On smoothing potentially non-stationary climate time series

Michael E. Mann

Department of Environmental Sciences, University of Virginia, Charlottesville, Virginia, USA

Received 23 January 2004; revised 10 March 2004; accepted 18 March 2004; published 15 April 2004.

[1] A simple approach to the smoothing of a potentially non-stationary time series is presented which provides an optimal choice among three alternative, readily motivated and easily implemented boundary constraints. This method is applied to the smoothing of the instrumental Northern Hemisphere (NH) annual mean and cold-season North Atlantic Oscillation (NAO) time series, yielding an objective estimate of the smoothed decadal-scale variations in these series including long-term trends. *INDEX TERMS:* 3309 Meteorology and Atmospheric Dynamics: Climatology (1620); 3394 Meteorology and Atmospheric Dynamics: Instruments and techniques; 4294 Oceanography: General: Instruments and techniques; 4594 Oceanography: Physical: Instruments and techniques. **Citation:** Mann, M. E. (2004), On smoothing potentially non-stationary climate time series, *Geophys. Res. Lett.*, *31*, L07214, doi:10.1029/2004GL019569.

1. Introduction

[2] Proper smoothing of climate time series, particularly those exhibiting non-stationary behavior (e.g., substantial trends late in the series) is essential for placing recent trends in the context of past variability. The smoothing of a time series can be posed as an inverse problem with non-unique boundary constraints [*Park*, 1992], for which additional objective considerations must be made to determine the behavior near the boundaries. Various different boundary constraints have recently been employed, for example, in the smoothing of the Northern Hemisphere mean temperature series [*Folland et al.*, 2001; *Mann and Jones*, 2003; *Soon et al.*, 2004]. The approach used by *Mann and Jones* [2003], as noted therein, employed a smoothing boundary constraint optimized to resolve the non-stationary late behavior of the time series in comparison with previously employed constraints involving e.g., the padding of the series with mean values after the boundary [*Folland et al.*, 2001; *Mann*, 2002; *Mann et al.*, 2003]. The approach used by *Mann and Jones* [2003] (which is incorrectly assumed by *Soon et al.* [2004] to be a 'wavelet' approach), is described in more detail in this study.

[3] The three lowest order boundary constraints that can be applied to a smooth [see, e.g., *Park*, 1992] involve the minimization near the boundaries of either: (1) the zeroth derivative of the smooth (yielding the 'smallest' or 'minimum norm' solution), (2) the 1st derivative of the smooth (yielding the 'minimum slope' constraint), and (3) the 2nd derivative of the smooth (yielding the smoothest or 'minimum roughness' solution). Application of constraint (1) favors the tendency of the smooth to approach the mean value (i.e., 'climatology') near the boundaries. Application of (2) favors the tendency of the smooth to approach a constant local value near the boundary. Application of (3) favors the tendency of the smooth to approach the boundary with a constant *slope*. The first two approaches will underestimate the behavior of the time series near the boundaries in the presence of a long-term trend, but the 3rd approach may lead to an extrapolation error in the presence of leverage by outliers near the boundaries. Without additional considerations, none of these three constraints can be favored on *a priori* grounds. An objective choice, nonetheless, can be motivated as that particular constraint of the three which minimizes some measure of misfit of the smooth with respect to the original time series.

[4] In this paper, we describe both time-domain and frequency-domain approaches to implementing each of these three alternative boundary conditions, and employ an objective measure of the quality of fit of the various candidate smooths.

[5] We note that while our focus here is on smoothing of time series, similar considerations can be applied to alternative statistical time series modeling such as change point analysis [*Tomé and Miranda*, 2004]. We provide applications to two relevant instrumental climate time series, the Northern Hemisphere (NH) annual mean series from 1856–2003 of *Jones et al.* [1999], and the cold-season North Atlantic Oscillation (NAO) times series of *Jones et al.* [1997] from 1825/26 to 1999/2000.

2. Method

[6] While constraints (1)–(3) can be applied explicitly in the frequency domain [e.g., *Park*, 1992; *Ghil et al.*, 2002], it is possible to implement reasonable approximations to these constraints in a simple manner in the time domain as follows: To approximate the 'minimum norm' constraint, one pads the series with the long-term mean beyond the boundaries (up to at least one filter width) prior to smoothing. To approximate the 'minimum slope' constraint, one pads the series with the values within one filter width of the boundary reflected about the time boundary. This leads the smooth towards zero slope as it approaches the boundary. Finally, to approximate the 'minimum roughness' constraint, one pads the series with the values within one filter width of the boundary reflected about the time boundary, and reflected vertically (i.e., about the "y" axis) relative to the final value. This tends to impose a point of inflection at the boundary, and leads the smooth towards the boundary with constant slope. Alternative approximate implementations of constraints (1)–(3) are of course possible.

[7] We first make use of a routine that we have written in the 'Matlab' programming language which implements constraints (1)–(3), as described above, making use of a 10 point "Butterworth" low-pass filter for smoothing; other

Copyright 2004 by the American Geophysical Union.
0094-8276/04/2004GL019569

filters can be substituted yielding similar results. Our Matlab routine is provided here: ftp://holocene.evsc.virginia.edu/pub/mann/Filter/lowpass.m [note that this routine requires access to the 'Matlab' time series analysis ('signal') toolbox]. For each of the three alternative smooths, the resulting mean-square error ('MSE') of the smooth is calculated as a fraction of the total variance in the series resolved by the smooth. That constraint providing the minimum MSE is arguably the optimal constraint among the three tested. MSE, which penalizes the mean-squared deviations, is only one possible measure of misfit. Alternative measures of misfit, particularly those which are relatively insensitive to outliers such as minimum absolute deviation ('MAD'), are also worthy of consideration. Our purpose here is not to favor any particular criterion, but rather, to emphasize the importance of employing *some* objective measure of misfit in determining the optimal smooth of a time series. It should be stressed that these criteria must be applied to smooths on the same timescale, and applied to identical intervals of the time series, for a meaningful comparison among alternative boundary constraints. Several of the instrumental NH series smooths compared by *Soon et al.* [2004], for example, were based on series with variable endpoints (from the year 2000 to 2003), precluding a meaningful comparison of alternative smoothing approaches.

3. Applications

3.1. Northern Hemisphere Mean Temperature Series

[8] We analyze the Northern Hemisphere mean annual surface temperature time series of the Climatic Research Unit (CRU) of the University of East Anglia [*Jones et al.*, 1999]. This series is updated continuously on a monthly basis at: http://www.cru.uea.ac.uk/ftpdata/tavenh2v.dat.

[9] We choose first, as in other recent analyses [*Folland et al.*, 2001; *Mann and Jones*, 2003; *Soon et al.*, 2004] to smooth the series on a 40 year and longer timescale (corresponding to a pass band boundary at f = 0.025 cycle/yr in the low-pass filter). As the MSE of the smooth is found to be insensitive to constraint choice (1)–(3) for the *early* boundary, we employ the 'minimum slope' constraint for the early boundary in all three cases, and examine the sensitivity with respect to the constraint on the *late* boundary of the series. Figure 1a compares the smooths for each of the three constraints. The 'minimum norm' constraint (MSE = 0.30) is observed to under-predict the late 20th century trend significantly, while application of the 'minimum slope' constraint (MSE = 0.23), the constraint effectively used by *Mann* [2002], *Folland et al.* [2001], and *Soon et al.* [2004], modestly underpredicts the late 20th century trend. Only the application of the 'minimum roughness' constraint (MSE = 0.21) used by *Mann and Jones* [2003] resolves the full 20th century trend. As application of this constraint indeed minimizes the mean-square error with respect to the three choices, it can be objectively favored among the three boundary constraints.

[10] A precise implementation of the boundary constraints (1)–(3) [see *Park*, 1992; *Ghil et al.*, 2002] can be provided in the frequency domain through multiple-taper methods [e.g., *Thomson*, 1982; *Percival and Walden*, 1993]. Routines to implement such an approach are available here: http://www.atmos.ucla.edu/tcd/ssa/. We compare the results

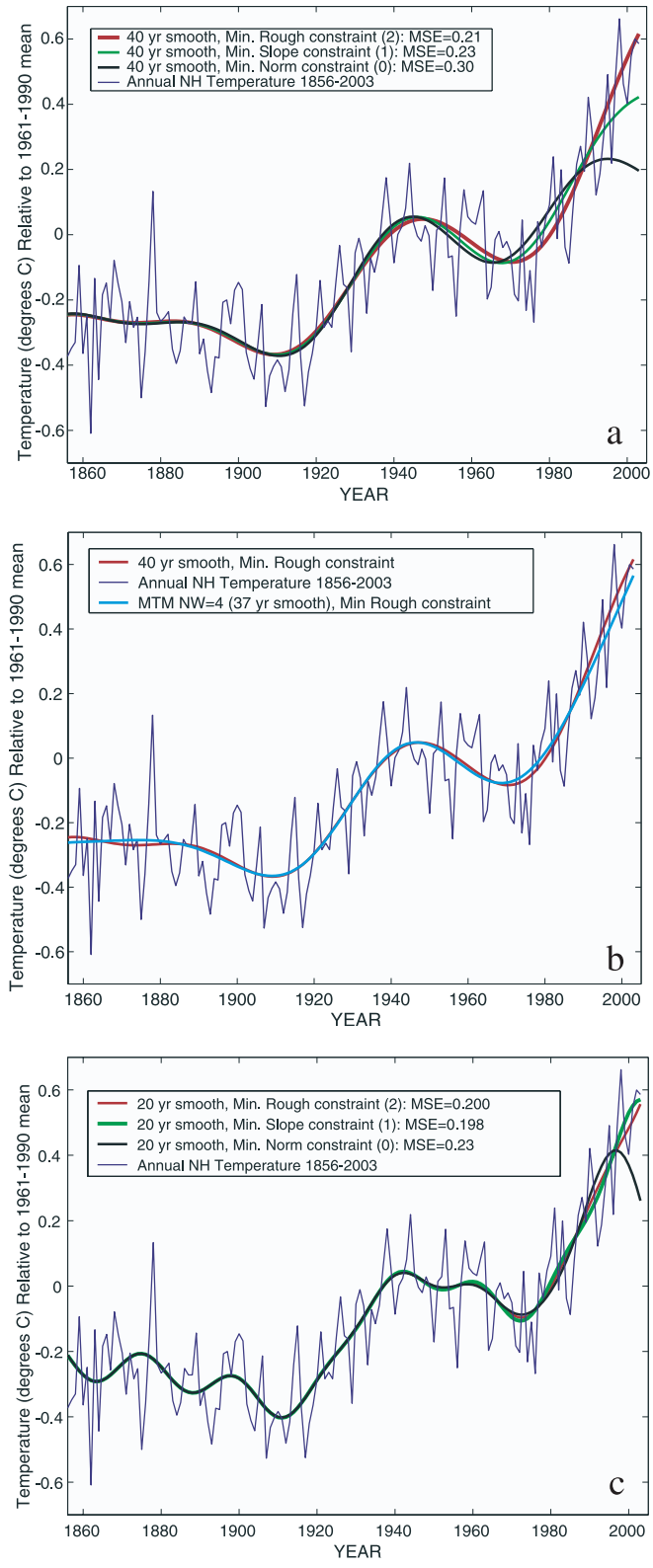

Figure 1. Annual mean NH series. (blue) shown along with (a) 40 year smooths of series based on alternative boundary constraints (1)–(3). Associated MSE scores favor use of the 'minimum roughness' constraint. (b) Comparison of 'minimum roughness' constraint from "a" with exact frequency-domain implementation of this constraint as described in text. (c) Same as "a" but employing a 20 year smoothing.

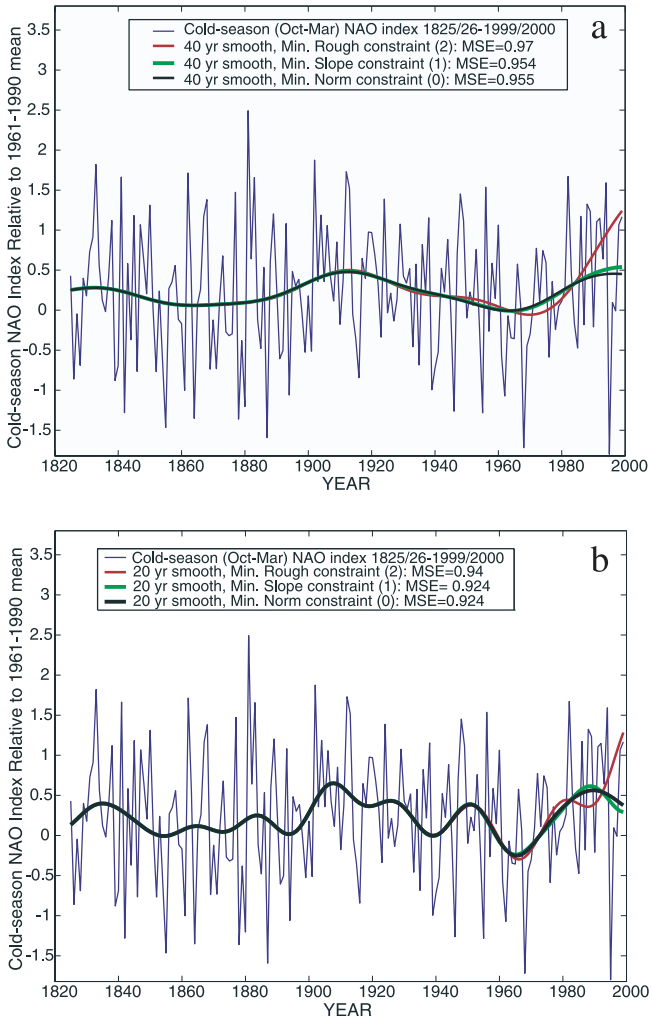

Figure 2. Cold season (Oct–Mar) mean NAO series. (a) Raw series shown along with 40 year smooths of series based on alternative boundary constraints (1)–(3). Associated MSE scores favor either the 'minimum norm' or 'minimum slope' constraint. (b) Same as "a" but employing a 20 year smoothing.

of application of the latter approach to those obtained above. A pass band boundary at approximately f = 0.025 cycle/year (40 year period) is afforded by the use a time-frequency bandwidth product of NW = 4 and 7 eigentapers in the multiple-taper approach [see, e.g., *Thomson*, 1982]. This yields an effective pass band boundary at f = 0.027 cycle/yr, corresponding to 37 year period. It is apparent that the approximate implementation of the minimum roughness constraint used earlier closely matches the result of the exact implementation of this constraint in the frequency domain (Figure 1b–note that the latter implementation employs the 'minimum roughness' constraint at *both ends* of the time series, leading to minor differences between the two smooths at the early end of the time series). The optimal nature of the smoothing of the instrumental NH temperature used by *Mann and Jones* [2003] is thus shown to be robust. For comparison, applications of constraints (1)–(3) are shown for the same series, applied to a 20 year, rather than 40 year, smooth (Figure 1c). In this case, both 'minimum roughness' and 'minimum slope' constraints yield similar results, and considerably lower MSE values than the 'minimum norm' constraint. Both the 20 year and 40 year optimal smooths suggest a net warming of roughly 0.8°C since the mid 19th century, with a roughly 1°C warming trend during the 20th century.

3.2. North Atlantic Oscillation (NAO) Series

[11] We provide an additional application of the procedure described above to the smoothing of the instrumental NAO series spanning nearly the past two centuries. The monthly series is available through CRU from 1821 through 1999 at http://www.cru.uea.ac.uk/ftpdata/nao.dat. We make use of the cold-season mean (Oct–Mar) series computed from the monthly data, which is available continuously from 1825/1826 through to 1999/2000. Application of constraints (1)–(3) to the smoothing of the series on 40 year and longer timescales yields the results shown in Figure 2a. In this example, the larger positive trend produced by use of the 'minimum roughness' constraint is likely spurious, as indicated by the larger MSE (0.97) relative to the application of 'minimum norm' and 'minimum slope' constraints, which have roughly the same MSE (0.955 and 0.954 respectively). We conclude that a positive trend over the past several decades, while nominally evident, is not clearly indicative of non-stationary behavior late in the series. Similar conclusions are drawn from a comparison of the applications of constraints (1)–(3) to a 20 year, rather than 40 year, smooth of the series (Figure 2b).

4. Conclusions

[12] We provide an easily implemented smoothing routine that yields objective estimates of the low-frequency variability of potentially non-stationary climate time series. The approximate implementation of the three most readily motivated boundary constraints ('minimum norm', 'minimum slope', and 'minimum roughness') closely reproduces the exact implementation of these constraints in the frequency domain. Applications of our approach to the NH annual mean temperature series demonstrates that an optimal 40 year smooth approaches the early 21st century boundary with a constant slope, suggestive of non-stationary behavior in the mean and a persistent positive trend late in the series. By contrast, application of the same procedure to the cold-season NAO index does not indicate clearly non-stationary late behavior. The analysis provided here also serves as a cautionary note with regard to some recently published comparisons of alternative time-domain smoothing boundary constraints to climate time series. Comparisons that are uninformed [e.g., *Soon et al.*, 2004] by objective evaluation criteria (e.g., MSE), are unlikely to provide useful insights into the relative merits of alternative boundary constraints. We thus urge the careful consideration of such criteria in choosing boundary constraints in the smoothing of climate time series, and warn against false conclusions based on unobjective statistical smoothing approaches.

[13] **Acknowledgments.** M.E.M. acknowledges support for this work by the NSF and NOAA-sponsored Earth Systems History (ESH) program (NOAA award NA16GP2913).

References

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., M. R. Allen, M. D. Dettinger, K. Ide, D. Kondrashov, M. E. Mann, A. W. Robertson, Y. Tian, F. Varadi, and P. Yiou (2002), Advanced spectral methods for climatic time series, *Rev. Geophys.*, *40*(1), 1003, doi:10.1029/2000RG000092.

Jones, P. D., T. Jónsson, and D. Wheeler (1997), Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibraltar and south-west Iceland, *Int. J. Climatol.*, *17*, 1433–1450.

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

Mann, M. E. (2002), The value of multiple proxies, *Science*, *297*, 1481–1482.

Mann, M. E., et al. (2003), On past temperatures and anomalous late 20th Century warmth, *Eos Trans. AGU*, *84*, 256–258.

Mann, M. E., and P. D. Jones (2003), Global surface temperatures over the past two millennia, *Geophys. Res. Lett.*, *30*(15), 1820, doi:10.1029/2003GL017814.

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.

Percival, D. B., and A. T. Walden (1993), *Spectral Analysis for Physical Applications*, Cambridge Univ. Press, New York.

Soon, W. W. H., D. R. Legates, and S. L. Baliunas (2004), Estimation and representation of long-term (>40 year) trends of Northern-Hemisphere gridded surface temperature: A note of caution, *Geophys. Res. Lett.*, *31*, L03209, doi:10.1029/2003GL019141.

Thomson, D. J. (1982), Spectrum estimation and harmonic analysis, *Proc. IEEE*, *70*, 1055–1096.

Tomé, A. R., and P. M. A. Miranda (2004), Piecewise linear fitting and trend changing points of climate parameters, *Geophys. Res. Lett.*, *31*, L02207, doi:10.1029/2003GL019100.

M. E. Mann, Department of Environmental Sciences, University of Virginia, Charlottesville, VA 22902, USA. (mann@virginia.edu)