The Annals of Applied Statistics
2011, Vol. 5, No. 1, 76–79
DOI: 10.1214/10-AOAS398B
Main article DOI: 10.1214/10-AOAS398

DISCUSSION OF: A STATISTICAL ANALYSIS OF MULTIPLE TEMPERATURE PROXIES: ARE RECONSTRUCTIONS OF SURFACE TEMPERATURES OVER THE LAST 1000 YEARS RELIABLE?[1,2]

BY JASON E. SMERDON

Lamont-Doherty Earth Observatory of Columbia University

McShane and Wyner [(2011); hereinafter MW11] reiterate a well-known and central challenge of paleoclimatology: it is fraught with uncertainties and based on noisy observations. Decades of research have aimed at characterizing these uncertainties and interpreting proxies through laboratory experiments, field observations, theory, process-based modeling, cross-record comparisons, and indeed through statistical modeling and hypothesis testing. It is against this larger backdrop that the problem addressed by MW11 must be considered. Attempts to reconstruct global or hemispheric temperature indices and fields using multi-proxy networks are an outgrowth of many efforts in paleoclimatology, but represent relatively recent pursuits in the field. They provide neither the principal scientific evidence supporting climate-proxy connections, nor the most compelling, and the inference by MW11 that their own findings demonstrate a widespread failure in the predictive capacity of climate proxies is at odds with most other independent lines of proxy research.

The above considerations notwithstanding, I focus on one principal argument by MW11 that uses cross-validation experiments to conclude that "*proxies are severely limited in their ability to predict average temperatures and temperature gradients.*" I demonstrate that this claim is based on a hypothesis test subject to Type II errors and therefore an inconclusive evaluation of the temperature sensitivity of proxy archives.

I perform additional cross-validation experiments using 283 time series that are randomly selected from the global CRU temperature field as infilled and subselected (1732 total grid cells) by Mann et al. (2008). I choose 283 samples based on the total number of unique $5^\circ \times 5^\circ$ grid cells that contain the 1209 proxies in the Mann et al. (2008) proxy network. Time series from these cells are used to create five predictor datasets spanning the instrumental period by adding 0, 50, 80 and 94% white noise by variance, and 86% red noise by variance ($\rho = 0.32$), the latter

Received September 2010.

[1]Lamont-Doherty Earth Observatory contribution number 7439.

[2]Supported in part by NSF Grant ATM-09-02436 and NOAA Grant NA07OAR4310060.

Key words and phrases. Climate change, paleoclimate, statistical climate reconstructions, Late-Holocene, common era, multi-proxy reconstructions, model validation.