L. Tierney. *Lisp-stat: an object-oriented environment for statistical computing and dynamic graphics*. Wiley: New York, 1990.

L. Tierney. **tkrplot**: simple mechanism for placing R graphics in a tk widget. URL http://cran.r-project.org/.

S. Urbanek and M. Theus. **iPlots**: high interaction graphics for R. In *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*, eds. K.Hornik, F.Leisch, A.Zeileis. Vienna, March 2003. ISSN: 1609-395X, 2003.

Adrian Bowman
Department of Statistics, University of Glasgow
adrian@stats.gla.ac.uk

Ewan Crawford
Department of Statistics, University of Glasgow
ewan@stats.gla.ac.uk

Richard Bowman
Churchill College, University of Cambridge
rwb27@cam.ac.uk

R's role in the climate change debate

by Matthew Pocernich

A contentious issue within the climate research community is the validity of a sequence of reconstructions of global surface temperatures which exhibit a hockey stick-like appearance. These graphs show relatively modest hemispheric temperature variations over the past millennium followed by a sharp increase over the 20th century (See Figure 1). Climate data collected instrumentally is largely limited to the last 150 years. For information prior to this, scientists rely on proxy sources such as tree rings, ice core samples and other types of information. The reconstructions created by Mann et al. (1998) and Mann et al. (1999) have attracted particular attention due to their prominence in the last International Panel on Climate Change reports (IPCC, 2001). Various criticisms against the Mann et al. reconstructions have been voiced over the years, but most recently, articles beginning with McIntyre and McKitrick (2003) questioned whether principal component analysis (PCA) was correctly used to summarize many sources of proxy data. The contention was that the implementation of PCA using a common centering convention limited to the overlap with the instrumental data rather than using a full-length centering could potentially produce a spurious offset, even in random data. This criticism has been transformed into a much broader challenge regarding the existence of climate change.

The National Research Council (NRC) established a committee to re-evaluate Mann's work in light of the subsequent criticisms. On June 22nd, a press conference was held to announce the findings of the report and answer questions. The Chair of the committee, Dr. Gerald North, made several remarks relating sharing code and data (Board on Atmospheric Sciences and Climate, 2006).

> Our view is that all research benefits from full and open access to published datasets and that a clear explanation of analytical methods is mandatory. Peers should have access to the information needed to reproduce published results, so that increased confidence in the outcome of the study can be generated inside and outside the scientific community. Other committees and organizations have produced an extensive body of literature on the importance of open access to scientific data and on the related guidelines for data archiving and data access (e.g., NRC 1995). Paleoclimate research would benefit if individual researchers, professional societies, journal editors, and funding agencies continued to improve their efforts to ensure that these existing open access practices are followed.

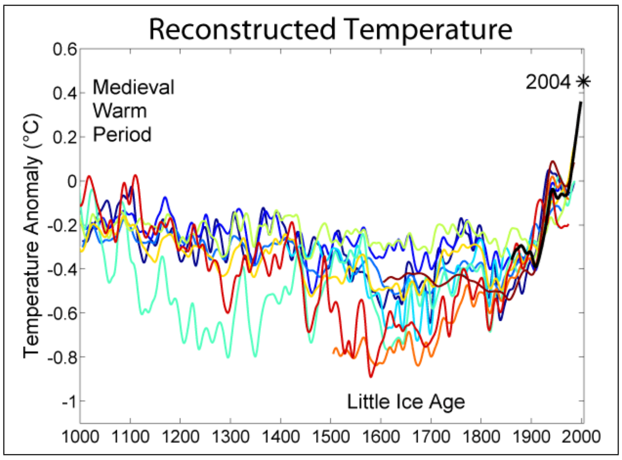

Figure 1: Reconstructed temperatures for the last millennium. Wahl and Ammann (2006)

The statistical and scientific details of this debate are not the point of this note, but rather the ways in which R has had a role in the dialog. A figure in the NRC report was created using R to illustrate how spurious trends could result from the use of principal component analysis (Figure 9-2). This is not too exciting by itself. What is more unique is that the code used to generate this figure was placed in the appendix, where it is merely referred as "R Code".

In a response to criticisms in McIntyre and McKitrick (2003), several articles were written including Wahl and Ammann (2006). Again, the issue here is not the content of the article. What is somewhat unique in the climate community is that all the code and data used to recreate the Mann's original analysis has been made available to the pubic. See www.cgd.ucar.edu/ccr/ammann/millennium/MBH_reevaluation.html. Since the analysis is in R, anyone can replicate the results and examine the methods. With such a contentious issue such as climate change, the strategy of sharing code and data may help to keep the focus on primary issue at hand: namely climate change.

R is increasingly a language of choice in the climate analysis arena. Members of the meteorological community have contributed R packages to facilitate the study climate and weather. Advanced spatial statistical tools such as thin plate spline fitting is found in **fields** (Nychka, 2005). The package **extRemes** (Gilleland et al., 2004) provides tools for applying extreme value statistics. **verification** (Pocernich, 2006) contains methods useful for verifying weather forecasts. **ensembleBMA** (Raftery et al., 2006) provides tools to create probabilistic forecasts out of an ensemble of predictions. **clim.pact** (Benestad, 2005) contains methods for downscaling data from climate models to weather models. Climate models can generate large amounts of data which is typically stored in netcdf files. The packages **ncdf** (Pierce, 2006) and **RNetCDF** (Michna, 2005) allow access to netcdf formatted files. Many of these packages have utility for those outside the weather and climate community. Undoubtedly this list is incomplete.

In discussing R and the study of climate, several points can be made. First, R code is included in the National Academy of Science report whose audience is a broad spectrum of the science and public policy community. In this report much is explained in detail or footnoted, but R was not further defined. Implicitly, this speaks to R's acceptance within the climate community. To further the discussion of climate change, in this report the NAS urges open access to statistical methods and data sharing. To this end, climate scientists such as Ammann have published their R code and data. Finally, many packages have been submitted to CRAN by the weather and climate community that both address both climatology-specific needs and more general statistical topics.

Bibliography

R. E. Benestad. *clim.pact: Climate analysis and downscaling package for monthly and daily data*. The Norwegian Meteorological Institute, 2005.

Board on Atmospheric Sciences and Climate. Surface temperature reconstructions for the last 2,000 years. Technical report, National Research Concil of the National Academies, 2006.

E. Gilleland, R. Katz, and G. Young. *extRemes: Extreme value toolkit*. National Center for Atmospheric Research, 2004. URL http://www.assessment.ucar.edu/toolkit/.

IPCC. *Climate Change 2001: The Scientific Basis*. Cambridge University Press, Cambridge, UK., 2001.

M. Mann, R. Bradley, and M. Hughes. Global-scale temperature patterns and climate forcing over the past 6 six centuries. *Nature*, 392:779–787, 1998.

M. E. Mann, R. S. Bradley, and M. K. Hughes. Northern hemisphere temperatures during the past millennium: Inferences, uncertainties, and limitations. *Geophysical Research Letters*, 26:759–762, 1999.

S. McIntyre and R. McKitrick. Corrections to the Mann et al. (1998) proxy data base and northern hemispheric average temperature series. *Energy Environment*, 14(6):751–771, 2003.

P. Michna. *RNetCDF: R Interface to NetCDF Datasets*, 2005. URL http://www.unidata.ucar.edu/packages/netcdf/. R package version 1.1-3.

D. Nychka. *fields: Tools for spatial data*. National Center for Atmospheric Research, 2005. URL http://www.image.ucar.edu/GSP/Software/Fields. R package version 2.3.

D. Pierce. *ncdf: Interface to Unidata netCDF data files*, 2006. URL http://cirrus.ucsd.edu/~pierce/ncdf. R package version 1.6.

M. Pocernich. *verification: Forecast verification utilities*. National Center for Atmospheric Research, 2006. R package version 1.08.

A. E. Raftery, J. M. Sloughter, and M. Polakowski. *ensembleBMA: Probabilistic forecasting using Bayesian Model Averaging*. University of Washington, 2006.

E. Wahl and C. Ammann. Robustness of the mann, bradley, hughes reconstruction of northern hemisphere surface temperatures: Examination of criticisms based on the nature and processing of proxy climate evidence. *Climatic Change*, in press, 2006.

Matthew Pocernich
Research Applications Laboratory
The National Center for Atmospheric Research
pocernic@rap.ucar.edu