Erroneous Model Field Representations in Multiple Pseudoproxy Studies: Corrections and Implications*

Jason E. Smerdon, Alexey Kaplan, and Daniel E. Amrhein

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

ABSTRACT

Pseudoproxy experiments evaluate statistical methods used to reconstruct climate fields from paleoclimatic proxies during the Common Era. These experiments typically employ output from millennial simulations by general circulation models (GCMs). It is demonstrated that multiple published pseudoproxy studies have used erroneously processed GCM surface temperature fields: the NCAR Community Climate System Model (CCSM), version 1.4, field was incorrectly oriented geographically and the GKSS ECHO-g FOR1 field was corrupted by a hemispheric-scale smoothing in the Western Hemisphere. These problems are not associated with the original model simulations; they instead arose because of incorrect processing of the model data for the pseudoproxy experiments. The consequences of these problems are evaluated for the studies in which the incorrect fields were used. Some quantitative results are invalidated by the findings: these include all experiments that used the corrupted ECHO-g field and those aspects of previous CCSM experiments that focused on Niño-3 reconstructions. Other results derived from the CCSM field can be reinterpreted based on the information provided herein and their qualitative characteristics remain similar.

1. Introduction

Pseudoproxy studies are used to evaluate the performance of statistical reconstruction methods in the context of proxy reconstructions during the Common Era (e.g., von Storch et al. 2004, 2006; Mann and Rutherford (2002); Mann et al. 2005, 2007a; Lee et al. 2007; Smerdon and Kaplan 2007; Smerdon et al. 2008a; Hegerl et al. 2007; Riedwyl et al. 2009; Christiansen et al. 2009). The majority of these studies use millennial integrations from general circulation models (GCMs), the output of which comprises the test bed for systematic and controlled experiments used to evaluate methodological behaviors. The two most commonly used millennial simulations for this purpose are the GKSS ECHO-g FOR1 (hereafter ECHO-g; Gonzalez-Rouco et al. 2003) and the National Center for Atmospheric Research (NCAR)

Community Climate System Model (CCSM), version 1.4 (Ammann et al. 2007), runs. The motivation of this note is to demonstrate that multiple published pseudoproxy studies have used erroneously processed output from these two simulations.

We have discovered that the geographic orientation of the CCSM field used by Mann et al. (2005, hereafter M05), Mann et al. (2007a, hereafter M07), and Mann et al. (2007b) was incorrect. Smerdon and Kaplan (2007) and Smerdon et al. (2008a) studied the effects of the incorrect standardization procedure applied in M05 and used the incorrectly oriented CCSM field obtained from the M05 supplemental Web site. We also have discovered that the ECHO-g field used in M07 was corrupted by a hemispheric-scale smoothing in the Western Hemisphere (WH). Inconsistencies in the M07 representation of the Northern Hemisphere (NH) mean of the ECHO-g field were already the subject of an exchange by Smerdon et al. (2008b) and Rutherford et al. (2008), with the latter explaining that the inconsistencies were caused by an interpolation of the original ECHO-g field to a 5° × 5° grid. This explanation did not describe the full extent to which the ECHO-g field was corrupted. Here we illustrate the problems with each of the employed fields and how pseudoproxy experiments based on the incorrect fields are either invalidated or require reinterpretation.
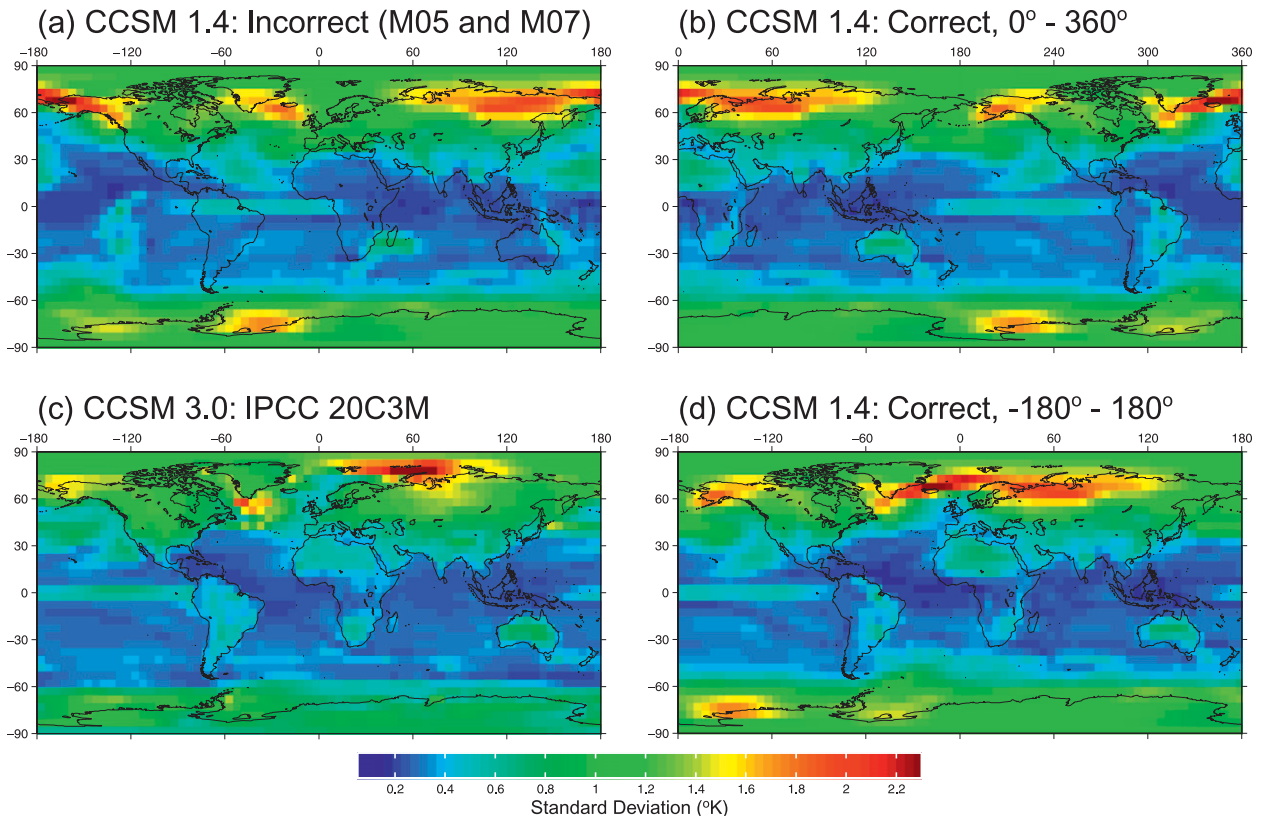
FIG. 1. Standard deviations (1880–1980 C.E.) of the annual surface temperature field for the NCAR CCSM simulations: (a) the incorrect M05 and M07 orientation of version 1.4, −180° to 180°; (b) the correct orientation of version 1.4, 0° to 360°; (c) the version 3.0, IPCC AR4 "Climate of the 20th Century" (20C3M) eight-member ensemble; and (d) the correct orientation of version 1.4, −180° to 180°.

2. Model fields as used in pseudoproxy studies

Both M05 and M07 used the CCSM data; the ECHO-g data were used only by M07. In each study, the model fields were first interpolated from original grid resolutions (T30 for ECHO-g and T31 for CCSM) to regular 5° × 5° grids. Only the resulting 5° × 5° gridded fields were used further by M05 and M07 in their experiments (i.e., to derive pseudoproxy records and masked "instrumental" fields). To our knowledge, the problems described in subsequent sections affecting these regridded fields did not affect the original model data.

We have downloaded and analyzed the archived versions of the regridded model fields from the M05 and M07 supplemental Web sites (http://fox.rwu.edu/~rutherfo/supplements/Pseudoproxy05/ and http://www.meteo.psu.edu/~mann/PseudoproxyJGR06/, respectively). Regridded CCSM annual surface temperature anomalies were available at either Web site as the variable northtosouth in a Matlab file named converted.mat, and ECHO-g annual means of absolute surface temperatures were at the latter Web site as the variable yearave in the Matlab file named gkssregrid.mat. Pseudoproxy networks generated from these datasets were also available at these Web sites and

located in directories organized according to experiments discussed in M05 and M07. At the time of this manuscript's submission, all of the aforementioned archived data were affected by the problems described hereafter and corrected versions of the data were not available at either supplemental Web site.

a. CCSM 1.4 dataset

Figure 1a plots the standard deviations of the CCSM temperature anomalies calculated from 1880 to 1980 C.E. using the version of the field archived by M05 and M07. The problems with the orientation are readily evident in the plotted field, particularly with regard to the near-rectangular area of higher standard deviations extending from the Atlantic to the Pacific Ocean along the equator. There is no obvious dynamical explanation for this feature.

Figure 1c presents the standard deviations based on the mean variance of the eight-member ensemble of CCSM3 simulations (Collins et al. 2006) from the Intergovernmental Panel on Climate Change Fourth Assessment Report (IPCC AR4) "Climate of the 20th Century" experiment archived by the Program for Climate Model Diagnosis and Intercomparison. Comparisons between
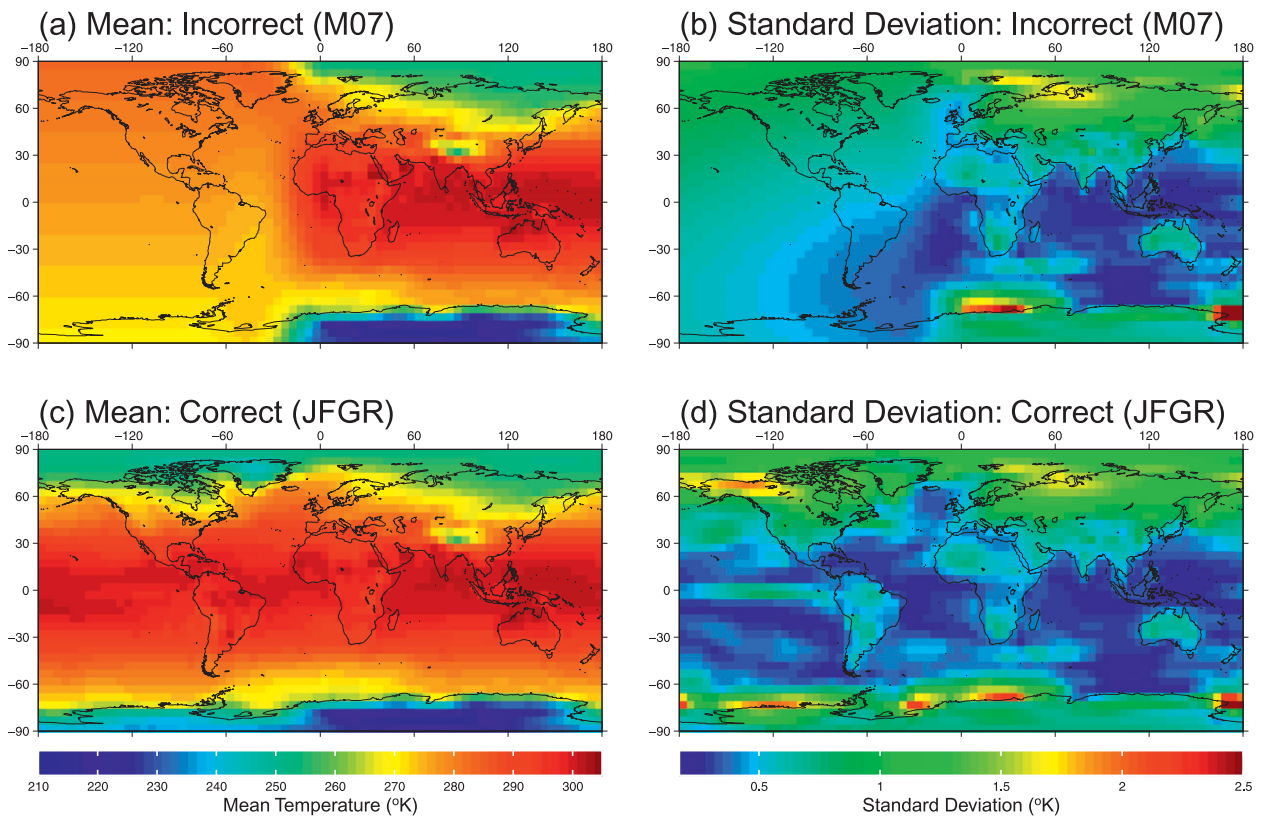
FIG. 2. ECHO-g FOR1 surface temperature statistics from 1880–1980 C.E. for the incorrect (M07) and correct (J. F. González-Rouco 2006, personal communication) versions: (a) means for the incorrect version, (b) standard deviations for the incorrect version, (c) means for the correct version, and (d) standard deviations for the correct version.

Figs. 1a and 1c make clear that the aforementioned equatorial feature in Fig. 1a is supposed to be in the equatorial Pacific and is due to the El Niño–Southern Oscillation (ENSO) phenomenon. Further comparison of Figs. 1a and 1c identifies the problem with the orientation in Fig. 1a: it has been transformed from the original 0° to 360° model domain (Fig. 1b) to the −180° to 180° domain of Fig. 1a using an incorrect transformation of longitudes x to x' in the form $x' = 180° - x$. This transformation establishes the symmetry of patterns in Fig. 1a (incorrect orientation) and Fig. 1b (correct orientation) with respect to a vertical axis drawn equidistantly between them. When the correct transformation ($x' = x$ for $x \leq 180°$ and $x' = x - 360°$ for $x > 180°$) is applied to Fig. 1b, it produces a standard deviation pattern (Fig. 1d) that is broadly similar to the CCSM3 pattern shown in Fig. 1c, despite the difference in the model versions and their forcings. These similarities confirm that Fig. 1d, not Fig. 1a, is correct.

b. ECHO-g FOR1 dataset

Figure 2 presents the mean and standard deviation fields from the ECHO-g simulation archived at the M07 supplementary Web site and from the correct version

acquired directly from J. F. González-Rouco (2006, personal communication). A comparison between the two versions shows that the Eastern Hemispheres (EHs) in the two fields are similar. An unrealistic smoothing throughout the WH of the M07 field is evident, however, and causes a false and unphysical spatial coherency in the hemisphere.

3. Implications for pseudoproxy experiments

The problems with the M05 and M07 versions of the CCSM and ECHO-g model fields were not simply errors in archiving. Smerdon and Kaplan (2007) reproduced the M05 results using the archived, incorrectly oriented CCSM field. Figure 6 in M07 also shows maps of model and reconstructed surface temperature means that are incorrectly oriented. Similarly, Figure S19 in M07 presents skill maps with equatorial minima in rectangular regions collocated with the displaced ENSO area shown in Fig. 1a. The ECHO-g NH mean index presented by M07 is also incorrect (Smerdon et al. 2008b), which below is confirmed as resulting from the corrupted WH. These findings clearly indicate that M05 and M07 actually used

(a) Intended Sampling
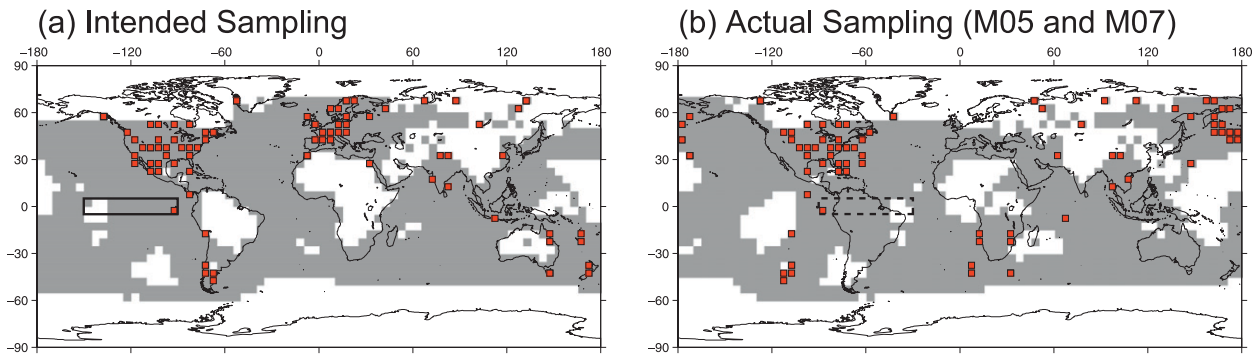
(b) Actual Sampling (M05 and M07)

FIG. 3. Instrumental and proxy locations for (a) the intended sampling distribution approximating the Jones et al. (1999) instrumental data locations and the Mann et al. (1998) proxy record locations and (b) the actual locations sampled in the archived M05 and M07 datasets. Solid and dashed boxes in (a) and (b) enclose the correct and incorrect Niño-3 regions, respectively.

the incorrect model fields as described in section 2 for their collective pseudoproxy experiments.

a. Instrumental mask and pseudoproxy sampling

The pseudoproxy experiments in M05 and M07 intended to reflect patterns of instrumental and proxy data availability as shown in Fig. 3a. Because of the incorrect orientation of the CCSM model field, however, the instrumental and pseudoproxy data were actually taken from the locations shown in Fig. 3b. The instrumental mask incorrectly excluded data from regions such as northern Europe and included data from regions such as continental Australia. Pseudoproxy sampling of the field was also critically changed; presumed dense sampling over Europe, for instance, was instead located over eastern Russia and the North Pacific Ocean.

b. Global patterns of reconstruction skill

The implications for pseudoproxy experiments that used the incorrectly oriented CCSM field are illustrated herein using reconstructions computed with the regularized expectation maximization method (Schneider 2001) and ridge regression (hereafter RegEM-Ridge); all employed pseudoproxies have signal-to-noise ratios of 0.5 (by standard deviation). All reconstruction settings are analogous to those of M05, except that global and nonhybrid reconstructions are performed herein. Spatial patterns of reconstruction skill, as measured by correlation coefficients between reconstructions and the model "truth" in the reconstruction interval (850–1855 C.E.), are presented in Fig. 4. Figure 4a corresponds to a scenario in which the model fields are incorrectly oriented but are thought to be correct, in emulation of the M05 and M07 representation. Figure 4b presents the same results as shown in Fig. 4a, but attributed to their actual geographic locations. The resulting reinterpretation moves areas of high skill to places where pseudoproxies were actually sampled (cf. Figs. 3b and 4b).

Figure 4c plots the correlation field for a RegEM-Ridge reconstruction using the correctly oriented CCSM field and the correct sampling distribution in Fig. 3a. Similar to Fig. 4b, the areas of highest correlation are located where the pseudoproxy network is densest, but now these regions reflect the originally intended sampling distribution. Figure 4d presents the results for the same pseudoproxy experiment as the one shown in Fig. 4c, but for the correct ECHO-g model field. Similar to the CCSM experiment, the highest correlations in the ECHO-g experiment are concentrated over dense pseudoproxy sampling areas. Reconstructions of the ECHO-g temperature field have systematically higher skill in the tropics (Fig. 4d), however, than those targeting the CCSM field (Fig. 4c). These differences are reflected in the global area-weighted averages of the correlation coefficients in Fig. 4c (0.465) and Fig. 4d (0.500). The former is also slightly higher than the average for the reconstruction using the incorrect sampling scheme in Fig. 3b and the CCSM field (0.460).

c. Representation of the Northern Hemisphere mean

Figure 5a plots the decadally filtered area-weighted NH means for the RegEM-Ridge reconstructions using the correct (Fig. 3a) and incorrect (Fig. 3b) distributions; the latter is analogous to the M05 and M07 results. While the two NH means are different, their correlations with the "true" model time series during the reconstruction interval are similar: 0.756 for the correct sampling distribution and 0.715 for the incorrect one. Note that both reconstructed NH mean time series also exhibit the variance losses and mean biases previously noted for the RegEM-Ridge method (cf. Smerdon and Kaplan 2007; Smerdon et al. 2008a).

Regarding the ECHO-g field, the corrupted M07 version has created confusion about its mean NH surface temperature index. An exchange between Smerdon et al. (2008b) and Rutherford et al. (2008) discussed the incorrect M07 representation of the index. The correct index
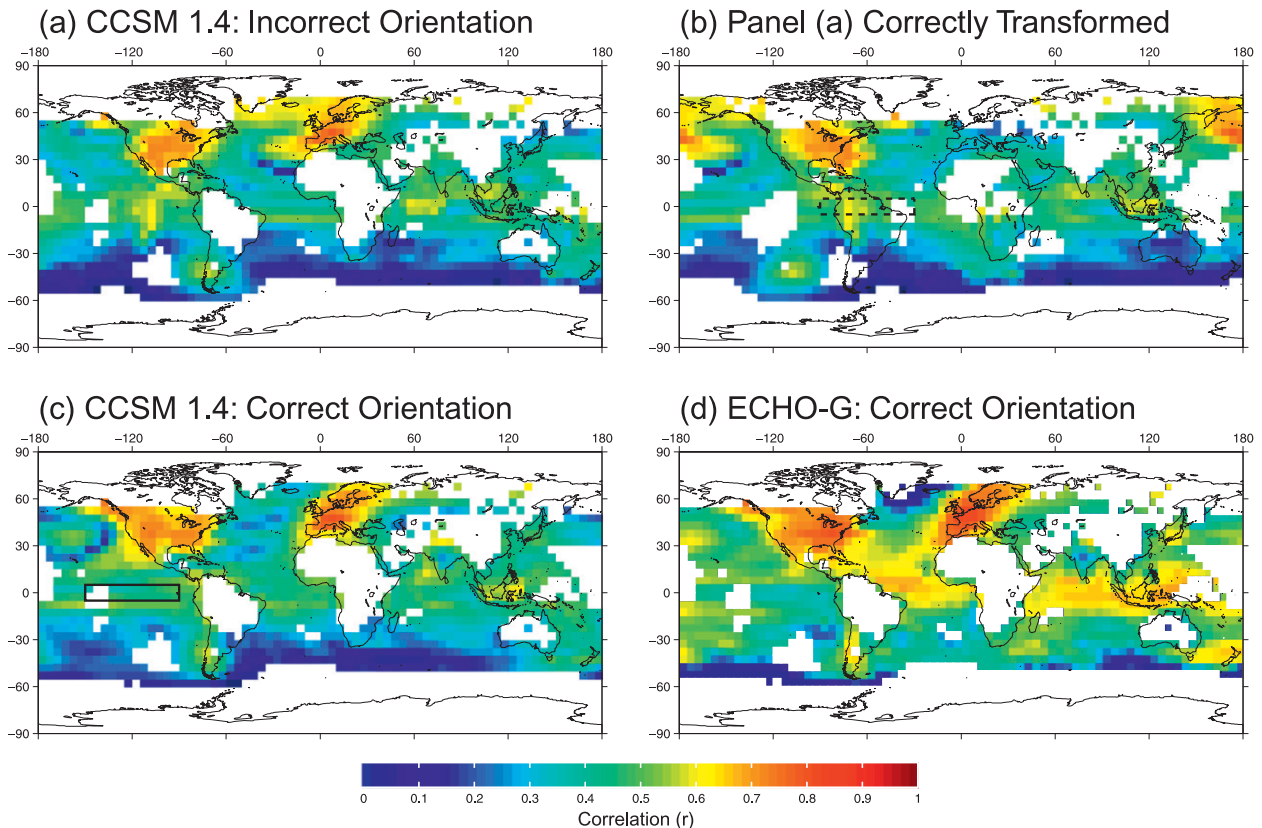
FIG. 4. Reconstruction-interval correlation coefficients between the RegEM-Ridge reconstructions and model "truth" using the CCSM and ECHO-g fields, pseudoproxies with a signal-to-noise ratio of 0.5, a calibration interval from 1856 to 1980 C.E., and reconstruction intervals of 850–1855 and 1001–1855 C.E., respectively. (a) The correlation field for the reconstruction using the incorrectly oriented CCSM field that was assumed to be correct, as in the M05 and M07 cases. (b) The correct transformation of (a) as the reconstruction skill for the unrealistic sampling distribution in Fig. 3b. (c),(d) The correlation fields for the correct sampling distribution shown in Fig. 3a and the correct CCSM and ECHO-g fields, respectively. Dashed and solid boxes in (b) and (c) enclose the incorrect and correct Niño-3 regions, respectively.

is shown in Fig. 5b, as well as the incorrect version presented by M07. Rutherford et al. (2008) explained the discrepancy as arising from interpolating the original ECHO-g field to a 5° × 5° grid using "the 'surface' function in the Generic Mapping Tools (GMT) (Wessel and Smith 1991) package to regrid the ECHO-g output after first averaging all values within a 5° × 5° box." They conceded that other interpolation methods preserve the NH mean better and provided new ECHO-g pseudoproxy results using a bilinearly interpolated field. Their explanation made no mention, however, of the hemispheric-scale smoothing in their version of the ECHO-g field, and implied that the differences were simply the result of two different interpolation choices. The NH mean index recalculated here from the corrupted ECHO-g field is identical to the time series published by M07 (Fig. 5b), indicating that the real problem with the M07 representation of the ECHO-g NH mean index was the corruption of the entire WH reported herein, not one legitimate interpolation choice over another.

4. Discussion and conclusions

We have identified problems with publicly available versions of model fields used in M05, M07, and subsequently published studies. The CCSM millennial integration was incorrectly oriented geographically and caused both visual misrepresentations of the field and spatial sampling that did not reflect reported distributions. We also have found that the ECHO-g model field employed by M07 was corrupted more than previously reported (Smerdon et al. 2008b; Rutherford et al. 2008) by a large-scale and unphysical spatial smoothing in the WH, making pseudoproxy experiments based on this field unrealistic.

Because of the errors reported herein, the quantitative results of all pseudoproxy experiments based on these fields are either invalidated or require reinterpretation. The M05, M07, and Mann et al. (2007b) results regarding NH mean reconstructions or global multivariate skill statistics using the incorrectly oriented CCSM field can be
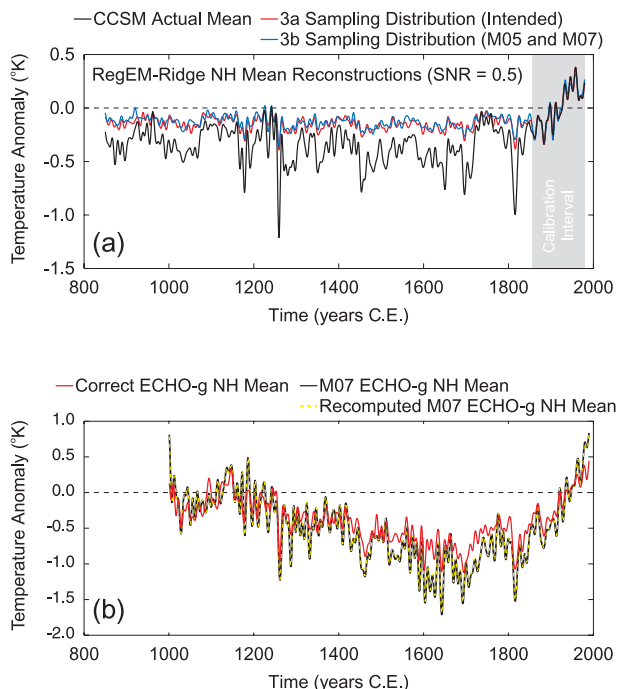
FIG. 5. (a) Area-weighted NH means derived from RegEM-Ridge reconstructions using the two sampling distributions shown in Fig. 3 and the CCSM field. (b) Area-weighted NH means calculated for the correct ECHO-g field and the version archived by M07. Also shown is the version of the NH mean published and archived by M07.

reinterpreted as reflecting results for the sampling distribution represented in Fig. 3b, but cannot be directly compared to future experiments that would employ distributions emulating real-world conditions. The results from CCSM pseudoproxy experiments in M07 that cannot be reinterpreted are the statistics reported for the Niño-3 index: instead of reconstructing the temperature index of the eastern equatorial Pacific (Figs. 3a and 4c), M07 reconstructed an index of land and ocean temperatures spanning the rectangular region indicated in Fig. 3b or 4b.

No pseudoproxy experiments based on the ECHO-g field containing the corrupted WH can be meaningfully reinterpreted. This version of the field is unphysical and has no analog for realistic reconstructions. Some of the M07 ECHO-g results were corrected by Rutherford et al. (2008), although changes in the updated reconstruction statistics were not correctly interpreted because the underlying problems with the M07 ECHO-g field were not accurately identified.

Smerdon and Kaplan (2007) and Smerdon et al. (2008a) used the incorrectly oriented CCSM field from the M05 supplemental Web site to reproduce selected M05 results and compare them with additional reconstructions. These comparisons illustrated how the use of data prior to the instrumental period, which cannot be used for standardization in real-world reconstructions, caused artificially high reconstruction skill (Smerdon and Kaplan 2007). Large mean biases and variance losses for the correct standardization procedure were traced to a systematic difference between the means of the instrumental and preinstrumental periods (Smerdon et al. 2008a). Similar to the M05, M07, and Mann et al. (2007b) CCSM results, the experiments of Smerdon and Kaplan (2007) and Smerdon et al. (2008a) must be reinterpreted as using the instrumental data mask and proxy locations shown in Fig. 3b. The maps in Figs. 6, 7, and 9 from Smerdon et al. (2008a) also must be transformed in the same way that Fig. 4a was transformed to Fig. 4b (or Fig. 1a to Fig. 1d). While these transformations are necessary to correctly interpret the experiments presented for illustration, the conclusions of both papers, as summarized above, remain unaffected.

Future studies of reconstruction methods will require cross-model comparisons that identify model-dependent characteristics in pseudoproxy experiments. These comparisons might illuminate the source of skill associated with specific methods in real-world reconstructions, but their success is dependent on consistent experimental settings, including the correct model field representations and common sampling distributions. Toward such ends, the correct 5° × 5° annual surface temperature fields from the ECHO-g and CCSM millennial integrations are provided online (at http://www.ldeo.columbia.edu/~jsmerdon/2010_jclim_supplement.html).

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. Nat. Acad. Sci. USA,* **104,** 3713–3718, doi:10.1073/pnas.0605064103.

Christiansen, B., T. Schmith, and P. Thejll, 2009: A surrogate ensemble study of climate reconstruction methods: Stochasticity and robustness. *J. Climate,* **22,** 951–976.

Collins, W. D., and Coauthors, 2006: The Community Climate System Model version 3 (CCSM3). *J. Climate,* **19,** 2122–2143.

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.

Hegerl, G. C., T. Crowley, M. Allen, W. T. Hyde, H. Pollack, J. Smerdon, and E. Zorita, 2007: Detection of human influence on a new, validated 1500-year climate reconstruction. *J. Climate,* **20,** 650–666.

Jones, P. D., M. New, D. E. Parker, S. Martin, and J. G. Rigor, 1999: Surface air temperature and its changes over the past 150 years. *Rev. Geophys.,* **37,** 173–199.

Lee, T. C. K., F. W. Zwiers, and M. Tsao, 2007: Evaluation of proxy-based millennial reconstruction methods. *Climate Dyn.,* **31,** 263–281, doi:10.1007/s00382-007-0351-9.

Mann, M. E., and S. Rutherford, 2002: Climate reconstruction using 'pseudoproxies.' *Geophys. Res. Lett.,* **29,** 1501, doi:10.1029/2001GL014554.

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

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

——, ——, ——, and ——, 2007b: Reply. *J. Climate,* **20,** 5671–5674.

Riedwyl, N., M. Küttel, J. Luterbacher, and H. Wanner, 2009: Comparison of climate field reconstruction techniques: Application to Europe. *Climate Dyn.,* **32,** 381–395.

Rutherford, S., M. E. Mann, E. Wahl, and C. Ammann, 2008: Reply to comment by Jason E. Smerdon et al. on "Robustness of proxy-based climate field reconstruction methods." *J. Geophys. Res.,* **113,** D18107, doi:10.1029/2008JD009964.

Schneider, T., 2001: Analysis of incomplete climate data: Estimation of mean values and covariance matrices and imputation of missing values. *J. Climate,* **14,** 853–871.

Smerdon, J. E., and A. Kaplan, 2007: Comments on "Testing the fidelity of methods used in proxy-based reconstructions of past climate": The role of the standardization interval. *J. Climate,* **20,** 5666–5670.

——, ——, and D. Chang, 2008a: On the origin of the standardization sensitivity in RegEM climate field reconstructions. *J. Climate,* **21,** 6710–6723.

——, J. F. González-Rouco, and E. Zorita, 2008b: Comment on "Robustness of proxy-based climate field reconstruction methods" by Michael E. Mann et al. *J. Geophys. Res.,* **113,** D18106, doi:10.1029/2007JD009542.

von Storch, H., E. Zorita, J. M. Jones, Y. Dimitriev, F. González-Rouco, and S. F. B. Tett, 2004: Reconstructing past climate from noisy data. *Science,* **306,** 679–682.

——, ——, ——, F. González-Rouco, and S. F. B. Tett, 2006: Response to comment on "Reconstructing past climate from noisy data." *Science,* **312,** 529c.

Wessel, P., and W. H. F. Smith, 1991: Free software helps map and display data. *Eos, Trans. Amer. Geophys. Union,* **72,** 441, doi:10.1029/90EO00319.