Climatic Change (2011) 107:267–276
DOI 10.1007/s10584-010-9998-z

On long range dependence in global surface temperature series

An editorial comment

Michael E. Mann

Received: 8 September 2010 / Accepted: 12 November 2010 / Published online: 23 December 2010

Abstract Long Range Dependence (LRD) scaling behavior has been argued to characterize long-term surface temperature time series. LRD is typically measured by the so-called "Hurst" coefficient, "H". Using synthetic temperature time series generated by a simple climate model with known physics, I demonstrate that the values of H obtained for observational temperature time series can be understood in terms of the linear response to past estimated natural and anthropogenic external radiative forcing combined with the effects of random white noise weather forcing. The precise value of H is seen to depend on the particular noise realization. The overall distribution obtained over an ensemble of noise realizations is seen to be a function of the relative amplitude of external forcing and internal stochastic variability and additionally in climate "proxy" records, the amount of non-climatic noise present. There is no obvious reason to appeal to more exotic physics for an explanation of the apparent scaling behavior in observed temperature data.

1 Introduction

A number of past studies (e.g. Bloomfield and Nychka 1992; Smith 1993; Gil-Alana 2005; Eichner et al. 2003; Fraedrich and Blender 2003; Kiraly et al. 2006; Mills 2007) have investigated the possibility of "long-range dependence" ("LRD"–also called "long term persistence" or "long memory") in surface temperature records. Indeed, the existence of LRD has sometimes been invoked (e.g. Koutsoyiannis et al. 2008) in support of the proposition that long-term temperature trends are a manifestation of

Electronic supplementary material The online version of this article (doi:10.1007/s10584-010-9998-z) contains supplementary material, which is available to authorized users.

M. E. Mann (✉)
Department of Meteorology and Earth and Environmental Systems Institute,
Pennsylvania State University, University Park, PA, USA
e-mail: mann@psu.edu

intrinsic long-term natural variability of the climate system. Were that the case, the observed warming of the past century, rather than reflecting the response to human-caused increases in greenhouse gas concentrations, might instead just be the result of chance, long-term natural fluctuations. It is thus rather important to address the issue of whether or not there is support for interpreting global temperature fluctuations as the manifestation of LRD.

Most studies testing for the presence of LRD in time series focus on the so-called "Hurst Coefficient", or simply "H". H provides a measure of possible simple power law scaling of the power spectrum $S(f)$ with frequency f (sometimes referred to as "self similar" behavior),

$$S(f) \sim f^d \quad (1)$$

where the scaling exponent d is given by

$$d \equiv 1 - 2H \quad (2)$$

In the context of this particular model for stochastic time series behavior, $H = 0.5$ corresponds to a frequency-invariant spectrum, i.e. white noise. Values of $0.5 < H < 1$ corresponds to "LRD", implying more persistence in the data at the lowest frequencies than would be expected for white noise (this sort of persistence is sometimes called "red noise", though there are forms of red noise which do not require this sort of power law scaling and LRD). For $H \geq 1$, the persistence becomes so extreme that the time series exhibits non-stationary behavior. Long-term fluctuations about the mean are unbounded, though the time series always eventually reverts toward the mean. Such behavior contrasts sharply with other types of non-stationary behavior where substantial departures from the mean may instead be associated with deterministic forced, rather than stochastic behavior. An obvious example is the steady warming of the globe expected from continuing increases in greenhouse gas concentrations associated with anthropogenic activity.

Past studies have noted that estimates of H in temperature records vary substantially between land and ocean and more generally as a function of region (e.g. Blender and Fraedrich 2003; Fraedrich and Blender 2003; Kiraly et al. 2006) as well as altitude in the atmosphere (and Vyushin and Kushner 2009). A reasonable hypothesis is that this variation in H arises due to the spatially variable amplitude of externally-forced vs. internal stochastic climate variability. The analysis of e.g. global mean surface temperature arguably helps to avoid this particular complication. Further potential complications arise from the possible influence of systematic biases in temperature observations on estimates of H, particularly with respect to data/model "reanalysis" products (see Vyushin and Kushner 2009). Additional complications arise with the interpretation of "proxy" temperature estimates, which are based on natural archives of information, such as tree-rings, corals, ice cores, and ocean and lake sediments, that may contain non-climatic sources of noise and bias.

In this issue of *Climatic Change*, Rea and co-authors (Rea et al. 2011) introduce a novel approach for assessing evidence for LRD in an array of very long proxy temperature series. Their approach exploits the fact that true LRD and scaling behavior should be evident throughout distinct sub-intervals of a time series. Rea et al. consequently motivate the analysis of the stability of H over distinct objectively-determined sub-intervals of the time series to determined whether persistent behavior in a time series can in fact be interpreted in the context of an LRD process. Here, I attempt to complement their empirical approach with some physical context, using the results of experiments with a simple climate model to guide the interpretation of actual temperature observations.

2 Use of climate model simulations to assess LRD

A useful null hypothesis or 'straw man' for interpreting Hurst coefficients (H) estimated from real-world surface temperature data is provided by the analysis of the behavior of theoretical models of the climate system, which are governed by established underlying physics of the ocean, atmosphere, and climate system. Several past studies have assessed the properties of H in forced and control simulations of global three-dimensional coupled ocean-atmosphere models and model/data re-analysis products (e.g. Koutsoyiannis et al. 2008; Vyushin and Kushner 2009; Vyushin et al. 2009). While some of these studies have argued for discrepancies between observations and model simulation results (Koutsoyiannis et al. 2008), other studies have argued for consistency between them (Vyushin and Kushner 2009; Vyushin et al. 2009).

Purely speaking, H has an appropriate interpretation only in the context of purely stochastic variability. The presence of deterministic forcing arguably renders the assumptions behind even the estimation of H inappropriate. One might argue that some external climate forcing (e.g. natural forcing by solar irradiance variations and volcanic eruptions)—and hence the response to that forcing—could itself represent the realization of a stochastic underlying process. However, such an argument is difficult to make for the non-stationary forcing by anthropogenic increases in greenhouse gas concentrations. In this context, it is therefore difficult to cleanly interpret analyses of LRD in single realizations of forced coupled model simulations which reflect a combination of stochastic and forced deterministic variability (Koutsoyiannis et al. 2008). Use of a combination of forced and control (no external forcing) coupled model simulations, however, can potentially provide further insights (e.g. Vyushin and Kushner 2009).

One limitation of previous analyses using coupled climate model simulations is that the physics of the models is sufficiently complex that it is difficult to interpret the results in terms of basic physical properties. There is some utility in turning instead to a far more basic class of climate models known as "Energy Balance Models" (e.g. North et al. 1981). Here, the physics of the model is simple, control experiments are easy to devise, and results can readily be interpreted in terms of basic physical processes. Such properties are attractive from the perspective of "Occam's Razor". To the extent that observations can be explained in terms of the behavior exhibited by such a simple model, one need not appeal to more elaborate or exotic physics to provide a plausible explanation for observed features of the climate.

3 A simple climate modeling framework

We have employed a simple zero-dimensional EBM (see e.g. North et al. 1981) of the form

$$C dT/dt = F - A - BT + w(t) \tag{3}$$

T is the temperature of Earth's surface (approximated as the surface of a 70 m depth mixed layer ocean covering 70% of Earth's surface area). $C = 2.08 \times 10^8\ \mathrm{JK^{-1}\ m^{-2}}$ is an effective heat capacity that accounts for the thermal inertia of the mixed layer ocean, but does not allow for heat exchange with the deep ocean as in more elaborate so-called "upwelling-diffusion models" (e.g. Wigley and Raper 1990).

F is defined as the external forcing, which includes both shortwave (i.e. solar insolation) and longwave (anthropogenic greenhouse gas) forcing, viz. $F = S(1-\alpha)/4 + F_{GHG}$. The solar constant S (approximately 1,370 $\mathrm{Wm^{-2}}$) and effective surface albedo α (the reflectivity of Earth looking down from the top of Earth's atmosphere, approximately 0.3), are allowed to vary as prescribed by estimated past changes in solar irradiance, and the reflective natural+anthropogenic atmospheric aerosols, respectively. We use the approximation $F_{GHG} = 5.35 \log(\mathrm{CO_2}e/280)$ (Myhre et al. 1998) where 280 ppm is the preindustrial CO_2 level and CO_2e is the "equivalent" anthropogenic CO_2, representing the combined effects of all anthropogenic greenhouse gases.

$A + BT = \varepsilon\sigma T^4$ represents a linearization of the outgoing longwave radiation from Earth's surface, where σ is the Stefan-Boltzman constant, and ε is the surface longwave emissivity. By choosing appropriate "gray body" values for the parameters A and B, we account, in a crude manner, for the impact of natural atmospheric longwave emissitivity i.e., the natural atmospheric "Greenhouse Effect" (see McGuffie and Henderson-Sellers 1997). The choice $A = 221.3\ \mathrm{WK^{-1}\ m^{-2}}$, $B = 1.25\ \mathrm{Wm^{-2}}$ yields a realistic pre-industrial global mean temperature $T = 14.8$ C and an equilibrium climate sensitivity (equilibrium warming for a doubling of atmospheric CO_2 concentrations relative to pre-industrial values) of $\Delta T_{2\times CO2} = 3.0°C$, consistent with mid-range IPCC estimates (IPCC 2007).

$w(t)$ represents a stochastic heat flux forcing of the ocean surface associated with atmospheric "weather noise", typically approximated as Gaussian white noise. This stochastic weather forcing is approximated as white noise.

The response of the model to white noise forcing exhibits a characteristic response timescale $\tau = C/B$ ($\tau \approx 5.3$ years in our case) so that even in the absence of any external radiative forcing, temperature time series are expected to display "red noise" behavior, with increasing spectral power at increasingly lower frequencies (Hasselmann 1976, see e.g.). However, measurements spaced several τ apart are effectively independent, so that the noise asymptotically (i.e. if sampled coarsely in time) exhibits white noise properties. Such red noise is characterized by first order autoregressive, so-called "AR(1)" time series behavior that, unlike LRD red noise, has bounded variance at zero frequency.

Rea et al. (2011) take note that, in the real world, there is evidence that climate noise is more complex than simple AR(1) red noise, due to the presence of interannual and multidecadal internal oscillations (Mann and Lees 1996, see e.g.). There is indeed some important physics behind the sources of this oscillatory behavior—e.g. the El Nino phenomenon on interannual timescales, and coupled ocean-atmosphere phenomena that can give rise to even longer multidecadal oscillations (e.g. Delworth and Mann 2000). AR(1) climate noise is nonetheless a useful approximation and, as we will see, suffices for our present purposes.

The model is solved numerically, using estimated natural and anthropogenic forcing series as described below. Matlab© source code is available through Online Supplementary Information: http://www.meteo.psu.edu/~mann/supplements/Springboard_ClimChange10.

4 Comparing long rang dependence in simulations and observations

There are a number of different statistical approaches for estimating H from available time series. We tested 8 different approaches and settled on a combination of methods that exhibited stable performance and returned correct theoretical result for a wide-ranging set of test cases: (1) a pure Gaussian white noise $y_t = \varepsilon_t$ for which $H = 0.5$; (2) AR(1) red noise process $y_t = \rho\, y_{t-1} + \varepsilon_t$ with $0 < \rho < 1$ for which $0 < H < 1$ [and $H = 0.5$ asymptotically, i.e. as long as the series is sampled at a spacing large compared to the autocorrelation time scale $\tau = (1+\rho)/(1-\rho)$]; and (3) the case of a "random walk"/ "brownian motion", $y_t = y_{t-1} + \varepsilon_t$ for which the true theoretical value is $H = 1.5$. Note that this latter case can be viewed as the limit of AR(1) red noise as $\rho \to 1$. Further details are available in Supplementary Information.

We drove the EBM described in Section 3 with estimated natural (solar irradiance and volcanic aerosol) and anthropogenic (greenhouse gas and tropospheric aerosol) forcing over the period AD 850–1999 (Crowley 2000; Ammann et al. 2007—see Supplementary Information for further details). The amplitude of the white noise forcing term was chosen to give an average breakdown of 65% forced vs. 35% stochastic variance, consistent with previous estimates by Crowley (2000). For each experiment, an ensemble of 1,000 simulations was performed using independent realizations of white noise stochastic forcing. Any apparent scaling behavior observed in these simulations is the result of a combination of the linear response to prescribed external radiative forcing and stochastic white noise forcing.

In the absence of any noise (i.e. modeling *only* the pure radiatively forced component of temperature variation), we obtained the following values of H for the full Period (AD 850–1999): 0.870, pre-instrumental period (AD 1850–1849): 0.838, and instrumental period (AD 1850–1999): 0.903. Even during this latter period of marked warming from anthropogenic forcing, H does not breach the non-stationary $H = 1$ threshold. Adding stochastic forcing however has the effect of producing a distribution of possible values of H, which indeed extend into the non-stationary range.

The actual instrumental series (Brohan et al. 2006) is observed to fall reasonably well within the ensemble of realizations in the full (external + stochastic forcing) simulations (Fig. 1a), though there are some features (e.g. the warming spike due to the extremely large 1877–1878 El Nino event) that the model, for reasons discussed earlier, is likely too simple to capture. The estimated value of H for the actual instrumental series ($H = 0.871$) falls well within (see Fig. 2a) the distribution obtained for the full simulations (median $H = 0.889$, 95% confidence interval from 0.826 to 1.13; 8% of the values fall above the non-stationary $H = 1$ threshold). More surprisingly, however, it also falls within (Fig. 2a) the distribution for the stochastic forcing-only simulations (median 0.768, 95% confidence interval from 0.571 to 0.898; 1% still fall above the $H = 1$ threshold). Given the relatively short instrumental period, both stochastic forcing only or external + stochastic forcing are sufficient to explain the observed value of H.

It is reasonable to suspect that the two alternatives models might be more distinguishable given a longer time period, particularly since the stochastic component—AR(1) red noise—has an asymptotic value of $H = 0.5$. Let us thus consider the simulations over the entire period back to AD 850 (Fig. 1b and c). We find that, indeed, the distributions of H for the full and stochastic-only simulations essentially

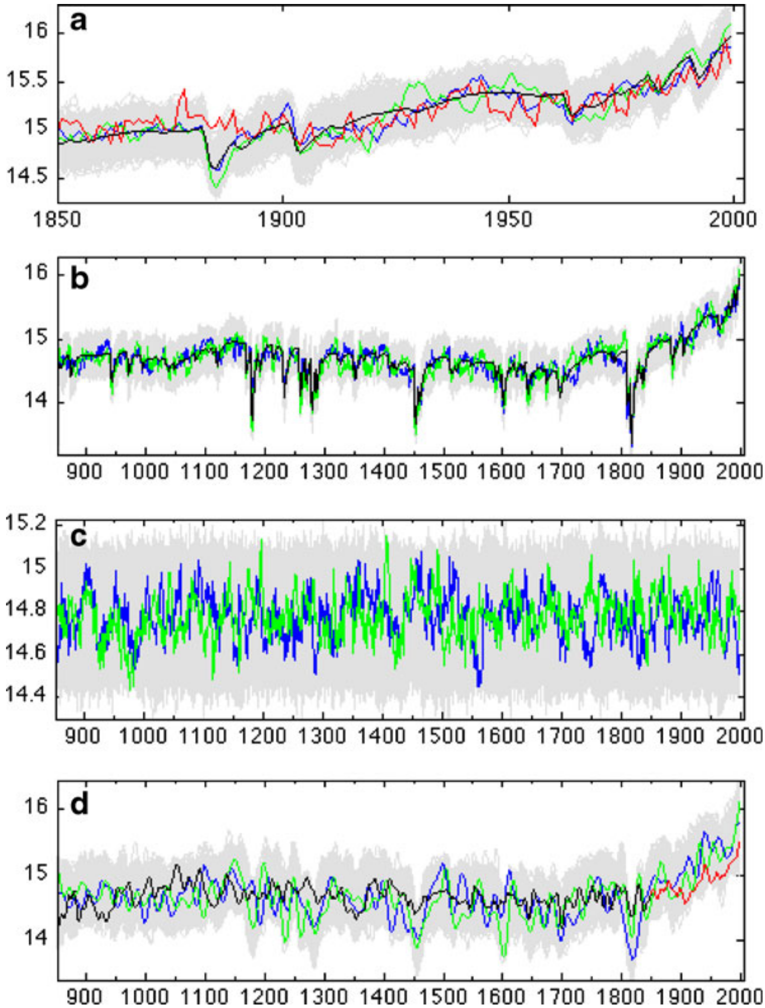

Fig. 1 Global mean surface temperature. **a** Period AD 1850–1999 from EBM simulation with radiative forcing only (*black*), radiative + stochastic forcing for two different noise realizations (*blue*; *green*) and full spread over 1,000 noise realizations (*gray*). The actual global mean surface temperature record is shown for comparison (*red*). **b** As in panel '**a**' but for the entire period AD 850–1999. **c** As in panel '**b**' but showing results for stochastic forcing only (i.e. fixed radiative forcing). **d** Model-generated "proxy" series for two realizations (*blue*; *green*) based on additive proxy noise as described in text. Shown for comparison (*black*) is proxy temperature reconstruction ("CPS" Northern Hemisphere land temperature reconstruction) of Mann et al. (2008) and instrumental temperature series (*red*). For the purposes of this comparison, all series have been decadally-smoothed (i.e. low-pass filtered with half-power at $f = 0.10$ cycle/year) using the method of Mann (2008)

separate when determined from the longer pre-instrumental time interval (Fig. 2b). The stochastic-only runs yield a median $H = 0.685$, with 95% confidence interval from 0.604 to 0.764. By contrast, the full simulations give a median $H = 0.816$, with 95% confidence interval from 0.782 to 0.848. Unlike the shorter instrumental interval

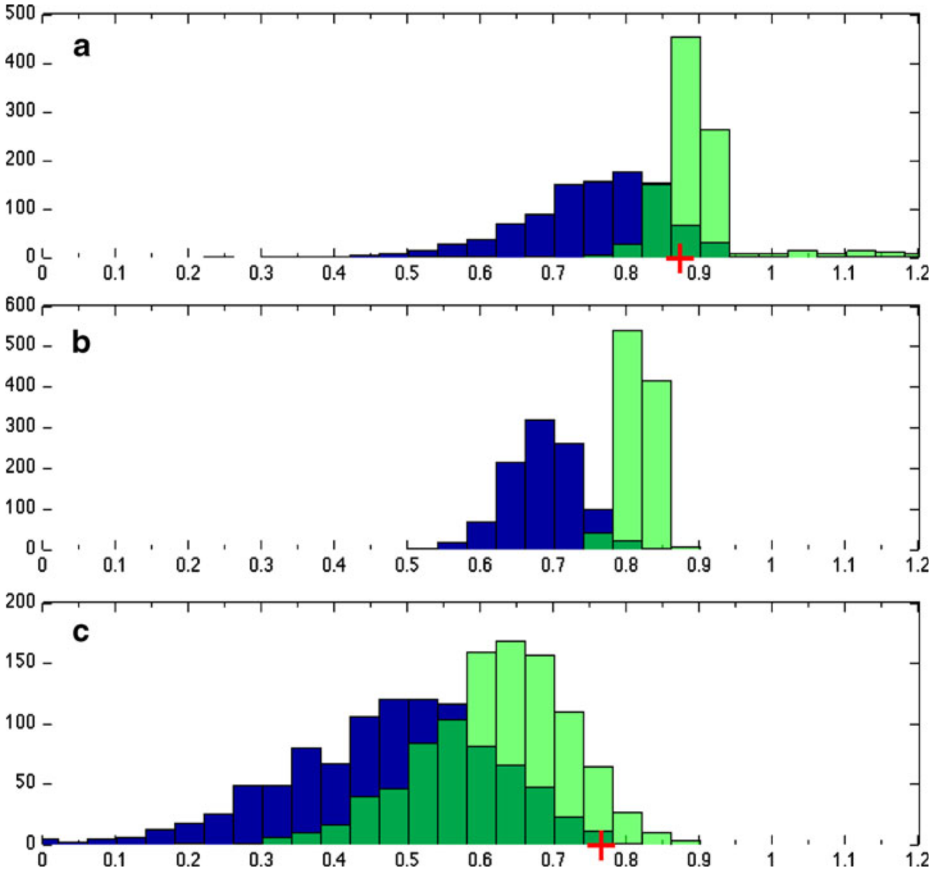

Fig. 2 Sampling distributions of H from model simulations for stochastic only (*blue*) and forced + stochastic (*green*) simulations. **a** Instrumental period AD 1850–1999. **b** Pre-instrumental period AD 850–1849. **c** Proxy temperature series AD 1850–1999 (series have been decimated for the purpose of calculating H owing to decadal resolution). Shown for comparison (*red cross symbols*) are observed values for instrumental record (in '**a**') and proxy temperature reconstruction of Mann et al. (2008) (in '**c**')

explored earlier, the $H = 1$ non-stationary limit is not breached for a single of the 1000 realizations for either of the two cases.

The fact that the 95% confidence intervals do not overlap for this longer time interval suggests that long-term "proxy" records of temperature (e.g. from tree-rings, corals, ice cores, and other natural indicators of climate) might be able to resolve which, if either, model (stochastic vs. forced + stochastic) best characterizes the power law scaling behavior for global mean temperature. Indeed it is such thinking that has lead Rea et al. (2011) (as well as others—e.g., Mills 2007) to investigate millennial length proxy records for evidence of possible LRD.

It turns out, however, that the situation isn't quite as simple as the comparison above might seem to suggest. Millennial reconstructions of climate from proxy data tend to have a substantial additional noise component that is due to physical, biological, or chemical properties of the proxy data unrelated to climate. Analyses of proxy data (e.g. Mann et al. 2007) suggest typically signal-to-noise amplitude ratios of about 0.4 (i.e. 86% noise by variance), while hemispheric composites, which tend to cancel out a substantial amount of this non-climatic noise, are closer to 40–60% noise by variance (see Mann et al. 2008). Importantly, the proxy noise appears somewhat red, approximated by AR(1) red noise with lag-one autocorrelation coefficient $\rho = 0.32$ (see Mann et al. 2007). For the purpose of comparison with actual decadal-resolution proxy reconstructions (Fig. 1d), we decadally-smooth and then decimate the records before estimating H. The stochastic-only runs give a median $H = 0.487$, with 95% confidence interval from 0.224 to 0.685. By contrast, the full simulations give a median $H = 0.629$ with 95% confidence interval from 0.441 to 0.773. The 95% confidence intervals for the two cases now do overlap significantly (Fig. 2c). However, the estimates for the actual proxy-reconstructed mean temperature series ($H = 0.764$) are inconsistent with stochastic-only simulations, and consistent, albeit just barely, with the full simulations.

5 Conclusions

Our analysis of climate model simulations and observational data provide no support for the hypothesis of non-stationary ($H > 1$) stochastic long-term behavior in global mean surface temperatures. Our findings are in this sense very much in accord with those of Rea et al. (2011), who conclude that:

In particular, the mean reverting nature of [statistical models such as random walks with $H > 1$] cannot be appealed to to give us comfort that recent observed temperature increases in instrumental records will naturally reverse themselves in the near future.

The EBM results analyzed provide a useful straw man for the interpretation of LRD (i.e. estimated values of the Hurst parameter H) in observational temperature time series. The values of H estimated in historical observed global mean temperature are entirely consistent with established physics as contained within a simple climate model. The short instrumental time interval, however, is found to be insufficient to distinguish the competing explanations of purely stochastic vs. externally forced +stochastic variability within the modeling framework.

Use of longer, millennial-length simulations more sharply distinguishes the underlying distributions expected for the two cases. The presence of non-climatic noise in proxy temperature data available for comparison over this longer timescale, however, complicates the interpretation, as the two distributions are found to overlap substantially when allowing for the additional influence of non-climatic proxy noise. Nonetheless, one of the two distributions (unforced, purely stochastic variability) is found to be inconsistent with the proxy data, while the other (externally forced + stochastic variability) is found to be compatible with the observations.

We conclude that the behavior of H in instrumental and long-term proxy temperature reconstructions appears consistent with the results of a simple climate model (EBM) forced by estimated natural and anthropogenic radiative forcing changes, and subject to white noise stochastic weather forcing. Nothing more exotic than the physics of such a simple model is necessary to explain the apparent scaling behavior in observed surface temperatures.

We argue that such a simple modeling framework should always serve as a starting point and, indeed, a null hypothesis, in any search for putative exotic behavior in observational climate data. If similar behavior can be found within such as simple theoretical framework, Occam's Razor dictates that it would be imprudent to look to more exotic, non-physically-based time series models for an explanation of observed variability and trends. More generally, any statistical analysis of climate data must be informed by an understanding of the physics underlying the observational phenomena at hand. A statistical analysis devoid of physical understanding is prone to false, and in some cases, extremely misleading inferences—such as the conclusion that global warming can be dismissed as the random excursions of a process displaying LRD.

Acknowledgements This article benefited from brief discussion with Rypdal Kristoffer, Gavin Schmidt, and Richard Smith, Dmitry Vyushin, Francis Zwiers,

References

Ammann CM, Joos F, Schimel DS, Otto-Bliesner BL, Tormas RA (2007) Solar influence on climate during the past millennium: results from transient simulations with the NCAR Climate System Model. Proc Natl Acad Sci U S A 104:3713–3718

Blender R, Fraedrich K (2003) Long time memory in global warming simulations. Geophys Res Lett 30(14):1769

Bloomfield P, Nychka D (1992) Climate spectra and detecting climate change. Clim Change 21:275–287

Brohan P, Kennedy JJ, Haris I, Tett SFB, Jones PD (2006) Uncertainty estimates in regional and global observed temperature changes: a new dataset from 1850. J Geophys Res 111:D12106. doi:10.1029/2005JD006548

Crowley TJ (2000) Causes of climate change over the past 1000 years. Science 289:270–277

Delworth TL, Mann ME (2000) Observed and simulated multidecadal variability in the northern hemisphere. Clim Dyn 16:661–676

Eichner J, Koscielny-Bunde E, Bunde A, Havlin S, Schellnhuber H (2003) Power-law persistence and trends in the atmosphere: a detailed study of long temperature records. Clim Dyn 68E:046133. doi:10.1103/PhysRevE.68.046133

Fraedrich K, Blender R (2003) Scaling of atmosphere and ocean temperature correlations in observations and climate models. Phys Rev Lett 90:108501. doi:10.1103/PhysRevLett.90.108501

Gil-Alana L (2005) Statistical modeling of the temperatures in the Northern Hemisphere using fractional integration techniques. J Climate 18:5357–5369

Hasselmann K (1976) Stochastic climate models. Part 1: theory. Tellus 28:473–485

IPCC (2007) Climate Change 2007: the physical science basis. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds) Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge

Király A, Bartos I, Janosi IM (2006) Correlation properties of daily temperature anomalies over land. Tellus A 58:593–600. doi:10.1111/j.1600-0870.2006.00195.x

Koutsoyiannis D, Efstratiadis A, Mamassis N, Christofides A (2008) On the credibility of climate predictions. Hydrol Sci J 53(4):671–684

Mann ME (2008) Smoothing of climate time series revisited. Geophys Res Lett 35:L16708. doi:10.1029/2008GL034716

Mann ME, Lees J (1996) Robust estimation of background noise and signal detection in climatic time series. Clim Change 33:409–445

Mann ME, Rutherford S, Wahl E, Ammann C (2007) Robustness of proxy-based climate field reconstruction methods. J Geophys Res 112:D12109. doi:10.1029/2006JD008272

Mann ME, Zhang Z, Hughes MK, Bradley RS, Miller SK, Rutherford S, Ni F (2008) Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia. Proc Natl Acad Sci 105:13252–13257

McGuffie K, Henderson-Sellers A (1997) Climate modeling primer, 2nd Edn. Wiley, Chichester

Mills T (2007) Time series modelling of two millenia of northern hemisphere temperatures: long memory or shifting trends? J R Stat Soc Ser A 170:83–94

Myhre G, Highwood EJ, Shine KP, Stordal F (1998) New estimates of radiative forcing due to well mixed greenhouse gases. Geophys Res Lett 25:2715–2718

North GR, Cahalan RF, Coakley JA (1981) Energy balance climate models. Rev Geophys 19:91–121

Rea W, Reale M, Brown J (2011) Long memory in temperature reconstruction. Clim Change (this issue)

Smith R (1993) Long-range dependence and global warming. In: Barnett V, Turkman F (eds) Statistics for the environment. Wiley, Chichester, pp 141–161

Vyushin DI, Kushner PJ (2009) Power-law and long-memory characteristics of the atmospheric general circulation. J Climate 22:2890–2904

Vyushin DJ, Kushner PJ, Mayer J (2009) On the origins of temporal power-law behavior in the global atmospheric circulation. Geophys Res Lett 36:L14706. doi:10.1029/2009GL038771

Wigley TML, Raper SCB (1990) Natural variability of the climate system and detection of the greenhouse effect. Nature 344:324–327. doi:10.1038/344324a0