Smoothing of climate time series revisited

Michael E. Mann[1]

Received 16 May 2008; revised 30 June 2008; accepted 25 July 2008; published 29 August 2008.

[1] We present an easily implemented method for smoothing climate time series, generalizing upon an approach previously described by Mann (2004). The method adaptively weights the three lowest order time series boundary constraints to optimize the fit with the raw time series. We apply the method to the instrumental global mean temperature series from 1850–2007 and to various surrogate global mean temperature series from 1850–2100 derived from the CMIP3 multimodel intercomparison project. These applications demonstrate that the adaptive method systematically out-performs certain widely used default smoothing methods, and is more likely to yield accurate assessments of long-term warming trends. **Citation:** Mann, M. E. (2008), Smoothing of climate time series revisited, *Geophys. Res. Lett.*, *35*, L16708, doi:10.1029/2008GL034716.

1. Introduction

[2] Climate time series often exhibit non-stationary statistical behavior, including long-term trends that are associated with exogenous forcing such as increasing greenhouse gas concentrations. This behavior presents problems for many traditional smoothing methods which assume stationary time series behavior. Particularly problematic in this context is the issue of how best to smooth time series with trends near the boundaries, where the smoothing window runs into the ends of the series. These considerations are non-trivial, because the smoothing of a time series near its boundaries is not uniquely defined; the behavior of the 'smooth' is dependent in part on the unknown behavior of the time series outside the interval for which data are actually available.

[3] Such issues have been examined in the past geosciences literature [e.g., *Park*, 1992; *Ghil et al.*, 2002; *Mann*, 2004; *Arguez et al.*, 2008]. *Mann* [2004] described a specific method for dealing with the complications of non-stationary behavior in the context of climate time series smoothing. The method was motivated by previously described 'inverse theory' based approaches to time series smoothing [*Park*, 1992] which treat assumptions regarding the unknown behavior of the time series outside the available interval as an additional constraint which must be imposed to yield a unique smooth of the time series. The three lowest-order possible such constraints are the so-called 'minimum norm' constraint, which tends to minimize the amplitude of the smooth near the boundaries, the 'minimum slope' constraint which tends to minimize the slope of the smooth near the boundaries, and the 'minimum roughness' constraint which seeks the smoothest behavior near the time series boundaries by tending to minimize the second derivative of the smooth near the boundaries. These constraints can be implemented exactly using the frequency domain approach described by *Park* [1992]. These constraints can also be implemented approximately as done by *Mann* [2004] which simulates the three constraints using simple procedures for padding the ends of the time series (padding the time series with climatology approximates the 'minimum norm' constraint, reflecting the time series horizontally about the final data point approximates the 'minimum slope' constraint, and reflecting the time series horizontally and vertically about the final data point approximates the 'minimum roughness' constraint). *Mann* [2004] argues that the constraint among these three which optimizes some objective metric of goodness-of-fit (e.g., which minimizes the mean-square error ('MSE') with respect to the raw time series), can be motivated as being the most appropriate for the particular time series and time interval being analyzed. In fact, this choice is a simplification of a more general adaptive approach (e.g., as used by *Mann and Park* [1996] in a multivariate setting, and *Ghil et al.* [2002] for the univariate problem of time series smoothing) where one instead selects as an objective choice that linear combination of the three above constraints which minimizes the MSE with respect to the raw time series.

[4] Often, much simpler boundary conditions are used. In the most recent Intergovernmental Panel on Climate Change (IPCC) assessment, for example, either a restricted version of the *Mann* [2004] time series smoothing approach wherein the 'minimum slope' condition alone is employed [*Trenberth et al.*, 2007], or an alternative 'mean padding' approach (where the series is padded with the mean over the final $N/2$ values where N is the width (in # of samples) of the smoothing window used [*Jansen et al.*, 2007]) was used. *Trenberth et al.* [2007] discuss the conservative nature of their constraint choice, noting specifically that it is prone to underestimate trends near the time series endpoints (this is even more the case with the 'mean padding' approach used by *Jansen et al.* [2007]). While the use of overly conservative approaches might be appropriate in the context of assessment reports which require a consensus among a large number of authors with differing views and preferences, we believe that the demonstrated performance of the method should in general be the guiding consideration.

[5] Here we demonstrate that the approaches used in the most recent IPCC assessment (the 'minimum slope' and 'mean padding' approaches) are indeed overly conservative, tending to understate the amplitude of long-term trends by artificially suppressing trends near the boundaries of the time series. This finding is established both through application of competing smoothing methods to interior intervals

[1]Department of Meteorology and Earth and Environmental Systems Institute, Pennsylvania State University, University Park, Pennsylvania, USA.

Copyright 2008 by the American Geophysical Union.
0094-8276/08/2008GL034716

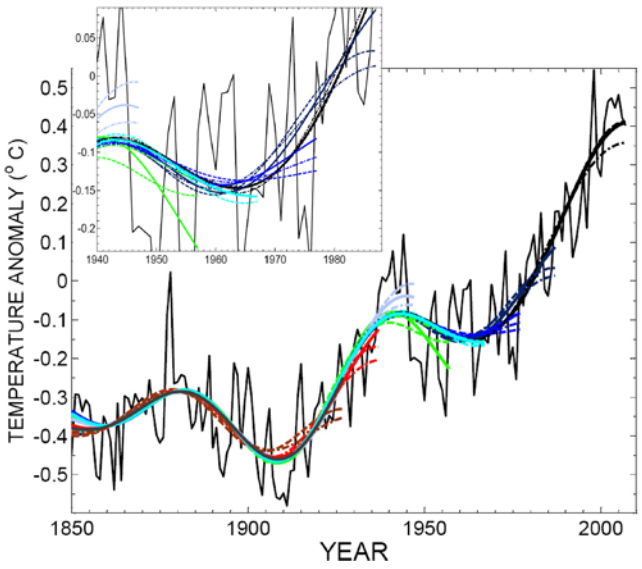

Figure 1. Instrumental annual global mean surface temperatures 1850–2007 smoothed on 40 year and longer timescales based on the 'adaptive' (solid), 'minslp' (dashed) and 'mean padded' (dot-dashed) approaches discussed in the text. Results are shown for the full 1850–2007 interval (black) and the various truncated intervals discussed in the text ending at 1987 (purple), 1977 (blue), 1967 (cyan), 1957 (green), 1947 (gray), 1937 (red), and 1927 (auburn). Inset focuses on late 20th century sub-interval.

of the observational series for which both the estimated and true smoothed behavior can be evaluated, and using simulated global mean temperature series taken from the multi-model suite of climate model simulations generated by the CMIP3 model intercomparison project. Our analysis instead motivates the use of an adaptive smoothing approach presented below, which treats the selection of time series smoothing boundary constraints as an empirical optimization problem.

2. Methods

[6] We employ a generalized 'adaptive' version of the method described by *Mann* [2004]. Rather than choosing a single best constraint, the adaptive approach seeks an optimal combination $S_0(t)$ of the three individual smooths $S_j(t)$ of time series $T(t)$ that result from application each of the three lowest-order (see section 1) boundary constraints *viz*.

$$S_0(t) = w_1 S_1(t) + w_2 S_2(t) + w_3 S_3(t) \quad (1)$$

$S_1(t)$, $S_2(t)$ and $S_3(t)$ represent the minimum smooth, slope, and roughness solutions respectively, and the coefficients w_j are chosen such that $\Sigma\,[S_0(t) - T(t)]^2$ is minimized, subject to the constraints $w_j \in [\mathbf{0}, \mathbf{1}]$ and $\Sigma w_j = 1$. The approach is thus free to choose the limiting cases of 'minimum norm' ($w_1 = 1, w_2 = w_3 = 0$), 'minimum slope' ($w_2 = 1, w_1 = w_3 = 0$), and 'minimum roughness' ($w_3 = 1$, $w_1 = w_2 = 0$), constraints, as well any mixture of those three constraints [*Mann and Park*, 1996; *Ghil et al.*, 2002]. As done by *Mann* [2004], we employ a 'Matlab' routine which uses a 10 point "Butterworth" lowpass filter of specified cutoff (half power) frequency f_0 for time series smoothing.

[7] The rough half-width of the smoothing window associated with the filter is $\Delta\tau = 1/f_0$. where f_0 is the cutoff (i.e., half-power drop-off) frequency. For example, if one wishes to retain variability on multidecadal timescales longer than 40 years using a filter with $f_0 = 0.025$ cycles/year, then $\Delta\tau \approx 40$ years and impacts of boundary conditions can be expected to influence the behavior of the smooth within $\Delta\tau = 40$ years (especially within $\Delta\tau/2 = 20$ years) of the time series boundaries. In practice, the influence can penetrate somewhat farther into the interior of the time series owing to the phenomenon of spectral leakage. The length of the series padding used here to implement boundary constraints has been increased from $\Delta\tau/2$ (used by *Mann* [2004]) to $3\Delta\tau/2$ to further guard against time series end effects. For simplicity, the simpler constraint procedure of *Mann* [2004] is used at the early end of the time series, while the adaptive constraint is applied to the late end whose behavior is of primary interest. The routine and required sub-routines are available at: http://www.meteo.psu.edu/~mann/smoothing08 (Matlab 'signal' toolbox required).

[8] Before proceeding, we used an example (the instrumental global mean temperature series used in section 3.1 below, smoothed on the multidecadal timescale $f_0 = 0.025$ cycles/year) to compare the Matlab smoothing procedure described above with an alternative version of the procedure based on multiple taper spectral analysis which implements the three lowest order boundary conditions exactly, and which employs a maximally spectral leakage-resistant filtering technique [see *Mann and Park*, 1996; *Ghil et al.*, 2002; also *Mann*, 2004]. Both of these implementations of the adaptive boundary constraint technique are found to yield remarkably similar smooths of the time series (see Figure S1 of the auxiliary material).[1] The Matlab smoothing procedure, though not optimal in terms of its statistical properties, thus gives results quite similar to more sophisticated and precise approaches, and offers the advantage of being less computationally cumbersome and easily adaptable (any other arbitrary low-pass filter available in Matlab can be substituted for the 'Butterworth' filter used in our routine). In this sense, we present the Matlab smoothing procedure more as a simple 'proof of concept' potentially to be improved upon, than a prescriptive approach for time series smoothing.

3. Applications

3.1. Instrumental Global Mean Surface Temperature Record

[9] We analyzed the annual 'HadCRUT3' global annual mean combined land + ocean instrumental surface temperature series from 1850–2007 [*Brohan et al.*, 2006] (available at http://www.cru.uea.ac.uk/cru/data/temperature). Both our adaptive smoothing method (henceforth termed '*adaptive*') and the mandated use of the 'minimum slope' constraint (henceforth termed '*minslp*') suggest (Figure 1) a net warming of roughly 0.88°C applied to the full series, while

[1] Auxiliary materials are available in the HTML. doi:10.1029/2008GL034716.

the 'mean padded' approach (henceforth termed '*mpad*') suggests a slightly more modest 0.83°C warming. The differences between the approaches are modest in this case, largely because the 'adaptive' method essentially chooses the pure 'minslp' constraint limit owing to the recent flattening of warming. However, the differences are considerably greater in other instances, as discussed below.

[10] In order to objectively evaluate the relative performance of the two different approaches in capturing the true smoothed behavior near the boundaries more generally, we performed additional smoothing experiments where the data interval was truncated at least $\Delta\tau/2 = 20$ years prior to the 2007 time series boundary (separate experiments were thus performed with termination dates of 1987, 1977, 1967, 1957, 1947, 1937, and 1927). In such cases, the true smoothed behavior of the time series at the termination date is known, because that date is far enough into the interior of the full series that its smooth at that point is largely insensitive to the constraint on the upper boundary. The relative skill of the different methods can then be measured by the misfit between the estimated and true smooths of the truncated series. We define this skill in terms of the ratio of variances in the misfit between the estimated and true smooth and the smooth itself (this is expressed as a fraction or percentage and is denote by "ε^2"). We removed the first $\Delta\tau = 40$ years from the error calculations to eliminate any influence of the lower boundary constraint on the smooth. Averaged over the 7 test intervals described above, the 'adaptive' method (see Table S1) yields an average relative misfit score of 1.0% while the 'minslp' method yields a score of 4.9% and the 'mpad' approach a score of 6.0%, factors of roughly five and six times greater relative error respectively. The 'adaptive' method out-performed the 'minslp' method for 6 of the 7 test intervals cases (the 'minslp' approach wins for the interval ending in 1957) while it outperformed the 'mpad' approach slightly less often (4 of the 7 test intervals) but typically more decisively. The optimal constraints selected by the 'adaptive' method vary, of course, over time. For example, a pure (or near pure) 'minimum roughness' constraint—consistent with assumption of a steady trend extending through the end of the series—is selected for terminal dates between 1999 and 2005 (see Figure S2), noted for example in the previous analysis by *Mann* [2004]. By contrast a nearly pure 'minimum slope' constraint (consistent with a persistence assumption) is selected for terminal dates in the early 1960s when a flattening of the warming trend is observed and, as noted above, for the most recent 2007 terminal date. Unfortunately, the relative performance of the different smoothing methods cannot be evaluated for the present day because the true multidecadal smoothed behavior cannot be defined through the present. However, we can turn to model simulations as surrogates for plausible 21st century global temperature scenarios, as discussed in the next section.

3.2. CMIP3 Multimodel Global Mean Surface Temperature Series

[11] We next analyzed the global mean temperature series from a total of 55 individual realizations of 21 different state-of-the-art model simulations from the World Climate Research Programme's (WCRP's) Coupled Model Intercomparison Project phase 3 (CMIP3) (see http://www-pcmdi.llnl.gov/ipcc/about_ipcc.php). For the purpose of our analyses, we combined corresponding realizations of the historical simulations (which were driven by historical estimated natural and anthropogenic radiative forcings from the late 19th through late 20th century) and the anthropogenic forced experiments from 2000–2100 resulting from the intermediate 'A1B' anthropogenic emissions scenario. These series were treated as 55 plausible surrogate 1850–2100 global mean surface temperature series.

[12] We restricted the analyses to the interval from 1900 forward during which all simulations are available. The 'true' smooth is defined by smoothing the model series through the late 21st century upper boundary (using the 'adaptive' smoothing method, but the series prior to the mid 21st century are insensitive to which boundary constraint is used). We performed smoothing tests similar to those outlined in section 3.1, but instead using seven truncated intervals with termination dates of: 1997, 2002, 2007, 2012, 2017, 2022 and 2027. These choices allow us to evaluate the relative performance of the two different smoothing methods under comparison for hypothetical termination dates ranging from the past decade through the next two decades.

[13] The 'adaptive' approach outperformed the 'minslp' approach for all but 75 of the 385 (55 models $\times$ 7 time intervals) cases considered (i.e., 81% of the time), similar to the 6/7 (85%) 'win' rate found in section 3.1 for the observed modern record. The mean error over all 385 cases was 0.8% with the 'adaptive' smoothing approach, and 1.5% with the 'minslp' approach. The adaptive method outperformed the 'minslp' method for each of the seven time intervals considered, averaged across the 55 different CMIP3 simulations (Table S2). The ratio of 'minslp' to 'adaptive' error variance ranged from 1.5 to 2.8, with the median and means both equal to 2.2. For nearly all (47 of 55) of the CMIP3 simulations, the 'adaptive' method outperformed the 'minslp' method averaged across the 7 time intervals (Figure 2; full details provided in Table S3), with the ratio of error variance ranging from 0.6 to 12.2 (median of 1.6 and mean of 2.4).

[14] Similar tests using the 'mpad' approach yielded even poorer performance than the 'minslip' approach (see Tables S2 and S3). The 'mpad' smoothing method was outperformed by the 'adaptive' method for 54 of the 55 simulations tested (Figure 2), with the error variance relative to the 'adaptive' approach ranging from 0.94 to 17.2 for the 55 simulations with median and mean of 3.3 and 4.5 respectively.

[15] It is clear upon visual inspection that the 'minslp' and (especially) 'mpad' methods tend to significantly underestimate the true trends near the time series upper boundary, while the 'adaptive' method reproduces those trends relatively faithfully. Specific examples (Figure 3) are shown both for a case where the underperformance relative to the 'adaptive' approach is quite large (simulation #34 in Table S2: a simulation of the MIUB ECHO-G coupled model) and a case where the underperformance is close to the mean over the ensemble of 55 models (simulation #10 in Table S2: a simulation of the CSIRO MK3.5 coupled model). In both cases, it is noteworthy that the 'minslp' ('mpad') methods typically underestimate the true long-

term trends by as much as 0.1°C (0.2°C). These are not trivial underestimates, given that the net 20th century warming is less than 1°C. The 'adaptive' method by contrast yields either little or no systematic underestimate of trend.

[16] Additional experiments using a *decadal* smoothing filter similar to that used by *Trenberth et al.* [2007] (cutoff frequency of $f_0 = 0.0625$ corresponding to 16 year period) yielded considerably more modest differences, with the mean and median error variance inflation relative to the 'adaptive' method considerably smaller for both the 'minslp' (mean of 1.7, median of 1.2) and 'mean padding' (mean of 1.7, median of 1.02) averaged across the 55 simulations. We conclude that the performance issues examined in this study are of primary relevance to comparisons of series smoothed on *multidecadal* [e.g., *Jansen et al.*, 2007; *Folland et al.*, 2001] rather than decadal [e.g., *Trenberth et al.*, 2007; *Rahmstorf et al.*, 2007] timescales.

4. Conclusions

[17] Our analyses indicate that some commonly used approaches for climate time series smoothing are likely to underestimate the amplitude of long-term changes due to artificial suppression of trends near the time series boundaries. Methods such as the adaptive boundary constraint approach presented here, which have the ability to preserve trends near time series boundaries, should instead be used. This is particularly true where non-stationary behavior, such as is associated with the response to anthropogenic forcing, is likely to be important.

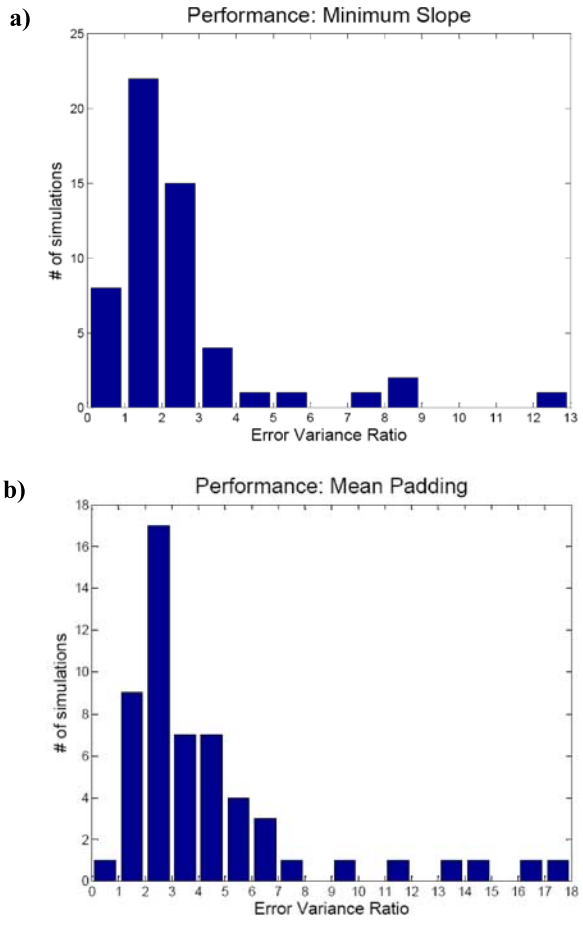

Figure 2. Distribution of the ratio of error variance of the (a) 'minimum slope' and (b) 'mean padding' smoothing approaches relative to the favored 'adaptive' smoothing approach for the 55 different CMIP3 simulations, averaged across the seven test intervals described in the text.

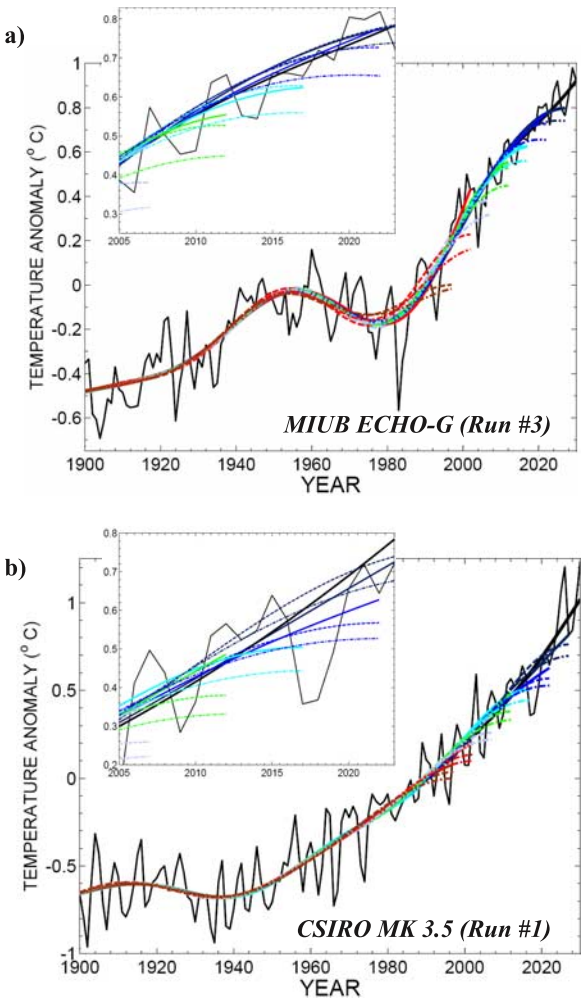

Figure 3. Global mean surface temperatures 1900–2027 for (a) MIUB ECHO-G (run #3) coupled model simulation and (b) CSIRO MK 3.5 (run #1) coupled model simulation, smoothed on 40 year and longer timescales based on the 'adaptive' (solid curves), 'minimum slope' (dashed) and 'mean-padded' (dot-dashed) approaches discussed in the text. Results are shown for smooths based on the full available interval ending at 2100 (black) and on the various truncated intervals discussed in the text ending at 2027 (purple), 2022 (blue), 2017 (cyan), 2012 (green), 2007 (gray), 2002 (red), and 1997 (auburn). Inset focuses on early 21st century sub-interval.

[18] **Acknowledgments.** We acknowledge the modeling groups, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) and the WCRP's Working Group on Coupled Modeling (WGCM) for their roles in making available the WCRP CMIP3 multi-model dataset. Support of this dataset is provided by the Office of Science, U.S. Department of Energy. This paper benefited from helpful comments from Gavin Schmidt, Stefan

Rahmstorf, Kevin Trenberth and an anonymous reviewer. We gratefully acknowledge technical assistance from Sonya Miller.

References

Arguez, A., P. Yu, and J. J. O'Brien (2008), A new method for time series filtering near endpoints, *J. Atmos. Oceanic Technol.*, *25*, 534–546.

Brohan, P., J. J. Kennedy, I. Harris, S. F. B. Tett, and P. D. Jones (2006), Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850, *J. Geophys. Res.*, *111*, D12106, doi:10.1029/2005JD006548.

Folland, C. K., T. R. Karl, J. R. Christy, R. A. Clarke, G. V. Gruza, J. Jouzel, M. E. Mann, J. Oerlemans, M. J. Salinger, and S.-W. Wang (2001), Observed climate variability and change, in *Climate Change 2001: The Scientific Basis*, edited by J. T. Houghton et al., pp. 99–181, Cambridge Univ. Press, New York.

Ghil, M., et al. (2002), Advanced spectral methods for climatic time series, *Rev. Geophys.*, *40*(1), 1003, doi:10.1029/2000RG000092.

Jansen, E., et al. (2007), Palaeoclimate, in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon et al., pp. 433–497, Cambridge Univ. Press, Cambridge, U. K.

Mann, M. E. (2004), On smoothing potentially non-stationary climate time series, *Geophys. Res. Lett.*, *31*, L07214, doi:10.1029/2004GL019569.

Mann, M. E., and J. Park (1996), Joint spatio-temporal modes of surface temperature and sea level pressure variability in the northern hemisphere during the last century, *J. Clim.*, *9*, 2137–2162.

Park, J. (1992), Envelope estimation for quasi-periodic geophysical signals in noise: A multitaper approach, in *Statistics in the Environmental and Earth Sciences*, edited by A. T. Walden and P. Guttorp, pp. 189–219, Edward Arnold, London.

Rahmstorf, S., A. Cazenave, J. A. Church, J. E. Hansen, R. F. Keeling, D. E. Parker, and R. C. J. Somerville (2007), Recent climate observations compared to projections, *Science*, *316*, 709.

Trenberth, K. E., et al. (2007), Observations: Surface and atmospheric climate change, in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon et al., pp. 235–336, Cambridge Univ. Press, Cambridge, U. K.

M. E. Mann, Department of Meteorology, Pennsylvania State University, University Park, PA 16802, USA. (mann@psu.edu)