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, NY, U.S.A.

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

†LDEO contribution number XXXX
ABSTRACT

Pseudoproxy experiments are used to evaluate the performance of statistical reconstruction methods used to estimate climate fields from paleoclimatic proxies during the Common Era. These experiments employ output from millennial simulations from coupled General Circulation Models (GCMs) as test beds for controlled and systematic methodological investigations. It is demonstrated that multiple published pseudoproxy studies have used erroneously processed GCM surface temperature fields: the NCAR CCSM 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. The consequences of these problems are evaluated for the pseudoproxy studies in which the incorrect fields were used. Some quantitative results are invalidated by the findings, but many can be partially salvaged by reinterpretations. For those experiments that can be reinterpreted, 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 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. Output from the GKSS ECHO-g FOR1 (González-Rouco et al. 2003, hereinafter ECHO-g) and NCAR CCSM 1.4 (Ammann et al. 2007, hereinafter CCSM) millennial simulations is commonly used for this purpose. The motivation of this note is to demonstrate that several 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, hereinafter M05), Mann et al. (2007a, hereinafter 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 used in M05. Both of these papers used the incorrectly oriented CCSM field obtained from the M05 supplemental website. We have also 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), the latter of whom explained that the inconsistency was caused by an interpolation of the original ECHO-g field to a $5^\circ \times 5^\circ$ 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 must be either reinterpreted or discarded.

2. Model Fields as Used in Pseudoproxy Studies

The CCSM simulation was used by both M05 and M07, while the ECHO-g simulation was used only by M07. In each study the model fields comprise mean annual surface air temperatures (provided as anomalies for the CCSM field and as absolute temperatures for the ECHO-g field) interpolated from their original model resolutions to $5^\circ \times 5^\circ$ gridded fields. These global fields are the underlying data from which the masked instrumental field and the pseudoproxy networks were sampled in the M05 and M07 experiments. The model data and pseudoproxy networks are publicly available at the M05 and M07 supplemental websites http://fox.rwu.edu/~rutherfo/supplements/Pseudoproxy05/ and http://www.meteo.psu.edu/~mann/PseudoproxyJGR06/, respectively.

a. CCSM 1.4 data set

We plot in Figure 1a the standard deviations of the CCSM temperature anomalies calculated from 1880-1980 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. Not only are there no obvious dynamical explanations for
this feature, it also shows no appreciable changes as it crosses the South American coastline.

Figure 1c presents the same statistics as Figure 1a, but computed from the eight-member ensemble of CCSM3 simulations (Collins et al. 2006) from the IPCC AR4 “Climate of the 20th Century” experiment archived at the Program for Climate Model Diagnosis and Intercomparison website. Comparisons between these panels make clear that the aforementioned equatorial feature in Figure 1a is supposed to be in the Equatorial Pacific and is due to the El Niño – Southern Oscillation (ENSO) phenomenon. Further comparison of Figure 1a and 1c identifies the problem with the 1a orientation: it has been transformed from the original $0^\circ$–$360^\circ$ model domain (Figure 1b) to the $-180^\circ$–$180^\circ$ domain of Figure 1a using an incorrect transformation of longitudes $x$ to $x'$ in the form $x' = 180^\circ - x$. This transformation establishes the mirror symmetry of patterns in Figures 1a (incorrect orientation) and 1b (correct orientation) with respect to a vertical axis drawn equidistantly between them. When the correct transformation ($x' = x$ for $\leq 180^\circ$ and $x' = x - 360^\circ$ for $x > 180^\circ$) is applied to Figure 1b, it produces a standard deviation pattern (Figure 1d) that is broadly similar to the CCSM3 patterns shown in Figure 1c, despite the differences in the model versions and the applied forcings. These similarities confirm that panel 1d, not 1a, is correct.

b. *ECHO-g FOR1 data set*

Figure 2 presents the mean and standard deviation fields from the ECHO-g simulation archived at the M07 supplementary website and from the correct version acquired directly from J. F. González-Rouco (personal communication). A comparison between the two versions shows that the Eastern Hemispheres (EH) in the two fields are similar. A unrealistic
smoothing in the entire WH of the M07 field is evident, however, and causes a false spatial coherency throughout the hemisphere.

3. Implications for Pseudoproxy Experiments

The problems associated with the M05 and M07 versions of the CCSM and ECHO-g model fields were not simply archiving errors. Smerdon and Kaplan (2007) successfully reproduced the M05 results using the M05 archived CCSM field that employed the incorrect geographic orientation. Figure 6 in M07 also shows maps of model and reconstructed surface temperature means that have the incorrect orientation. Similarly, Figure S19 in M07 presents maps of skill statistics with equatorial minima in rectangular regions collocated with the displaced ENSO area of enhanced variability in Figure 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 the pseudoproxy experiments performed by M05 and M07 used the erroneous model fields as described in Section 2.

a. Instrumental mask and pseudoproxy sampling

The pseudoproxy experiments in M05 and M07 were intended to reflect patterns of instrumental and proxy data availability as shown in Figure 3a. Because of the incorrect orientation of the CCSM model field, however, the pseudo-instrumental and pseudoproxy data were actually taken from the locations shown in Figure 3b. The instrumental data mask incorrectly
excluded data from regions such as Northern Europe and included data from regions such as Australia. Pseudoproxy sampling of the field was also critically changed. Presumed dense sampling over Europe, for instance, was instead located over parts of 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 here using RegEM-Ridge reconstructions (Schneider 2001) and pseudoproxies with a signal-to-noise ratio of 0.5 (by standard deviation). All reconstruction settings are analogous to those of M05, except that global and non-hybrid reconstructions are used. 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 Figure 4. Figure 4a corresponds to the case when 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 4a, but with a corrected geographic orientation. As a result of this reinterpretation, areas of high skill move to the places where pseudoproxies were actually sampled (cf. Figure 4b with Figure 3b).

Figure 4c plots the correlation field for the RegEM-Ridge reconstruction using the correctly oriented CCSM field and the correct sampling distribution in Figure 3a. Similar to Figure 4a and 4b, the areas of highest correlation are located where the pseudoproxy network is the densest, but in this case these regions are correctly located geographically and correspond to the correct grid points in the model field. Figure 4d presents the results for the
same pseudoproxy experiment as the one shown in Figure 4c, but for the correct ECHO-g model field. Broad similarities are observed between the skill fields in the two model simulations: the highest correlations are still concentrated over North America and Europe. Reconstructions of the temperature fields simulated by ECHO-g have systematically higher skill in the Tropics (Figure 4d), however, than those targeting the CCSM field (Figure 4c). These differences are reflected in the global area-weighted averages of the correlation coefficients in Figure 4c (0.465) and 4d (0.487). The former is also slightly higher than that for the reconstruction using the incorrect sampling scheme in Figure 3b and the correctly oriented CCSM field (0.460).

c. Representation of the Northern Hemisphere Mean

Figure 5a plots the area-weighted NH means for the RegEM-Ridge reconstructions using the correct (Figure 3a) and incorrect (Figure 3b) distributions; the latter is analogous to the M05 and M07 results. While the two NH means are different, their skills do not differ by much: the correlation coefficient between the reconstructed and “true” model time series during the reconstruction interval are only slightly higher for the correct sampling distribution (0.756) than for the one resulting from the incorrect model orientation (0.715).

In the case of 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 representation of the index by M07. The correct version of the index is shown in Figure 5b, as well as the incorrect version presented by M07. Rutherford et al. (2008) indicated that the explanation of the discrepancy was that the
original ECHO-g field was interpolated to a $5^\circ \times 5^\circ$ 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 $5^\circ \times 5^\circ$ 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 present in their ECHO-g field, and implied that the differences were simply the result of two different interpolation choices. But the NH mean index recalculated here from the corrupted ECHO-g field proved to be identical to the time series published by M07 (Figure 5b). The true problem with the M07 representation of the ECHO-g NH mean index was thus the hemispheric-scale smoothing of the WH reported herein, not one legitimate interpolation choice over another.

4. Discussion and Conclusions

We have identified problems with the publicly available M05 and M07 versions of the model fields that were used in these and other published studies. The NCAR CCSM millennial integration was incorrectly oriented geographically and caused both visual misrepresentations of the field and spatial sampling that did not reflect the actual availability of instrumental or proxy data. We have also found that the ECHO-g model field was corrupted more than previously reported (Smerdon et al. 2008b; Rutherford et al. 2008). A large-scale and unphysical spatial smoothing in the WH is present in the ECHO-g data employed by M07 that made pseudoproxy experiments based on this field unrealistic.

Because of the errors reported herein, the quantitative results of all pseudoproxy experi-
ments based on these fields either require reinterpretation or must be discarded. 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 reinterpreted as reflecting results for the sampling distribution represented in Figure 3b. Results based on this sampling distribution cannot be directly compared to future experiments that employ distributions correctly reflecting real-world conditions (e.g., Figure 3a). The M07 results from the CCSM pseudoproxy experiments that cannot be reinterpreted are the statistics reported for the Niño3 index: instead of reconstructing the temperature index of the Eastern Equatorial Pacific (Figure 3a, 4c), M07 reconstructed an index of land and ocean temperatures spanning the region shown in Figures 3b and 4b.

No pseudoproxy experiments based on the ECHO-g field containing the corrupted WH can be usefully reinterpreted. This field is unphysical and has no clear analog for realistic reconstructions. It is also unclear whether the updated reconstructions in Rutherford et al. (2008) corrected the hemispheric-scale smoothing in the WH because that feature of the field was not discussed.

Smerdon and Kaplan (2007) and Smerdon et al. (2008a) used the incorrectly oriented CCSM field provided at the M05 supplemental website to reproduce selected M05 reconstructions and compare them with reconstructions derived using different parameter values. These comparisons illustrated how the use of data prior to the instrumental period, which cannot be used for standardization in realistic pseudoproxy experiments, 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 pre-instrumental 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 Figure 3b. The maps in Figures 6, 7, and 9 from Smerdon et al. (2008a) also must be transformed in the same way that Figure 4a was transformed to Figure 4b (or Figure 1a to 1d). While these changes are necessary for the correct interpretation of 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 will provide insights into the source of skill associated with specific methods in real-world reconstructions, but their success is dependent on consistent experimental settings that include the correct model field representations and common sampling distributions. Toward such ends, the correct 5°×5° annual surface temperature fields from the ECHO-g FOR1 and CCSM 1.4 millennial integrations are provided at the following website: http://www.ldeo.columbia.edu/~jsmerdon/2010_jclim_supplement.html

Acknowledgments.

This research was supported by the National Science Foundation (grant ATM-0902436 to JES and AK) and the National Oceanic and Atmospheric Administration by grants NA07OAR4310060 and NAOAR4320912 to JES and AK, the latter under the Cooperative Institute for Climate Applications Research. We thank the PCMDI for providing access to the CCSM3 output and Dr. J. F. González-Rouco for providing the correct ECHO-g
field.
REFERENCES


Mann, M. E., S. Rutherford, E. Wahl, and C. Ammann, 2007a: Robustness of


List of Figures

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°–180°; (b) the correct orientation of version 1.4, 0°–360°; (c) version 3.0, IPCC AR4 “Climate of the 20th Century” (20C3M) 8-member ensemble; and (d) the correct orientation of version 1.4, -180°–180°.

2. ECHO-g FOR1 surface temperature statistics from 1880-1980 C.E. for the incorrect (M07) and correct (J. F. González-Rouco, 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.

3. Instrumental and proxy locations for (a) the intended sampling distribution approximating the Jones et al. (1999) instrumental locations and the Mann et al. (1998) proxy locations; and (b) the actual locations sampled in the archived M05 and M07 data sets. Solid and dashed boxes in panels (a) and (b) enclose the correct and incorrect Niño3 regions, respectively.
Reconstruction-interval correlation coefficients between the RegEM-Ridge reconstructions and model “truth” using the CCSM and ECHO-g fields, pseudo-proxies with a signal-to-noise ratio of 0.5, a calibration interval from 1856-1980 C.E., and reconstruction intervals of 850-1855 and 1001-1855 C.E., respectively. Panel (a) displays the correlation field for the reconstruction using the incorrectly oriented CCSM field that was assumed to be correct in the M05 and M07 cases. Panel (b) plots the correct reinterpretation of panel (a) as the reconstruction skill for the unrealistic sampling distribution in Figure 3b. Panels (c) and (d) plot the correlation fields for the sampling distribution shown in Figure 3a for the correct CCSM and ECHO-g fields, respectively. Solid and dashed boxes in panels (b) and (c) enclose the incorrect and correct Niño3 regions, respectively.

(a) Area-weighted NH means derived from RegEM-Ridge reconstructions using the two sampling distributions shown in Figure 3 and the correctly oriented CCSM field. (b) Area-weighted NH means calculated for the correct ECHO-g field and the corrupted field archived by M07. Also shown is the NH mean published and archived by M07.
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°–180°; (b) the correct orientation of version 1.4, 0°–360°; (c) version 3.0, IPCC AR4 “Climate of the 20th Century” (20C3M) 8-member ensemble; and (d) the correct orientation of version 1.4, -180°–180°.
Fig. 2. ECHO-g FOR1 surface temperature statistics from 1880-1980 C.E. for the incorrect (M07) and correct (J. F. González-Rouco, 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.
Fig. 3. Instrumental and proxy locations for (a) the intended sampling distribution approximating the Jones et al. (1999) instrumental locations and the Mann et al. (1998) proxy locations; and (b) the actual locations sampled in the archived M05 and M07 data sets. Solid and dashed boxes in panels (a) and (b) enclose the correct and incorrect Niño3 regions, respectively.
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 1856-1980 C.E., and reconstruction intervals of 850-1855 and 1001-1855 C.E., respectively. Panel (a) displays the correlation field for the reconstruction using the incorrectly oriented CCSM field that was assumed to be correct in the M05 and M07 cases. Panel (b) plots the correct reinterpretation of panel (a) as the reconstruction skill for the unrealistic sampling distribution in Figure 3b. Panels (c) and (d) plot the correlation fields for the sampling distribution shown in Figure 3a for the correct CCSM and ECHO-g fields, respectively. Solid and dashed boxes in panels (b) and (c) enclose the incorrect and correct Niño3 regions, respectively.
Fig. 5. (a) Area-weighted NH means derived from RegEM-Ridge reconstructions using the two sampling distributions shown in Figure 3 and the correctly oriented CCSM field. (b) Area-weighted NH means calculated for the correct ECHO-g field and the corrupted field archived by M07. Also shown is the NH mean published and archived by M07.