternary Science Reviews,* 25(17-18): 2082-2102.

Kucera, M., Schneider, R., Weinelt, M., Eds., 2005. Multiproxy approach for the reconstruction of the glacial ocean surface (MARGO). *Quaternary Science Reviews*, v. 24, issue 7-9, pp. 813-1107.

MARGO Deep reference:

Paul, A. Mulitza, S., 2009. Challenges to understanding ocean circulation during the Last Glacial Maximum, *Eos, Transactions American Geophysical Union,* 90 (19), 169, doi:10.1029/2009EO190004