Comment on "Heat capacity, time constant, and sensitivity of Earth's climate system" by S. E. Schwartz

Grant Foster,[1] James D. Annan,[2] Gavin A. Schmidt,[3] and Michael E. Mann[4]

Received 10 September 2007; revised 30 April 2008; accepted 6 May 2008; published 2 August 2008.

Citation: Foster, G., J. D. Annan, G. A. Schmidt, and M. E. Mann (2008), Comment on ''Heat capacity, time constant, and sensitivity of Earth's climate system'' by S. E. Schwartz, *J. Geophys. Res.*, *113*, D15102, doi:10.1029/2007JD009373.

1. Introduction

[1] In a recent paper, *Schwartz* [2007] (hereinafter referred to as S07) estimated climate sensitivity using globally averaged observed time series. He estimated the planet's effective heat capacity and "pertinent time constant" from ocean heat content and observed global temperature records, assuming that the latter can be modeled as a first-order Markov (AR(1)) process after removing the trend due to climate forcing by subtracting a linear trend.

[2] As S07 correctly states, the autocorrelation of an AR(1) process at lag Δt follows the simple relation

$$\rho(\Delta t) = e^{-\Delta t/\tau}, \tag{1}$$

where τ is the e-folding time constant. This makes it possible in principle to estimate τ from the sample autocorrelation estimate $\hat{r}$ of a finite time series:

$$\hat{\tau} = -\Delta t / \ln(\hat{r}(\Delta t)). \tag{2}$$

[3] This is the crux of S07's method for estimating the time constant of the global climate system from the 125-year record of global annually averaged temperature anomaly. The result for GISTEMP annual data is shown in S07's Figure 5g, for monthly data in S07's Figure 7; he concludes that the relevant time constant for global climate is 5 ± 1 years. It is important to recognize that S07's analysis does not depend on detailed assumptions concerning the Earth's climate system, only requiring that it obeys simple zero-order energy balance principles. Therefore, if his analysis method is valid, it should apply equally to all models which obey these same principles. In this comment, we test the performance of his analysis technique on a range of complex (section 3) and simple (section 4) models. First, in section 2, we explore the plausibility of the AR(1) approximation.

2. Autocorrelation of an AR(1) Process

[4] Even if the stochastic part of the temperature time series can be approximated by an autoregressive process, it may still exhibit numerous time scales. The climate system is likely to show such behavior, as it has multiple components with physically distinct characteristics: the atmosphere, land, upper ocean, deep ocean, and cryosphere. In fact the principal physical mechanism which leads us to believe that not all committed greenhouse gas warming has yet been experienced, and a substantial amount remains "in the pipeline," is the warming of the deep ocean [*Hansen et al.*, 2005; *Wigley*, 2005].

[5] Such a multicomponent physical system cannot be expected to act with a single time scale. Even if the system evolves according to an AR(1) process, it must be a vector AR(1) process with many distinct time scales. The observed surface temperature will be some function of the vector describing the entire system, and its time series is likely to exhibit all these time scales given a long enough set of data. Hence, basic physical considerations argue strongly against the notion that the global average surface temperature has a single characteristic time scale, or time constant.

[6] Furthermore, much of the low-frequency variability in temperature time series is very likely not stochastic, but represents a response to radiative forcing changes that are external to the climate system (both natural and anthropogenic). Under the simplifying assumption that the forced and internal variability "noise" components are linearly additive, the latter component alone might be considered, as a null hypothesis, to conform approximately to an AR(1) process [*Gilman et al.*, 1963; *Hasselmann*, 1976] in the asymptotic limit where complicating processes such as advective and diffusive exchanges of heat with the deep ocean are ignored [e.g., *Wigley and Raper*, 1990]. However the increase in forcing that has taken place over the 20th century is both strong and not linear. Therefore it seems unreasonable to expect that an AR(1) process should adequately describe observed temperature variations, even after detrending [*Mann and Lees*, 1996].

[7] For an AR(1) process, the ratio $\Delta t/\ln(\hat{r}(\Delta t))$ should be roughly constant. The results in S07, both for annual and

[1]Island Data Corporation, Carlsbad, California, USA.
[2]Frontier Research Center for Global Change, JAMSTEC, Yokohama, Japan.
[3]NASA Goddard Institute for Space Studies, New York, New York, USA.
[4]Department of Meteorology and Earth System Science Center, Pennsylvania State University, University Park, Pennsylvania, USA.

Copyright 2008 by the American Geophysical Union.
0148-0227/08/2007JD009373

Table 1. Estimated Climate Sensitivity From Five Realizations of Model Output for the GISS ER AOGCM for a Simulation of 20th Century Annual Global Mean Temperature Change Using the Analysis Method of *Schwartz* [2007][a]

Run	C1 (W a m^{-2} K^{-1})	C2 (W a m^{-2} K^{-1})	$\hat{\tau}$ (years)	S1 (K m^2 W^{-1} [C])	S2 (K m^2 W^{-1} [C])
1	25.7	30.2	2.5	0.096 [0.4]	0.082 [0.3]
2	21.4	25.3	2.9	0.135 [0.5]	0.115 [0.5]
3	24.1	24.3	2.3	0.094 [0.4]	0.093 [0.4]
4	24.9	27.6	6.9	0.275 [1.1]	0.248 [1.0]
5	23.4	26.7	7.6	0.324 [1.3]	0.284 [1.1]
Average	23.9	26.8	4.41	0.185 [0.7]	0.164 [0.7]

[a]S1 = tau/C1, S2 = tau/C2, and values in square brackets are equivalent warmings for a doubling of CO_2.

monthly data sets, show a distinctly increasing ratio as a function of lag; the estimated time scale is, for low lag, roughly proportional to the lag. This casts severe doubt on the applicability of equation (1), which is the basis for the estimate of the time constant. The monthly data result (Figure 7 in S07) is in clear contradiction to the AR(1) model.

[8] S07 tests for the appropriateness of the AR(1) model by regressing the detrended GISTEMP annual surface air temperature data T_s against its lag-1 value, then examining the residuals. The conclusion is (S07, paragraph 24) "Satisfaction of the assumption of a first-order Markov process was assessed by examination of the residuals of the lag-1 regression, which were found to exhibit no further significant autocorrelation." However, it appears this test was applied only to the annual time series, which consists of only 125 data points. For such a short time series the detection threshold is about $\simeq\pm0.2$, which makes the null hypothesis rather hard to refute conclusively, but there is still strong evidence that a higher-order model is more appropriate [e.g., *Smith et al.*, 2003].

[9] Nevertheless, treating the system as an energy balance system with a single time constant may still be a reasonable approximation under some circumstances, so we do not reject the approach in S07 out of hand. Instead, we apply the method to some numerical experiments where the correct answer is known, in order to investigate its accuracy. First we use a climate model which simulates the multiple time scales of fast atmospheric and slower oceanic processes, and then we test the method on the simple energy balance model from which it was derived by S07. Finally, we provide a simple theoretical analysis to explain a significant source of bias and uncertainty which is intrinsic to the method.

3. Application to Climate Model Results

[10] We have applied the method of S07 to the output of 5 simulations of the GISS-ER GCM [*Hansen et al.*, 2007] run with observed forcing for the same time period, 1880 to 2003. All the estimated parameters were derived as described by S07, using the total heat content anomaly (derived from the radiative imbalance at the top of the atmosphere) and the global mean surface temperature anomaly; the results are presented in Table 1.

[11] The time constant was estimated by computing $-\Delta t/\ln(\hat{r}(\Delta t))$ up to the last nonnegative $\hat{r}(\Delta t)$ value, for the linearly detrended T_s time series. In order to compute estimates that are the most favorable possible for the S07 method, we took the highest estimated $\hat{\tau}$ of all these values. We tested both methods given in S07 for estimating the effective heat capacity [*Wigley and Raper*, 1990], first (columns C1 and S1) by regressing the heat content anomaly (H) and temperature anomaly (T_s) from 1956 to 2002 (bisector of the slope of regressing H against T_s and the inverse of the slope of the regression of T_s against H), and second (columns C2 and S2) by taking the ratio of the estimates of dH/dt and dT_s/dt.

[12] The average of all these sensitivity estimates for the model runs is merely 0.175 K/(W/m^2), and none of the estimates exceeds 0.324 K/(W/m^2). These values imply warmings of 0.7 C and 1.3 C, respectively, for a CO_2-doubling experiment (for the GISS-ER model, a doubling of CO_2 results in a forcing of around 4 W/m^2). Yet this model is known to exhibit a true equilibrium climate sensitivity of 2.7 C under doubled CO_2 conditions.

[13] Although the effective heat capacities of these simulations seem rather high compared to the S07 observationally based estimate, the estimated time constants appear to be the greater problem with this analysis. A comparison of the autocorrelation estimates for the five runs of GISS-ER, and the GISTEMP time series data, is shown in Figure 1. It is apparent from Figure 1 (bottom) that taking the highest single value for the estimated time scale is, for some of the model runs, unrepresentative of the average behavior. Yet even with this built-in high bias, while the GISTEMP data yield the same very short time scale estimated by S07 (slightly different here since we used data through 2006), the model runs show even shorter and highly variable estimates of time scale. However, it is known that the model takes a number of decades to equilibrate after a change in external forcing, and that its sensitivity is many times greater than these estimates obtained via the S07 method. Hence this time scale analysis method does not appear to correctly diagnose the properties of the model.

4. Application to an Energy Balance Model

[14] S07 invokes the ansatz of a zero-dimensional energy balance model to justify the analysis. We show in Figure 2 (top) the results from a large ensemble of integrations of the first-order energy balance model

$$dH/dt = F - \lambda T + \epsilon \tag{3}$$

where $H = CT$ is the heat anomaly, C is the effective heat capacity, T is the surface temperature anomaly, F is the external forcing anomaly, λ is the radiative feedback

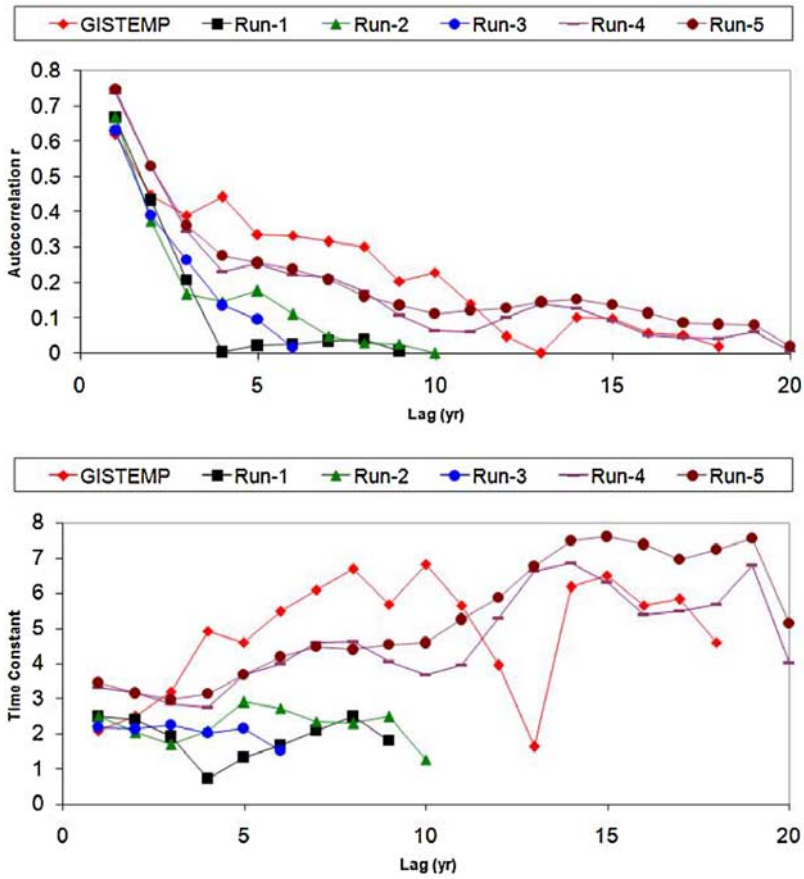

Figure 1. Autocorrelation, and estimated time constant by S07's method, for GISTEMP temperature time series and the output of five runs of the GISS-ER GCM.

(inverse of the sensitivity) and ϵ is white noise weather forcing.

[15] The effective heat capacity was set at 19 W a m^{-2} K^{-1}, and three values of τ were used, 5, 15 and 30 years (giving sensitivities of 1, 3 and 6 C for 3.7 W m^{-2} of forcing, equivalent to a doubling of CO_2). External forcing (comprising volcanic aerosols, GHGs, solar variability, and anthropogenic aerosols) was taken from [*Crowley*, 2000] with noise sampled from the Gaussian distribution $N(0, 1.5)$ (which allows for a "very unlikely" probability of 1–5% that natural variability could exceed the amplitude of forced variations). After an initial spin-up, the final 125 years (1874–1998) were analyzed. The method used by S07, of calculating $-\Delta t/\ln(\hat{r}(\Delta t))$ from the detrended time series, can be seen to generate highly inaccurate and strongly biased estimates. Each histogram is based on 1000 estimated values, each of which is the average of the diagnosed time scales for 5–15 year lags in an integration where all these values were defined. The results are insensitive both to these choices, and to the amount of noise added. So even in this case, where the underlying system does have a unique and well-defined relaxation time scale, this method of analysis fails to diagnose it, and, contrary to the confident estimate in S07 of $\tau = 5 \pm 1$, a time scale as high as 30 years (corresponding to a sensitivity of 6C for a doubling of CO_2) or even greater is compatible with the S07 analysis of global surface temperature. In this simple energy balance model, the effective heat capacity is precisely determined by the ratio of heat anomaly to surface temperature, so these erroneous estimates of τ directly feed though into equivalently erroneous estimates for the sensitivity of the model.

[16] Results from a similar analysis of 14 state-of-the-art GCMs from the World Climate Research Programme's (WCRP's) Coupled Model Intercomparison Project phase 3 (CMIP3) multimodel data set are also shown in Figure 2. Lags of 5–10 years were used in this analysis, in order not to exclude too many models from the sample. As with the GISS model analysis presented earlier, the estimates of time scale produced by this method are generally unrealistically low in comparison to the known behavior of the models in response to changes in GHG forcing. A more detailed analysis, considering also their effective heat capacities, confirms that the sensitivity estimates provided by the method of S07 are inaccurate and strongly biased [*Knutti et al.*, 2008].

5. Application to a Pure AR(1) Process

[17] As the length N of a sampled time series increases, the sample autocorrelation $\hat{r}_j$ converges to the true autocorrelation ρ_j, and follows an unbiased Gaussian distribution with variance given by Bartlett's formula [*Bartlett*, 1946]. What is widely overlooked is that the usual (Yule-Walker) estimate of autocorrelation is intrinsically biased [*Priestly*,

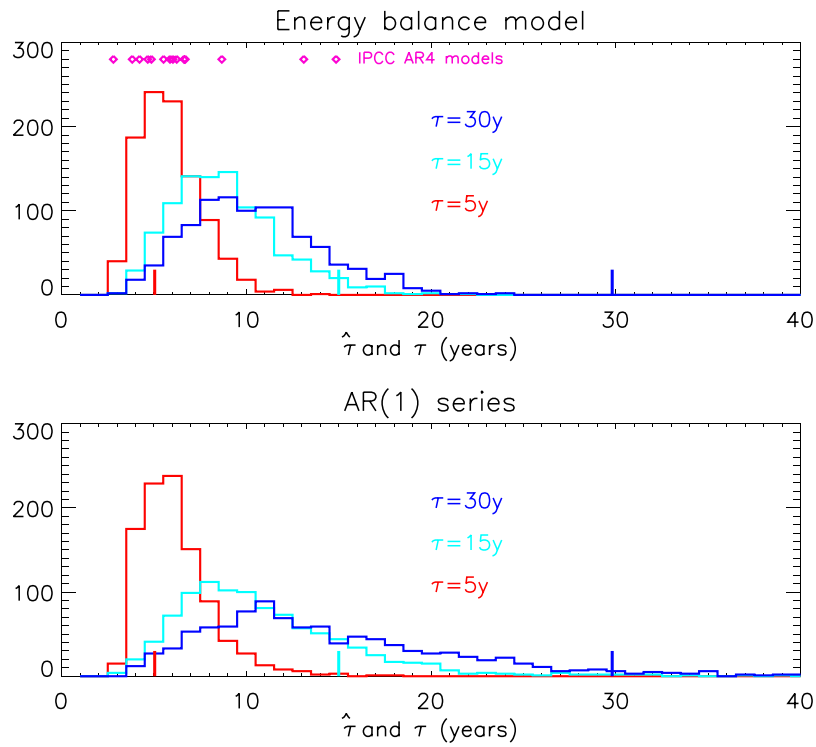

Figure 2. Estimation of time scale: (top) EBM and GCM results. (bottom) AR(1) process. Red, cyan, and blue indicate results for true τ values of 5, 15, and 30 years, respectively (true value marked by vertical lines).

1982; *Percival and Walden*, 1993], more so when the time series is short and the time scale is long.

[18] Figure 2 (bottom) shows the results obtained when the deterministic external forcing is omitted from the energy balance model, resulting in a pure AR(1) process with known time scale. Even with a true time scale as long as 30 years, there is a significant probability of obtaining an estimate as low as 5. No detrending was performed in this experiment; doing so would have reduced the estimated time scale still further.

[19] This possibly surprising result is readily explicable when considering the statistical behavior of the Yule-Walker estimate of correlation coefficients for short time series in the presence of strong autocorrelation. The lowest-order bias in the lag-1 autocorrelation estimate for an AR(1) process is [*Tjostheim and Paulsen*, 1983]

$$\langle \hat{r}_1 \rangle - \rho_1 = -N^{-1}(1 + 4\rho_1). \tag{4}$$

[20] For a series with $N = 125$ data points and time scale $\tau = 30$, the true value of ρ_1 is 0.9672, but the expected value of $\hat{r}_1$ is only 0.9283. This value would lead to an estimated time scale $-1/\ln(\hat{r}_1)$ of only 13.4. In fact, as the time scale increases, the expected value of the estimated lag-1 autocorrelation approaches a limiting value of 0.96, which implies an estimated time scale of only 24.5.

[21] For higher lags, bias in the autocorrelation estimate reduces the estimated time scale even more, and linear detrending reduces it still further. For $\tau = 30$ with $N = 125$, the biases conspire to lower the mean estimate of the autocorrelation time scale to less than half its actual value, with a significant probability of even lower results, as observed in our numerical tests.

6. Conclusions

[22] S07 has proposed an analysis method based on approximating the climate system as a linear trend plus an autoregressive process of order 1, forced by random noise. There are strong physical arguments why this approach is likely to be an oversimplification, and as we have shown, the data contradict this hypothesis. The S07 analysis method generates strongly biased results when applied to a climate model of known sensitivity, and even when applied to the simple energy balance model S07 invokes to justify the approach. In fact the S07 method for estimating the time scale of an AR(1) process is strongly biased for realistic parameter values, and we have provided a simple demonstration and explanation of this effect. We suggest that such credibility checks, which are not difficult to perform, should be considered a first step when a novel analysis technique such as that presented by S07 generates results that are inconsistent with previous work. Previous research indicates that, when correctly analyzed, the 20th century trends do not strongly constrain climate sensitivity other than to rule out low values [e.g., *Gregory et al.*, 2002; *Forest et al.*, 2006]. The belief that climate sensitivity is likely to lie in the range 2–4.5 C and very unlikely less than 1.5 C [*Intergovernmental Panel on Climate Change*, 2007] is supported by numerous diverse analyses of the climate system [e.g., *Annan and Hargreaves*, 2006; *Hegerl et al.*,

2006] and we conclude that S07 presents no substantive basis to challenge this view.

[23] **Acknowledgments.** We acknowledge the modeling groups, the Program for Climate Model Diagnosis and Intercomparison (PCMDI), and the WCRP's Working Group on Coupled Modelling (WGCM) for their roles in making available the WCRP CMIP3 multimodel data set. Support of this data set is provided by the Office of Science, U.S. Department of Energy. We are grateful for helpful comments from the reviewers and J. C. Hargreaves. J. D. A. was supported by the Global Environment Research Fund (S-5) of the Ministry of the Environment, Japan.

References

Annan, J. D., and J. C. Hargreaves (2006), Using multiple observationally-based constraints to estimate climate sensitivity, *Geophys. Res. Lett.*, *33*, L06704, doi:10.1029/2005GL025259.

Bartlett, M. S. (1946), On the theoretical specification and sampling properties of autocorrelated time series, *J. R. Stat. Soc.*, *8*, suppl., 27–41.

Crowley, T. J. (2000), Causes of climate change over the past 1000 years, *Science*, *289*, 270–277.

Forest, C. E., P. H. Stone, and A. P. Sokolov (2006), Estimated PDFs of climate system properties including natural and anthropogenic forcings, *Geophys. Res. Lett.*, *33*, L01705, doi:10.1029/2005GL023977.

Gilman, D. J., F. J. Fuglister, and J. M. Mitchell Jr. (1963), On the power spectrum of red noise, *J. Atmos. Sci.*, *20*, 182–184.

Gregory, J. M., R. J. Stouffer, S. C. B. Raper, P. A. Stott, and N. A. Rayner (2002), An observationally based estimate of the climate sensitivity, *J. Clim.*, *15*(22), 3117–3121.

Hansen, J., et al. (2005), Earth's energy imbalance: Confirmation and implications, *Science*, *308*, 1431–1435, doi:1410.1126/science.1110252.

Hansen, J., et al. (2007), Climate simulations for 1880–2003 with GISS modelE, *Clim. Dyn.*, *29*, 661–696, doi:10.1007/s00382-007-0255-8.

Hasselmann, K. (1976), Stochastic climate models, part 1. Theory, *Tellus*, *28*, 473–485.

Hegerl, G. C., T. J. Crowley, W. T. Hyde, and D. J. Frame (2006), Climate sensitivity constrained by temperature reconstructions over the past seven centuries, *Nature*, *440*, 1029–1032.

Intergovernmental Panel on Climate Change (2007), *Climate Change 2007: The Physical Science Basis*, edited by S. Solomon et al., Cambridge Univ. Press, New York.

Knutti, R., S. Krähenmann, D. J. Frame, and M. R. Allen (2008), Comment on "Heat capacity, time constant, and sensitivity of Earth's climate system" by S. E. Schwartz, *J. Geophys. Res.*, *113*, D15103, doi:10.1029/2007JD009473.

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

Percival, D. B., and A. T. Walden (1993), *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques*, pp. 190–195, Cambridge Univ. Press, New York.

Priestly, M. B. (1982), *Spectral Analysis and Time Series*, Academic, London.

Schwartz, S. E. (2007), Heat capacity, time constant, and sensitivity of Earth's climate system, *J. Geophys. Res.*, *112*, D24S05, doi:10.1029/2007JD008746.

Smith, R. L., T. M. L. Wigley, and B. D. Santer (2003), A bivariate time series approach to anthropogenic trend detection in hemispheric mean temperatures, *J. Clim.*, *16*(8), 1228–1240.

Tjostheim, D., and J. Paulsen (1983), Bias of some commonly-used time series estimates, *Biometrika*, *70*, 389–399.

Wigley, T. M. L. (2005), The climate change commitment, *Science*, *307*(5716), 1766–1769.

Wigley, T. M. L., and S. C. B. Raper (1990), Natural variability of the climate system and detection of the greenhouse effect, *Nature*, *344*, 324–327.

J. D. Annan, Frontier Research Center for Global Change, JAMSTEC, 3173-25 Showamachi, Kanazawa-ku, Yokohama, Kanagawa 236-0001, Japan. (jdannan@jamstec.go.jp)

G. Foster, Island Data Corporation, 2386 Faraday Avenue, Suite 280, Carlsbad, CA 92008, USA. (grant.foster@islanddata.com)

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

G. A. Schmidt, NASA Goddard Institute for Space Studies, 2880 Broadway, New York, NY 10025, USA. (gschmidt@giss.nasa.gov)