Reply to "Comments on 'Erroneous Model Field Representations in Multiple Pseudoproxy Studies: Corrections and Implications'"*

Jason E. Smerdon and Alexey Kaplan

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

Daniel E. Amrhein

Massachusetts Institute of Technology–Woods Hole Oceanographic Institution Joint Program in Oceanography, Cambridge, Massachusetts

(Manuscript received 29 March 2012, in final form 9 January 2013)

ABSTRACT

The commenters confirm the errors identified and discussed in Smerdon et al., which either invalidated or required the reinterpretation of quantitative results from pseudoproxy experiments presented or used in several earlier papers. These errors have a strong influence on the spatial skill assessments of climate field reconstructions (CFRs), despite their small impacts on skill statistics averaged over the Northern Hemisphere. On the basis of spatial performance and contrary to the claim by the commenters, the Regularized Expectation Maximization method using truncated total least squares (RegEM-TTLS) cannot be considered a preferred CFR technique. Moreover, distinctions between CFR methods in the context of the discussion in the original paper are immaterial. Continued investigations using accurately described and faithfully executed pseudoproxy experiments are critical for further evaluation and improvement of CFR methods.

Rutherford et al. (2013, hereinafter R13) confirm the errors that were identified and discussed in Smerdon et al. (2010, hereinafter S10). These errors were associated with the processing of the millennium-length National Center for Atmospheric Research (NCAR) Community Climate System Model, version 1.4 (CCSM1.4) (Ammann et al. 2007), and the global Hamburg Ocean Primitive Equation (ECHO-G) (González-Rouco et al. 2003) simulations by Mann et al. (2005) and Mann et al. (2007, hereinafter M07). R13 also clarify that related papers published after M07 were not affected by the errors described in S10. This is an important clarification. Below we respond to several additional arguments raised by R13.

R13 emphasize a distinction between the two versions of the Regularized Expectation Maximization (RegEM) method (Schneider 2001). They imply that RegEM using truncated total least squares (RegEM-TTLS) is a better climate field reconstruction (CFR) method than RegEM using ridge regression (RegEM-Ridge), the latter of which was used by S10 to illustrate some consequences of the data-processing errors. We first note that any CFR method could have been used to demonstrate the errors discovered by S10, making methodological distinctions in this context immaterial. Second, it is true that RegEM-TTLS has been shown in pseudoproxy studies to better reconstruct the Northern Hemisphere (NH) mean (see Smerdon 2012 for a review), but both of the RegEM methods are meant to reconstruct temperature fields. Spatial reconstruction skill therefore is a fundamental measure of their performance. To date, the only comprehensive comparisons of the spatial skill of multiple methods for global temperature CFRs did not find RegEM-TTLS to be a clear frontrunner (Smerdon et al. 2011; Li and Smerdon 2012). To the contrary, RegEM-TTLS performs similarly to other multivariate regression methods in several

* Lamont-Doherty Earth Observatory Contribution Number 7669.

Corresponding author address: Jason E. Smerdon, Lamont-Doherty Earth Observatory, Columbia University, 61 Route 9W, P.O. Box 1000, Palisades, NY 10964.
E-mail: jsmerdon@ldeo.columbia.edu

DOI: 10.1175/JCLI-D-12-00165.1

spatial skill metrics, and all of the evaluated methods have important spatial errors. The advocacy of one multivariate linear CFR method over another is therefore premature.

R13 also claim that similar results are obtained from pseudoproxy experiments using the correctly and incorrectly oriented CCSM1.4 fields. This point requires qualification: the statistics reported in lines three and four of R13's Table 1 are similar only because they are NH averages. The spatial performance of RegEM-TTLS and other CFR methods is nevertheless strongly dependent on the distribution of the pseudoproxy network (Smerdon et al. 2011; Werner et al. 2013; Annan and Hargreaves 2012). Any perceived similarity between results presented by Mann et al. (2005), M07, and R13 therefore only holds for NH-averaged statistics, while regional skill statistics (e.g., for Niño-3) would expose important differences between experiments with correct and incorrect sampling as demonstrated in S10.

Maintaining consistent and correctly documented records of pseudoproxy tests is critical for evaluating CFR methods. The advantage of such tests lies in their ability to serve as common testbeds on which reconstruction methods can be systematically evaluated and compared [see Smerdon (2012) for a review]. This advantage can only be realized if pseudoproxy experiments are accurately described and correctly executed. Timely corrections to pseudoproxy tests are therefore vital for avoiding the perpetuation of errors and inconsistencies in the published literature.

Acknowledgments. This research was supported in part by the National Science Foundation by Grant ATM-0902436 and by the National Oceanic and Atmospheric Administration by Grant NA07OAR4310060.

REFERENCES

Ammann, C. M., F. Joos, D. S. Schimel, B. L. Otto-Bliesner, and R. A. Tomas, 2007: Solar influence on climate during the past millennium: Results from transient simulations with the NCAR Climate System Model. *Proc. Natl. Acad. Sci. USA,* **104,** 3713–3718, doi:10.1073/pnas.0605064103.

Annan, J. D., and J. C. Hargreaves, 2012: Identification of climate state with limited proxy data. *Climate Past,* **8,** 1141–1151, doi:10.5194/cp-8-1141-2012.

González-Rouco, F., H. von Storch, and E. Zorita, 2003: Deep soil temperature as proxy for surface air-temperature in a coupled model simulation of the last thousand years. *Geophys. Res. Lett.,* **30,** 2116, doi:10.1029/2003GL018264.

Li, B., and J. E. Smerdon, 2012: Defining spatial assessment metrics for evaluation of paleoclimatic field reconstructions of the Common Era. *Environmetrics,* **23,** 394–406.

Mann, M. E., S. Rutherford, E. Wahl, and C. Ammann, 2005: Testing the fidelity of methods used in proxy-based reconstructions of past climate. *J. Climate,* **18,** 4097–4107.

——, ——, ——, and ——, 2007: Robustness of proxy-based climate field reconstruction methods. *J. Geophys. Res.,* **112,** D12109, doi:10.1029/2006JD008272.

Rutherford, S., M. E. Mann, E. Wahl, and C. Ammann, 2013: Comments on "Erroneous model field representations in multiple pseudoproxy studies: Corrections and implications." *J. Climate,* **26,** 3482–3484.

Schneider, T., 2001: Analysis of incomplete climate data: Estimation of mean values and covariance matrices and imputation of missing values. *J. Climate,* **14,** 853–887.

Smerdon, J. E., 2012: Climate models as a test bed for climate reconstruction methods: Pseudoproxy experiments. *Wiley Interdiscip. Rev: Climate Change,* **3,** 63–77, doi:10.1002/wcc.149.

——, A. Kaplan, and D. E. Amrhein, 2010: Erroneous model field representations in multiple pseudoproxy studies: Corrections and implications. *J. Climate,* **23,** 5548–5554.

——, ——, E. Zorita, J. F. González-Rouco, and M. N. Evans, 2011: Spatial performance of four climate field reconstruction methods targeting the Common Era. *Geophys. Res. Lett.,* **38,** L11705, doi:10.1029/2011GL047372.

Werner, J. P., J. Luterbacher, and J. E. Smerdon, 2013: A pseudoproxy evaluation of Bayesian hierarchical modeling and canonical correlation analysis for climate field reconstructions over Europe. *J. Climate,* **26,** 851–867.