

# Comparing variability and trends in observed and modelled global-mean surface temperature

John C. Fyfe,<sup>1</sup> Nathan P. Gillett,<sup>1</sup> and David W. J. Thompson<sup>2</sup>

Received 8 June 2010; revised 1 July 2010; accepted 6 July 2010; published 19 August 2010.

[1] The observed evolution of the global-mean surface temperature over the twentieth century reflects the combined influences of natural variations and anthropogenic forcing, and it is a primary goal of climate models to represent both. In this study we isolate, compare, and remove the following natural signals in observations and in climate models: dynamically induced atmospheric variability, the El Niño–Southern Oscillation, and explosive volcanic eruptions. We make clear the significant model-to-model variability in estimates of the variance in global-mean temperature associated with these natural signals, especially associated with the El Niño–Southern Oscillation and explosive volcanic eruptions. When these natural signals are removed from time series of global-mean temperature, the statistical uncertainty in linear trends from 1950 to 2000 drops on average by about half. Hence, the results make much clearer than before where some model estimates of global warming significantly deviate from observations and where others do not. **Citation:** Fyfe, J. C., N. P. Gillett, and D. W. J. Thompson (2010), Comparing variability and trends in observed and modelled global-mean surface temperature, *Geophys. Res. Lett.*, 37, L16802, doi:10.1029/2010GL044255.

## 1. Introduction

[2] The twentieth century evolution of global-mean surface temperature reflects the combined influences of naturally occurring climate fluctuations (e.g., oceanic, volcanic, dynamical and possibly solar variability), anthropogenic emissions of various gases and aerosols (e.g., greenhouse gases, sulphate aerosols, and black carbon aerosols), and land use changes (e.g., deforestation). In a recent study some naturally occurring fluctuations were identified and removed by subtracting from the observed record of global-mean surface temperature three known signals of natural climate variability, namely, dynamically induced atmospheric variability associated with anomalous temperature advection over the Northern Hemisphere land masses, the El Niño–Southern Oscillation (ENSO), and explosive volcanic eruptions [Thompson *et al.*, 2009]. The residual time series so obtained reveals near monotonic warming since about 1950, presumably of anthropogenic origin, as well as interannual variations tied to land-sea coupling. Here, we take the logical and important next step of extending this calculation to the

multi-model ensemble of twentieth century simulations performed in support of the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report [Intergovernmental Panel on Climate Change, 2007]. The objectives of this study are two-fold: 1) To isolate and compare the observed and simulated signals of natural climate variability. 2) To compare the observed and simulated post-1950 trends in the presence and absence of these naturally occurring climate fluctuations.

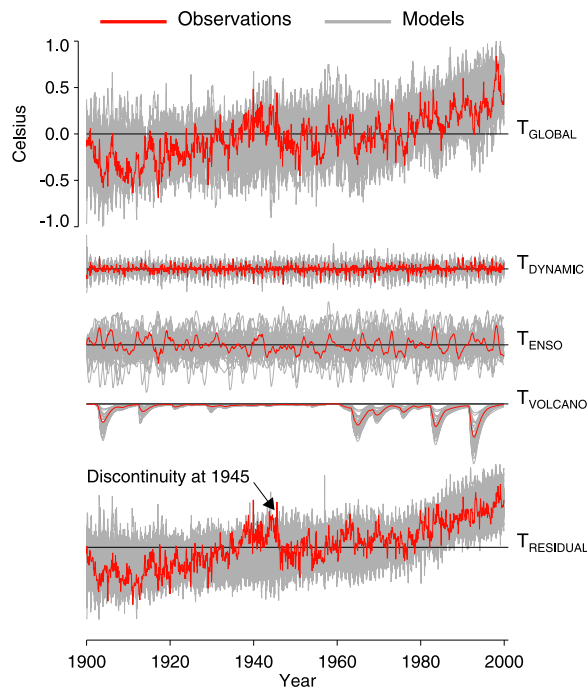
## 2. Data and Models

[3] The observed temperature data used in this study are the HadCRUT3 combined land surface temperature and sea surface temperature (SST) datasets [Brohan *et al.*, 2006]. The temperature dataset is available via the Climatic Research Unit at the University of East Anglia and is provided in monthly-mean form on a 5° (latitude) by 5° (longitude) grid, and is expressed as anomalies with respect to the 1961 to 1990 base period. The observed sea-level pressure (SLP) data are provided by the National Center for Atmospheric Research Data Support Section and are also given as monthly-means on a 5° (latitude) by 5° (longitude) grid as described by Trenberth and Paolino [1980]. The seasonal cycle is removed from the SLP data by subtracting the long-term mean calculated for the period 1950 to 2000 as a function of calendar month.

[4] The Atmosphere Ocean Coupled Climate Model (AOGCM) data used in this study are from the Coupled Model Intercomparison Project 3 (CMIP3) [Meehl *et al.*, 2007] database of climate model simulations archived at the Program for Climate Model Diagnostics and Intercomparison (PCMDI) at the Lawrence Livermore National Laboratory. We consider a multi-model ensemble of twentieth century simulations driven with observed greenhouse gas and sulphate aerosol forcing, and in some cases, volcanic forcing. The 24 models considered are BCC-CM1 (A,4), BCCR-BCM2-0 (B,1), CCCMA-CGCM3-1 (C,5), CCCMA-CGCM3-1-T63 (D,1), CNRM-CM3 (E,1), CSIRO-MK3-0 (F,3), CSIRO-MK3-5 (G,1), GFDL-CM2-0 (H,3), GFDL-CM2-1 (I,3), GISS-AOM (J,2), GISS-MODEL-E-H (K,5), GISS-MODEL-E-R (L,2), IAP-FGOALS1-0-G (M,3), INGV-ECHAM4 (N,1), INMCM3-0 (O,1), IPSL-CM4 (P,1), MIROC3-2-MEDRES (Q,3), MIROC3-2-HIRES (R,3), MIUB-ECHO-G (S,5), MPI-ECHAM5 (T,1), MRI-CGCM2-3-2A (U,5), NCAR-CCSM3-0 (V,1), NCAR-PCM1 (W,4) and UKMO-HADCM3 (X,2); with model identifier and number of independent realizations analyzed shown in parentheses. The simulated surface temperature and SLP fields are interpolated onto the same 5° (latitude) by 5° (longitude) grid as the observations and are thereafter treated identically to the observations including masking with observed coverage.

<sup>1</sup>Canadian Centre for Climate Modelling and Analysis, Environment Canada, University of Victoria, Victoria, British Columbia, Canada.

<sup>2</sup>Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado, USA.



**Figure 1.** (top) The monthly-mean global-mean surface temperature anomalies; (middle) the natural signals; (bottom) the residual time series found by removing the natural signals from  $T_{GLOBAL}$ .

All significance testing in this study is at the 95% confidence level.

### 3. Methodologies

[5] The natural climate signals in monthly-mean global-mean surface temperature  $T_{GLOBAL}(t)$  are estimated as follows using the method of *Thompson et al.* [2009].

#### 3.1. Dynamically Induced Signal

[6] The dynamically induced signal is found on a month-by-month basis by regressing SLP anomaly maps onto normalized land-ocean temperature difference time series for the Northern Hemisphere. The coefficient time series of these SLP regression patterns define the dynamically induced signal  $\hat{T}_{DYNAMIC}(t)$ .

#### 3.2. ENSO Signal

[7] The ENSO signal is estimated using a simple oceanic mixed layer model  $C(d/dt)\hat{T}_{ENSO}(t) = F(t) - \hat{T}_{ENSO}(t)/\beta$ , where  $F(t)$  is an estimate of the anomalous flux of sensible and latent heat in the eastern tropical Pacific;  $\beta$  is a linear damping coefficient set to  $2/3 \text{ K W}^{-1} \text{ m}^{-2}$ ; and  $C$  is an effective heat capacity of the atmosphere-ocean system per unit area. The heat capacity is obtained such that the correlation coefficient between  $\hat{T}_{ENSO}(t)$  and  $T_{GLOBAL}(t)$  is maximum based on de-trended data from 1950 through 2000. The results in this study are largely insensitive to the choice of  $\beta$ , or computed values of  $C$ .

#### 3.3. Volcanic Signal

[8] The volcanic signal  $\hat{T}_{VOLCANO}(t)$  is estimated from the oceanic mixed layer model as above where  $F(t)$  is derived

from optical thickness data;  $\beta$  is set to  $2/3 \text{ K W}^{-1} \text{ m}^{-2}$ ; and  $C$  is obtained optimally as above. For the observations we use *Sato et al.*'s [1993] optical thickness data. For each model simulation that includes volcanic aerosols we use the reconstruction used to force that simulation [i.e., *Sato et al.*, 1993; *Amman et al.*, 2003], or a very similar reconstruction.

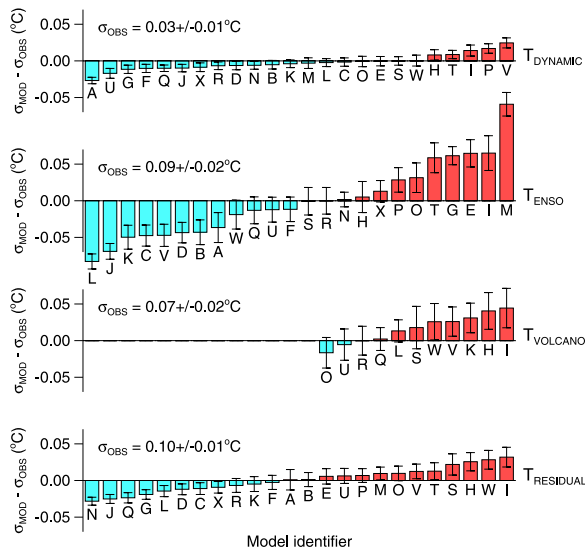
[9] Using  $\hat{T}_{DYNAMIC}(t)$ ,  $\hat{T}_{ENSO}(t)$ , and  $\hat{T}_{VOLCANO}(t)$ , a multivariate regression was carried out using a prescribed first order autoregressive, AR(1), model for the noise  $\varepsilon(t)$ . (The results are insensitive to higher orders of the AR noise model.) In this way we obtain  $T_{GLOBAL}(t) = c + \alpha t + T_{DYNAMIC}(t) + T_{ENSO}(t) + T_{VOLCANO}(t) + \varepsilon(t)$  with  $T_{DYNAMIC}(t) = \mu \hat{T}_{DYNAMIC}$ ,  $T_{ENSO}(t) = \nu \hat{T}_{ENSO}$ , and  $T_{VOLCANO}(t) = \xi \hat{T}_{VOLCANO}$ . The constants  $\alpha$ ,  $\mu$ ,  $\nu$ , and  $\xi$ ; are the regression parameters and  $c$  is a constant. Finally, we compute  $T_{RESIDUAL} = T_{GLOBAL} - T_{DYNAMIC} - T_{ENSO} - T_{VOLCANO} = c + \alpha t + \varepsilon$  where  $T_{RESIDUAL}$  represents the component of  $T_{GLOBAL}$  that is linearly unrelated to the natural climate signals. One of the key advantages of this approach over, say, decadal averaging of the data is that it removes estimates of these components of natural variability without reducing the time resolution of the data.

## 4. Results

### 4.1. Climate Evolution and the Natural Signals

[10] Figure 1 (top) shows the evolution of observed and simulated twentieth century global-mean surface temperature anomalies. As a group the models show good agreement with the observed global-mean warming. Figure 1 (middle) gives the impression that for many models the variance in global-mean temperature associated with the ENSO signal is overestimated, and that there is a significant range of estimates associated with the volcanoes. We will quantify these visual impressions shortly. Figure 1 (bottom) shows that with the natural climate signals removed from the observed global-mean temperature time series there is a large discontinuity around 1945, which has been ascribed to uncorrected biases in the SST record [*Thompson et al.*, 2008]. The discontinuity does not appear in the climate model simulations, and hence the mismatch in global-mean surface temperature for the decade around 1945. Because of this discontinuity all subsequent calculations, including the multivariate regressions, are for the period after 1950.

[11] Figure 2 shows the difference in standard deviation between each model and observations for the natural time series (Figure 2, top) and residual time series with the linear trend excluded (Figure 2, bottom). The bars reflect uncertainties arising from the multivariate regression. Note that the null hypothesis of no significant difference is rejected when the observed and simulated uncertainty ranges taken together exclude zero. The simulated ENSO values show a wide range, with 15 of 24 values deemed inconsistent with observed. Five of the simulated ENSO values are greater than 1.5 times observed, and one is about 2.5 times observed. The simulated volcanic values also show a wide range, with 2 of 11 values deemed inconsistent with observed, at about 1.5 times observed. The simulated dynamically-induced and de-trended residual time series have 7 and 8 values inconsistent with observed, respectively. In short, the models produce a significant range of estimates of the variance in global-mean surface temperature associated with these climate signals.



**Figure 2.** The difference in the standard deviation of temperature between each model and the observations. For the models the standard deviation and uncertainty is the average over all available realizations.

#### 4.2. Trends and Uncertainties

[12] We now turn to an analysis of the trends  $\gamma$  and  $\alpha$  in  $T_{GLOBAL}(t) = d + \gamma t + \eta(t)$  and  $T_{RESIDUAL}(t) = c + \alpha t + \varepsilon(t)$ , respectively, where  $\eta(t)$  and  $\varepsilon(t)$  are the regression residuals assumed to be AR(1) noise. Figure 3a shows that the observations and models display significant warming trends in both  $T_{GLOBAL}$  (red bars) and  $T_{RESIDUAL}$  (blue bars). The observed and model-mean trends in  $T_{GLOBAL}$  are statistically indistinguishable, with values of about  $0.099^\circ\text{C}/\text{decade}$  and  $0.100^\circ\text{C}/\text{decade}$ , respectively. Similarly, in  $T_{RESIDUAL}$  the observed and simulated trends are about  $0.100^\circ\text{C}/\text{decade}$  and  $0.107^\circ\text{C}/\text{decade}$ , respectively. However, the variance across the model trends is about 25% smaller in  $T_{RESIDUAL}$  than in  $T_{GLOBAL}$ , indicating a tighter multi-model estimate absent the natural signals. Figure 3a also shows that of the models with volcanic forcing the majority underestimate the observed trend in  $T_{GLOBAL}$ , but exhibit  $T_{RESIDUAL}$  trends much closer to that observed.

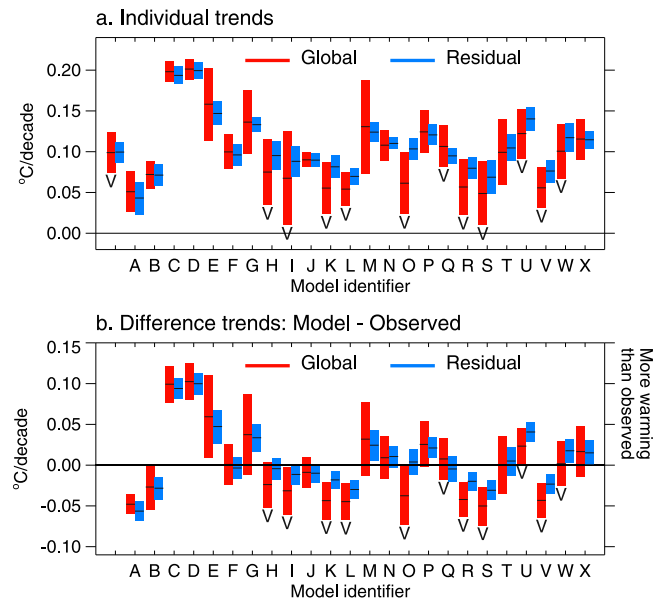
[13] Figure 3a shows that the individual confidence intervals in the  $T_{RESIDUAL}$  are much narrower than in  $T_{GLOBAL}$ , especially for the models that include volcanic forcing. It follows that the degree of consistency between a simulated and observed trend could be different depending on whether or not the natural signals are removed. To illustrate this consider the null hypothesis that there is no significant difference between a model and observed trend – which would be rejected if the confidence interval for the difference time series were not overlapping zero. Here the difference time series is the difference between the given model and observed times series. Figure 3b shows that the outcome of the hypothesis test is changed for 7 models going from  $T_{GLOBAL}$  to  $T_{RESIDUAL}$ . For example, the null hypothesis is accepted in  $T_{GLOBAL}$  of Models B, G, M, P and W, but rejected in  $T_{RESIDUAL}$  because of the drop in uncertainty. On the other hand, the null hypothesis is rejected in  $T_{GLOBAL}$  of Models I and O but accepted in  $T_{RESIDUAL}$  because of the closer to observed trend estimates. It is also notable that there is no

clear relationship between skill in reproducing trends and skill in reproducing natural variance (cf. Figure 2). For example, Models I and T have large ENSO biases but near perfect residual trends. This highlights the difficulty in evaluating the ability of climate models to project future climate change based on their ability to simulate past natural variability.

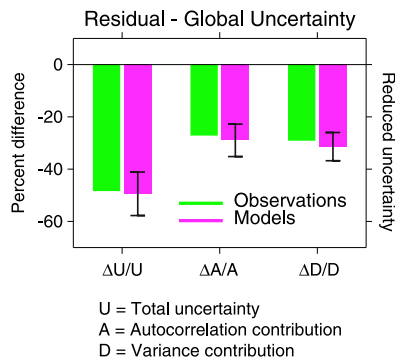
[14] To understand the effect of natural signal removal on trend estimates we appeal to *Santer et al.* [2000, section 4.1] who provide trend and uncertainty formulae. To an excellent approximation these formulae recover the trends and uncertainties in Figure 3a (not shown). Uncertainty  $U$  involves the variance of regression residuals about the trend line  $D^2 = \sum e(t)^2$ , as well as effective sample size  $n_e \approx n(1 - a_1)/(1 + a_1)$ , where  $n$  is the number of time samples and  $a_1$  is the lag-1 autocorrelation of  $e(t)$ . We want to understand how this uncertainty depends partly on the autocorrelation (more autocorrelation means more uncertainty) and partly on the variance around the trend line (more variance means more uncertainty). Accordingly, we use a simple algebraic manipulation of the expressions of *Santer et al.* [2000] to obtain  $U = A(a_1)D$  where  $A(a_1)$  is then, by definition, a monotonically increasing function of  $a_1$ . For small differences in trend uncertainty between the global and residual trends we can approximate that difference as arising from parts related to the autocorrelation and the variance:  $\Delta U/U \approx \Delta A/A + \Delta D/D$ . Figure 4 shows that the observed and model-mean uncertainty is about 50% smaller in  $T_{RESIDUAL}$  than in  $T_{GLOBAL}$ , due in about equal measure to lower autocorrelation and variance effects in the absence of the natural climate signals.

#### 5. Conclusions

[15] A physically-based approach to isolating the signals associated with dynamically induced atmospheric variability,



**Figure 3.** (a) Individual trends. For the models both the mean and uncertainties are realization-means. (b) Difference trends based on the difference between the model and observed time series. A bar not overlapping zero indicates a rejected null hypothesis that there is no significant difference between the given model trend and the observed trend. “V” denotes the presence of volcanic forcing.



**Figure 4.** Percent difference in the trend uncertainty  $U$  (as in Figure 3a) between  $T_{RESIDUAL}$  and  $T_{GLOBAL}$ . As described in section 4 the trend uncertainty  $U$  is such that  $\Delta U/U \sim \Delta A/A + \Delta D/D$ , where  $\Delta A/A$  and  $\Delta D/D$  measure contributions arising from differing autocorrelation and differing variance, respectively. Error bars are multi-model confidence intervals.

El Niño–Southern Oscillation (ENSO), and explosive volcanic eruptions has been used in this study to compare and remove the signals of naturally occurring variability in observed and climate model simulated global-mean surface temperature over the twentieth century. We have made clear the significant model-to-model variability in estimates of the variance in global-mean temperature associated with these natural signals, especially with ENSO and explosive volcanic eruptions. We have also shown that the observed and simulated uncertainty in 1950–2000 trends drops by about half when the natural signals are removed, making clearer where the anthropogenic response in some models deviates significantly from observed. The simulated and observed global-mean temperature trends are statistically indistinguishable in 12 of 24 models for the raw data, but in 8 of 24 models for the residual data. Hence, the methodology employed in this study has allowed us to separate the different contributions to variance in global-mean surface temperature, and this has led us to identify some key strengths and weaknesses in: 1) climate models ability to simulate the observed variance in natural phenomena and 2) climate models ability to simulate the observed global-warming of the second half of the 20th century. Interestingly, the results reveal that there is no clear

relationship between a model’s ability to simulate the variance of natural phenomena and the observed global warming.

[16] **Acknowledgments.** With thank Greg Flato, Bill Merryfield, Slava Kharin, and two anonymous reviewers for their insightful comments. 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 multi-model dataset.

## References

- Ammann, C. M., G. A. Meehl, W. M. Washington, and C. S. Zender (2003), A monthly and latitudinally varying volcanic forcing dataset in simulations of 20th century climate, *Geophys. Res. Lett.*, *30*(12), 1657, doi:10.1029/2003GL016875.
- 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.
- Intergovernmental Panel on Climate Change (2007), *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., 996 pp., Cambridge Univ. Press, Cambridge, U. K.
- Meehl, G. A., C. Covey, T. Delworth, M. Latif, B. McAvaney, J. F. B. Mitchell, R. J. Stouffer, and K. E. Taylor (2007), The WCRP CMIP3 multi-model dataset: A new era in climate change research, *Bull. Am. Meteorol. Soc.*, *88*(9), 1383–1394.
- Santer, B. D., T. M. L. Wigley, J. S. Boyle, D. J. Gaffin, J. J. Hnilo, D. Nychka, D. E. Parker, and K. E. Taylor (2000), Statistical significance of trends and trend differences in layer-average atmospheric temperature time series, *J. Geophys. Res.*, *105*(D6), 7337–7356.
- Sato, M., J. E. Hansen, M. P. McCormick, and J. B. Pollack (1993), Stratospheric aerosol optical depths, 1850–1990, *J. Geophys. Res.*, *98*(D12), 22,987–22,994.
- Thompson, D. W. J., J. J. Kennedy, J. M. Wallace, and P. D. Jones (2008), A large discontinuity in the mid-twentieth century in observed global-mean surface temperature, *Nature*, *453*, 646–649.
- Thompson, W. J., J. M. Wallace, P. D. Jones, and J. J. Kennedy (2009), Identifying signatures of natural climate variability in time series of global-mean surface temperature: Methodology and insights, *J. Clim.*, *22*, 6120–6141.
- Trenberth, K. E., and D. A. Paolino (1980), The Northern Hemisphere sea level pressure data set: Trends, errors and discontinuities, *Mon. Weather Rev.*, *108*, 855–872.

J. C. Fyfe and N. P. Gillett, Canadian Centre for Climate Modelling and Analysis, Environment Canada, University of Victoria, PO Box 1700, Victoria, BC V8W 2Y2, Canada. (john.fyfe@ec.gc.ca)

D. W. J. Thompson, Department of Atmospheric Science, Colorado State University, 200 West Lake St., Fort Collins, CO 80523, USA.