Supplementary material to "A New Framework for Inferring Earth's Past Climate"

6 September 2011

Martin Tingley, Institute for Mathematics Applied to Geosciences, National Center for Atmospheric Research, Boulder, Colorado

Eugene Wahl, National Climate Data Center, National Oceanic and Atmospheric Administration, Boulder, Colorado

Edward Cook, Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York

Citation:

Tingley, M., E. Wahl, and E. Cook (2011), A new framework for inferring Earth's past climate, *Eos Trans. AGU*, *92*(36), 299, doi:10.1029/2011EO360003. **[Full Article]**

Key Speakers

Dr. Andrew Gelman of Columbia University gave the opening keynote address, and Drs. Martin Tingley of NCAR, Bo Li of Purdue University, Johannes Werner of the University of Geissen (Germany), Matthew Schofield of the University of Kentucky, and Naresh Devineni of Columbia University led the workshop with regard to the use and implementation of BHMs for spatial climate reconstruction.

Extended Discussion by Topic Area

• A strong focus was put on separating model building, per se, from inference of model parameters. It was noted that climate scientists sometimes mix these two concepts, which can result in significant attention being paid to inference issues and comparisons of performance within a closely-related set of models rather than to the more general issue of developing conceptually appropriate yet empirically tractable physical models. example, Bürger et al (2006) demonstrated a lack of robustness in reconstructions resulting from a number of methods, all of which can be thought of as variations on a single, basic regression model. As with other papers in this field, (e.g., Christiansen, 2009), the focus is on the sensitivity of results to various robust inference techniques applied to a common model. In this regard, the key shift in thinking is not to Bayesian methods but to models – which would likely be hierarchical in nature. Inference can then been conducted using a range of tools, but as models become more involved, Bayesian inference strategies may be the (conceptually) simplest option. This kind of shift in thought may be helpful in ensuring that future methodological debates in the paleoclimate literature focus on issues surrounding the construction of scientifically based models, rather than the details of the statistical inference.

• A closely related focus was put on using particular scientific questions to motivate model construction. For example, if the goal of the analysis is to understand how estimated forcings interact with climate, then forcing information should properly be included directly into the hierarchical model. Simultaneous estimation on the past climate and the parameters defining the relationship or interaction between the climate system and the forcings then allows for statistically coherent error estimation and uncertainty propagation. This point was debated from the logical position that separately-derived estimates of forcings and reconstructed climate variables may be needed for completely independent sensitivity estimation.

• It was noted that a particular advantage of reconstruction techniques that produce an ensembles of possible reconstructions is that uncertainty estimates for any function of the target field can be derived by applying the function to each ensemble member and then taking percentiles. Ensemble-based climate reconstruction techniques can thus be used to estimate, in a logically sound manner, quantities such as the likelihood that the most recent decade was the warmest of the past millennium - which requires inference on the distribution of decadal averages. Such inference is not readily attainable from traditional, frequentist point estimates and confidence intervals for each year. More generally, Bayesian inference provides an estimate of the joint probability density functions of the climate variables of interest, along with any model parameters, conditional on the prior probability distributions, model assumptions, and data.

Dendroclimatology has traditionally focused on minimizing bias by validating the accuracy of results. Typically, a portion of the climate data is withheld from the statistical estimation procedure and is compared with estimates from the calibrated proxy data to assess mean biases. Full uncertainty quantification, however, must also carefully consider the precision, or, equivalently, the variance of a reconstruction. In the former context, the well-established validation tools are very useful; in the latter, quantities like the width of confidence or credible intervals, and the nominal versus actual coverage rates of these intervals become additional important validation tools. In any case, it is necessary to consider both accuracy and precision since neither approach alone can provide the range of information needed to fully assess the quality of a climate reconstruction.

Incorrect modeling assumptions can affect both the the accuracy and precision of a climate reconstruction, regardless of the inference technique. While Christiansen *et al* (2009) argue that the key shortcoming of many reconstruction techniques is bias resulting from model misspecifiction, Tingley and Huybers (2010b), show that hierarchical models with Bayesian inference can produce reconstructions that are robust to departures from the modeling assumptions. Pseudo-proxy experiments, based on applying reconstruction techniques to corrupted climate model output, are an important tool for investigating these issues.

• It was generally acknowledged that standard stationary and isotropic covariance functions are likely not sufficiently flexible in the context of reconstructing past climate. Techniques for assimilating empirically derived EOFs, well-known spatial teleconnection patterns, or other complex spatial structures into the hierarchical modeling framework were discussed on several occasions. Dr. Werner showed one possible method of how this might be done, but noted that initial implementation did not lead to improved spatial field skill in parts of the field that are devoid of proxy information. The revised version of the SAMSI theoretical paper on BHM for paleoclimate (Tingley et al., 2010) will provide additional discussion on how this issue might be addressed. Conceptually, specific spatial patterns can be included in the process-level of the BHM, which describes the statistical structure of the target climate field. The key modeling choice is whether to include a specific spatial pattern in the specification of the mean structure or the covariance structure, and the appropriate choice likely depends on the particulars of the analysis. In either case, the residual spatial covariance should be accounted for using one of the standard spatial covariance functions - such as an exponential decay of correlation with separation - and not spatially uncorrelated noise.

It is critical to note that including empirically derived spatial patterns involves assuming that these structures are stationary through time. That is, while the weightings of the patterns may change as a function of time and time scale (for example, annual means versus decadal means), the spatial pattern themselves are constant. Similarly, the parameters of a standard spatial covariance function are assumed stationary in time, so the issue in essence becomes the extent to which particular assumptions of temporal stationarity are considered valid. The position was advanced that it is important to always focus on the development of physically motivated process-level models, rather than technical details, such as the number of EOFs to retain.

• BHMs are not "one size fits all." A given model, such as BARCAST (Tingley and Huybers, 2010a), may be appropriate for one set of data and target process – in the sense that all diagnostics indicate the modeling assumptions are suitable, the Markov Chain Monte Carlo estimation process converges, and the resulting ensemble of draws has reasonable properties. However, the same model may produce results that are physically unreasonable if applied to a different data set, or used to infer a different target process. Such results can often be interpreted as an indication of model misspecification, and it was stressed that model building is an iterative process. Akin to the residual analysis that follows standard linear regression, BHMs allow for posterior analysis of the suitability of the model assumptions for the data under analysis. Both successful and unsuccessful results for the BARCAST algorithm were shown and evaluated, and the unsuccessful applications provided insights into both the data and how to improve upon the analysis model.

• Hierarchical modeling and Bayesian inference are still in their infancy in the context of paleoclimate field reconstructions. As a single model cannot be reasonable for all data and for answering all scientific questions, the first generation of BHMs will certainly not be the best implementation of this emerging reconstruction paradigm. A key goal of workshops such as this one is to develop a common language and to focus on formalizing scientific understanding, with the objective of improving the second and following generations of Hierarchical Models. At this point, simple BHMs may underperform relative to more established single-level regression models with frequentist inference. A clear example was demonstrated by comparing the BARCAST results for inferring Palmer Drought Severity Index reconstructions in the North American Colorado Plateau region with those from a revised implementation of the point-by-point regression method of Cook et al. (2004). This kind of outcome is to be expected in the early stages of applying a new paradigm into an established field, as the established methods were themselves developed iteratively over the course of decades by researchers with deep understanding of the data and climate system.

References

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

Christiansen, B., T. Schmith, and P. Thejll (2009). A surrogate ensemble study of climate reconstruction methods: Stochasticity and Robustness. Journal of Climate, 22 (4), 951–976.

Cook, E.R., C.A. Woodhouse, C.M. Eakin, D.M. Meko, and D.W. Stahle (2004). Long-Term Aridity Changes in the Western United States. Science, 306, 1015–1018.

Tingley, M.P. and P. Huybers (2010a). A Bayesian Algorithm for Reconstructing Climate Anomalies in Space and Time. Part I: Development and Applications to Paleoclimate Reconstruction Problems. Journal of Climate, 23 (10), 2759–2781.

Tingley, M. and P. Huybers (2010b). A Bayesian Algorithm for Reconstructing Climate Anomalies in Space and Time. Part 2: Comparison with the Regularized Expectation-Maximization Algorithm. *Journal of Climate*, 23 (10), 2782-2800

Tingley, M.P., P.F. Craigmile, M. Haran, B. Li, E. Mannshardt-Shamseldin, and B. Rajaratnam (2010). Piecing together the past: Statistical insights into paleoclimatic reconstructions. Technical Report 2010–09, Stanford University, Department of Statistics. http://statistics.stanford.edu/~ckirby/techreports/GEN/2010/2010-09.pdf