JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 106, NO. D22, PAGES 28,033–28,059, NOVEMBER 27, 2001

Accounting for the effects of volcanoes and ENSO in comparisons of modeled and observed temperature trends

B. D. Santer,[1] T. M. L. Wigley,[2] C. Doutriaux,[1] J. S. Boyle,[1] J. E. Hansen,[3] P. D. Jones,[4] G. A. Meehl,[2] E. Roeckner,[5] S. Sengupta,[6] and K. E. Taylor[1]

Abstract. Several previous studies have attempted to remove the effects of explosive volcanic eruptions and El Niño–Southern Oscillation (ENSO) variability from time series of globally averaged surface and tropospheric temperatures. Such work has largely ignored the nonzero correlation between volcanic signals and ENSO. Here we account for this collinearity using an iterative procedure. We remove estimated volcano and ENSO signals from the observed global mean temperature data, and then calculate trends over 1979–1999 in the residuals. Residual trends are sensitive to the choice of index used for removing ENSO effects and to uncertainties in key volcanic parameters. Despite these sensitivities, residual surface and lower tropospheric (2LT) trends are almost always larger than trends in the raw observational data. After removal of volcano and ENSO effects, the differential warming between the surface and lower troposphere is generally reduced. These results suggest that the net effect of volcanoes and ENSO over 1979–1999 was to reduce globally averaged surface and tropospheric temperatures and cool the troposphere by more than the surface. ENSO and incomplete volcanic forcing effects can hamper reliable assessment of the true correspondence between modeled and observed trends. In the second part of our study, we remove these effects from model data and compare simulated and observed residual trends. Residual temperature trends are not significantly different at the surface. In the lower troposphere the statistical significance of trend differences depends on the experiment considered, the choice of ENSO index, and the volcanic signal decay time. The simulated difference between surface and tropospheric warming rates is significantly smaller than observed in 51 out of 54 cases considered. We also examine multiple realizations of model experiments with relatively complete estimates of natural and anthropogenic forcing. ENSO and volcanic effects are not removed from these integrations. As in the case of residual trends, model and observed raw trends are in good agreement at the surface but differ significantly in terms of the trend differential between the surface and lower troposphere. Observed and simulated lower tropospheric trends are not significantly different in 17 out of 24 cases. Our study highlights the large uncertainties inherent in removing volcano and ENSO effects from atmospheric temperature data. It shows that statistical removal of these effects improves the correspondence between modeled and observed temperature trends over the satellite era. Accounting for volcanoes and ENSO cannot fully explain the observed warming of the surface relative to the lower troposphere, or why this differential warming is not reproduced in the model simulations considered here.

1. Introduction

Recently, several groups have attempted to perform climate model experiments that incorporate a combination of natural and anthropogenic forcings [*Hansen et al.*, 1997a, 1997b; *Bengtsson et al.*, 1999; *Tett et al.*, 1999; *Stott et al.*, 2000]. Over the period relevant for comparison with observations, such experiments yield profiles of vertical temperature change that are markedly different from those simulated in experiments with changes in greenhouse gases only [*Santer et al.*, 1996; *Tett et al.*, 1996; *Hansen et al.*, 1998].

A number of recent studies suggest that improved realism in the mix of applied forcings and in the models themselves [*Gates et al.*, 1999] translates to better statistical agreement between simulated and observed temperature behavior [*Santer et al.*, 1996, 2000a; *Tett et al.*, 1996, 1999; *Hegerl et al.*, 1997; *Wigley et al.*, 1998; *Knutson et al.*, 1999; R. L. Smith et al., A bivariate time series approach to anthropogenic trend detection in hemispheric mean temperatures, submitted to *Journal of Climate*, 2001, hereinafter referred to as Smith et al., submitted manuscript, 2001]. There are, however, problems that

[1]Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, Livermore, California, USA.
[2]National Center for Atmospheric Research, Boulder, Colorado, USA.
[3]NASA Goddard Institute for Space Studies, New York.
[4]Climatic Research Unit, University of East Anglia, Norwich, England, UK.
[5]Max-Planck Institute for Meteorology, Hamburg, Germany.
[6]Engineering Directorate, Lawrence Livermore National Laboratory, Livermore, California, USA.

Copyright 2001 by the American Geophysical Union.

Paper number 2000JD000189.
0148-0227/01/2000JD000189$09.00

complicate comparisons of modeled and observed temperature trends, even in recent coupled model experiments with more realistic forcing estimates.

The first problem relates to different realizations of natural internal variability. Even if a given model realistically captured the observed statistical characteristics and structure of the El Niño–Southern Oscillation (ENSO), the precise timing of ENSO events (and hence of ENSO-induced temperature variability) would differ from that in the real world. A key question is whether the comparison of underlying externally forced signals can be facilitated by identifying and removing ENSO effects from observations and model results.

A second, related problem is the influence of volcanically induced cooling on estimates of ENSO signals. Over the period of the satellite temperature record, two major volcanic eruptions (El Chichón in April 1982 and Pinatubo in June 1991) occurred at the same time as El Niño events (T. M. L. Wigley and B. D. Santer, Differential ENSO and volcanic effects on surface and tropospheric temperatures, submitted to *Journal of Climate*, 2001, hereinafter referred to as Wigley and Santer, submitted manuscript, 2001 (WS)). These eruptions had a marked impact on atmospheric temperature [*Hansen et al.*, 1992, 1993; *Graf et al.*, 1993; *Robock and Mao*, 1995; *Robock*, 2000; WS]. This means that volcano and ENSO temperature signals are not statistically independent over the period of the satellite record. For example, correlations over 1979–1999 between various ENSO indices and the *Sato et al.* [1993] monthly mean estimates of global average stratospheric aerosol optical depth range from 0.384 (SATO versus Niño 3 sea surface temperatures) to 0.464 (SATO versus the Southern Oscillation Index).

The problem of correlation between two or more predictor variables is referred to as collinearity in regression problems [*Mosteller and Tukey*, 1977]. Here it complicates the separation of volcano and ENSO signals. Previous work in this area has either neglected collinearity [*Angell*, 1988; *Jones*, 1994a; *Michaels and Knappenberger*, 2000], or recognized the possibility of collinearity effects, but concluded that they were unimportant [*Christy and McNider*, 1994; *Robock and Mao*, 1995; *Angell*, 2000]. A useful summary of such work is given by WS, who show that collinearity affects estimates of the regression coefficient between various ENSO indices and atmospheric temperature. Thus the first problem alluded to above, the identification and removal of ENSO effects from atmospheric temperature data, is linked to the second problem of removing volcanic effects.

The estimation of volcanic signals is also important in making comparisons between observed temperature changes and model results derived from experiments with incomplete volcanic forcing. A case in point is the GSOP experiment [*Bengtsson et al.*, 1999; *Roeckner et al.*, 1999] performed with the ECHAM4/OPYC model of the Max-Planck Institute for Meteorology. GSOP included stratospheric aerosols produced by the Pinatubo eruption, but neglected the volcanic forcing from El Chichón. The comparison of model-observed trend differences would be facilitated by removal of the Pinatubo signal from GSOP and of El Chichón and Pinatubo effects from the observations.

Our aim here is to apply to observations and relevant model perturbation experiments an automated version of the iterative procedure developed by WS for removing volcano and ENSO effects. We focus on estimating the uncertainties inherent in quantifying these effects, and show that these are much larger than has been assumed in previous work [e.g., *Christy and McNider*, 1994; *Michaels and Knappenberger*, 2000].

The structure of the paper is as follows. In section 2 we introduce the observed and simulated surface and 2LT data used in our study. Section 3 describes a modified version of the iterative volcano/ENSO estimation procedure employed by WS. Results from the application of this procedure to the observations and to data from the ECHAM4/OPYC model are presented in sections 4 and 5, respectively. Section 6 analyzes "raw" atmospheric temperature trends in model experiments with relatively complete estimates of natural and anthropogenic forcing. A summary and conclusions are given in section 7.

2. Data

2.1. Observational Data

The lower tropospheric temperature data used here are from the satellite-based Microwave Sounding Unit (MSU), which uses the upwelling microwave radiation from oxygen molecules to monitor the vertically weighted temperature of deep atmospheric layers [*Spencer and Christy*, 1992]. We rely on a recent version (May 3, 2000) of the MSU lower tropospheric temperature retrieval (2LT) [*Christy et al.*, 1998, 2000]. This differs slightly from MSU data (version d03) we have employed in previous work [*Santer et al.*, 2000a; WS], primarily through inclusion of data from the NOAA 15 Advanced Microwave Sounding Unit (AMSU) after September 1998.

The MSU data were available as monthly means for 2.5° × 2.5° latitude/longitude grid boxes. The data have full global coverage and span the 252-month period January 1979 through December 1999. The peak of the 2LT weighting function is at roughly 740 hPa.

We use near-surface temperature data that are a combination of marine sea surface temperatures (SSTs) and land surface air temperatures (SATs). For full details of the data set, refer to *Jones* [1994b], *Parker et al.* [1994], and *Jones et al.* [1999]. Data were available in the form of monthly mean anomalies (relative to climatological monthly means over 1961–1990) for 5° × 5° latitude/longitude grid boxes, and span the period January 1856 through December 1999.

Our iterative procedure for separation of ENSO and volcanic signals relies on global mean data. We therefore computed area-weighted global mean, monthly mean temperatures from the gridded MSU 2LT and near-surface data. Temperatures were expressed as anomalies relative to climatological monthly means calculated over a common reference period, generally January 1979 through either December 1997 or December 1999.

We use three different indices for removing ENSO effects from global mean 2LT and near-surface temperature data: the Southern Oscillation Index (SOI) and SSTs in the Niño 3 and 3.4 regions (5°S–5°N, 150°W–90°W and 5°S–5°N, 120°W–170°W, respectively). The version of the SOI used here [*Ropelewski and Jones*, 1987; *Können et al.*, 1998] was in the form of monthly means from January 1866 through December 1999. Area-averaged SSTs for the Niño 3 and 3.4 regions were computed from the gridded *Jones et al.* [1999] near-surface temperature data.

2.2. Model Data

We used model data from three sources: the Max-Planck Institute for Meteorology in Hamburg (MPI), the Goddard

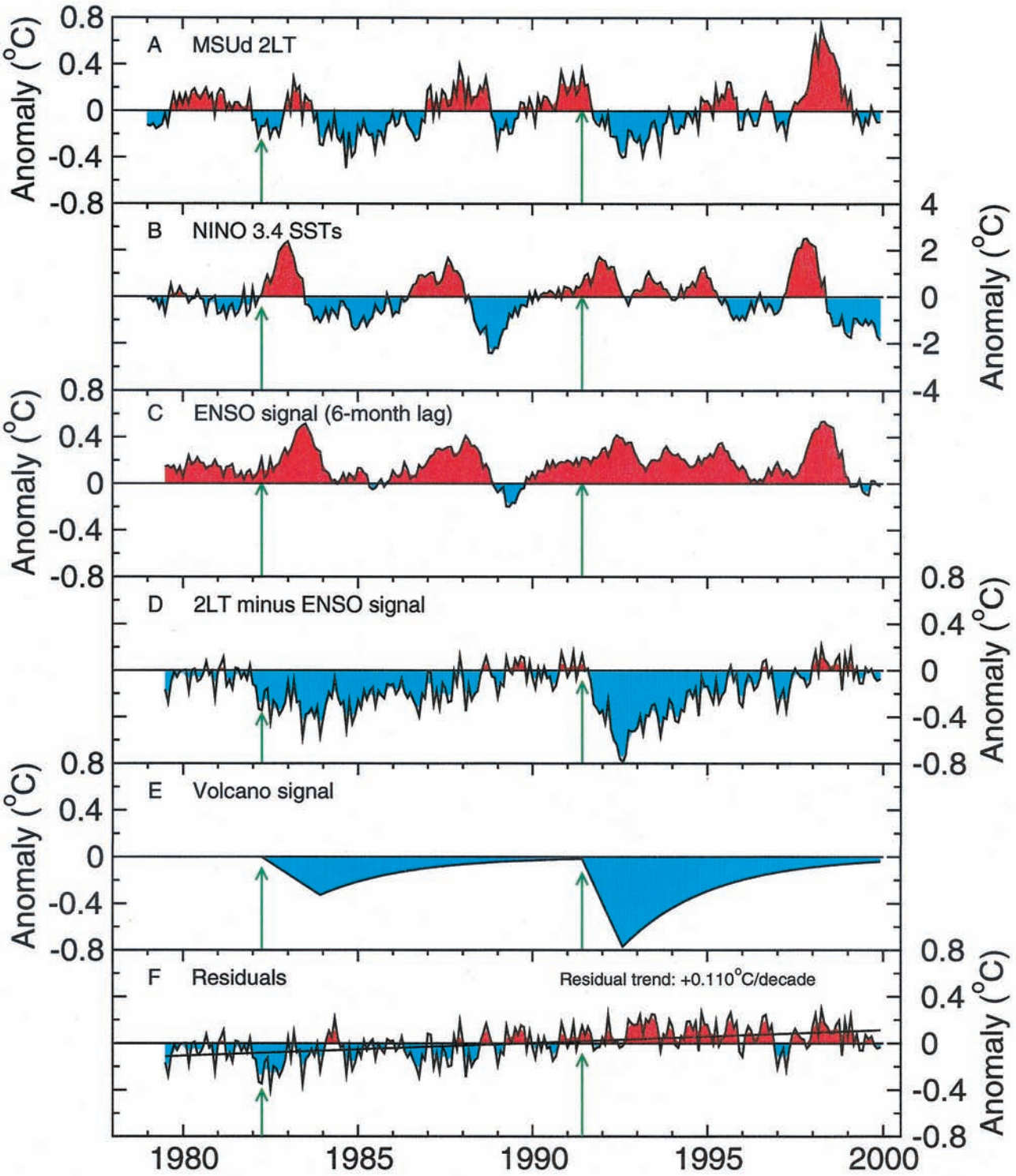

Plate 1. Estimated ENSO and volcanic signals (in °C) in observed (MSUd 2LT) monthly mean lower tropospheric temperatures. (a) Raw MSUd 2LT temperatures [*Christy et al*., 1998] and (b) Niño 3.4 time series [*Jones et al*., 1999]. (c) Regression-based estimate of the ENSO effect on 2LT. Subtraction of the ENSO signal from the raw 2LT data yields the residuals in Plate 1d. These were used to estimate the volcanic signal (Plate 1e). The residuals after subtraction of ENSO and volcanic signals from the raw 2LT data are plotted in Plate 1f. All panels except Plate 1e show both unfiltered results (black lines) and data smoothed with a five-term binomial filter (color fill). The signals (Plates 1c and 1e) and residuals (Plates 1d and 1f) were obtained with an iterative procedure (section 3.2). Results are for the 252-month period January 1979 through December 1999, τ = 30 months, and t_{base} = 6 months. Green vertical lines mark the times of the El Chichón and Pinatubo eruptions.

Institute for Space Studies in New York (GISS), and the National Center for Atmospheric Research in Boulder (NCAR). The MPI data were from perturbation experiments performed with the atmospheric General Circulation Model (AGCM) ECHAM4 (jointly developed by the European Centre for Medium-Range Weather Forecasts and the MPI in Hamburg) coupled to OPYC (an ocean isopycncal model [*Oberhuber*, 1993]). The coupled model and the climate change experiments performed with it have been described in detail by *Bengtsson et al*. [1999] and *Roeckner et al*. [1999]. The ECHAM4 model has 19 atmospheric levels and was run at T42 spectral truncation, equivalent to a horizontal resolution of 250–300 km in the tropics. The OPYC ocean component has comparable horizontal resolution to the atmosphere poleward of 36°, with gradually decreasing meridional grid spacing equatorward of 36°, down to 0.5° at the equator. OPYC has 11 vertical layers. The magnitude of ENSO variability in this model is comparable to observations [*Roeckner et al*., 1996].

In the present study, we analyze the GSOP, GSO1, and GSO2 integrations. All three span the 228-month period from January 1979 through December 1997, and include estimates of historical forcing by well-mixed greenhouse gases, direct and indirect sulfate aerosol effects, and tropospheric and stratospheric ozone. Additionally, GSOP incorporates stratospheric aerosols from the eruption of Pinatubo in June 1991.

The GSO1 and GSO2 experiments use January 1979 oceanic initial conditions from a climate change experiment which commenced in 1860. GSO1 and GSO2 have the same forcing, but start from slightly different atmospheric initial conditions in January 1979. GSOP's initial conditions were taken from June 1991 in GSO1. We also analyzed ECHAM4/OPYC data from a 300-year control integration with no changes in natural or anthropogenic forcings.

The GISS perturbation experiments analyzed here involve combined forcing (from 1979 onward) by well-mixed greenhouse gases, stratospheric aerosols from volcanic eruptions, solar irradiance, and tropospheric and stratospheric ozone [*Hansen et al*., 1997a]. They differ from the ECHAM4/OPYC simulations in their inclusion of solar variability and the effects of El Chichón, and in their exclusion of tropospheric aerosols. Additionally, the GISS experiments incorporate an estimated "initial disequilibrium forcing" of 0.65 W m^{-2} in order to account for the neglect of "unrealized" radiative forcing prior to 1979.

Four different experimental configurations were used. All involve the so-called GISS SI95 AGCM, which has a 4° × 5° latitude/longitude grid and nine vertical levels. In the first configuration the GISS AGCM used prescribed, time-varying SSTs over 1979–1996 and the climatological mean seasonal cycle of sea ice over the decade 1979–1988 [*Hansen et al*., 1997a]. In the second configuration the GISS AGCM was coupled to a mixed-layer ocean with fixed horizontal transport of oceanic heat. Heat anomalies from the mixed layer were diffused into the deeper ocean with a geographically varying diffusion coefficient. The third and fourth configurations were fully coupled A/OGCMs, involving the so-called GISS ocean model [*Russell et al*., 1995], and the GFDL *Bryan and Cox* [1972] ocean model (as implemented by *Miller and Jiang* [1996]).

Ocean components for the third and fourth model configurations both have the same horizontal resolution (4° × 5°) as the GISS SI95 atmospheric GCM. The vertical resolution is 13 and 16 levels in the GISS and GFDL ocean models (respectively). *Miller and Jiang* [1996] have documented that the magnitude of ENSO SST variability in the SI95 AGCM/GFDL OGCM is smaller than observed by a factor of roughly 3.

The original *Hansen et al*. [1997a] experiments covered the 17-year period 1979–1996. Experiments performed with the three coupled configurations were extended through to December 1999. For each configuration there are five independent realizations of the perturbation experiment.

We also analyzed data from a 300-year control integration performed with the Parallel Climate Model (PCM) jointly developed by NCAR and the Los Alamos National Laboratory [*Washington et al*., 2000], and employing version 3 of NCAR's Community Climate Model, CCM3 [*Boville and Hurrell*, 1998]. The atmospheric component of PCM has a T42 spectral truncation (roughly 2.8° × 2.8° horizontal resolution) and 18 vertical levels. The PCM ocean component has 32 vertical levels and 2/3° × 2/3° horizontal resolution, decreasing to 0.5° at the equator. The amplitude of the El Niño/La Niña SST cycle in this model is comparable to observations [*Meehl et al*., 2000].

The comparison of modeled and observed atmospheric temperatures was facilitated in a number of ways. First, we used a static global mean weighting function to compute "equivalent" MSU 2LT temperatures from the ECHAM4/OPYC and PCM data. The GISS equivalent 2LT data were calculated with a radiative transfer code. The two methods give very similar results for large-scale spatial averages, such as global and hemispheric means [*Santer et al*., 1999]. Second, we merged simulated SSTs over ocean and SATs over land to form model near-surface temperature data sets that are more directly comparable with observations. Third, model SSTs were averaged over the same areas as the observations (section 2.1) to form Niño 3 and 3.4 time series. Simulated SOI time series were computed with sea level pressure data from the model grid points closest to Tahiti and Darwin, and were normalized as in the observations. Fourth, temperature and ENSO index time series in models and data were always expressed as anomalies relative to a common reference period (generally either 1979–1997 or 1979–1999).

We do not subsample globally complete model surface data with incomplete observational coverage. Nor do we subsample globally complete modeled and observed 2LT data with observational surface coverage. Accounting for such effects can explain roughly one third of the observed difference between warming rates at the surface and in the lower troposphere [*Santer et al*., 2000a]. Our primary aim here is to quantify uncertainties in estimates of ENSO and volcanic effects on atmospheric temperature. We want to separate this from the issue of coverage differences. Joint consideration of the effects of coverage differences, ENSO, and volcanoes will be addressed in a subsequent paper.

3. Method

3.1. Description of WS Procedure

The WS approach involves successive estimation and removal of volcano and ENSO signals from atmospheric temperature data. In the work of WS the functional form of the global mean temperature response to volcanic forcing is characterized by three parameters: ΔT_{max}, the maximum global mean volcanically induced cooling; t_{ramp}, the time (in months) between the start of the eruption and ΔT_{max}; and τ, the exponential decay time (in months) of the volcanic signal. The first two parameters are estimated directly from the data and are

determined separately for each eruption considered. The third parameter, τ, is model-derived and is identical for each eruption.

There are several reasons why it is difficult to estimate τ from observations. First, the volcano signal is obscured by natural internal variability and by the signals associated with other slowly varying external forcings, such as changes in greenhouse gas concentrations and in solar variability. Second, there is some subjectivity involved in determining the frequency characteristics that should be emphasized if τ is estimated by RMS fitting.

WS obtained estimates of τ in the following way. They used a slightly idealized form of the *Sato et al*. [1993] volcanic forcing data (henceforth SATO) to drive an energy-balance climate model [*Wigley and Raper*, 1992]. As WS point out, the SATO forcing has an e-folding time of roughly 12 months, while the e-folding time of the temperature response is substantially longer due to the thermal inertia of the ocean [*Wigley and Schlesinger*, 1985]. Since the latter depends on climate sensitivity, τ is also a function of uncertainties in estimates of climate sensitivity. WS obtained τ values of 32–39 months for a canonical climate sensitivity range of 1.5°–4.5°C [*Kattenberg et al*., 1996]. This range encompasses estimates of τ obtained from GCMs [*Hansen et al*., 1993] and from observations [*Hansen et al*., 1997b].

Having first obtained a plausible range of estimates for τ, WS then systematically explored the effect of uncertainties in τ on estimates of volcano and ENSO signals. For each of a range of τ values, (20, 30, and 40 months), WS estimated and subtracted a volcano signal from the raw temperature data. Using these residuals, a regression-based approach was used to determine the ENSO effect on temperature. The ENSO signal was then removed from the original temperature data, and the volcano signal was reestimated from the new residuals. This procedure was repeated until convergence. Values of $\Delta T_{\max}$ and t_{ramp} were determined subjectively.

Here, because we have more time series to examine, we decided to automate the determination of $\Delta T_{\max}$ and t_{ramp} (section 3.2). One advantage of the automated procedure is that it is reproducible and common to all data series examined. Implementing an automated approach, however, requires a number of subjective decisions. For example, in order to compute $\Delta T_{\max}$ we need to determine T_{ref}, the preeruption reference level temperature. We estimate T_{ref} by averaging over different time periods of the preeruption temperature data. There is no single "best guess" value for the length of averaging period.

Determination of the maximum cooling after each eruption is also sensitive to the filter choices made. One might decide that $\Delta T_{\max}$ was simply T_{ref} minus the temperature in the coldest month after the eruption (with some realistic cutoff, e.g., 24 months). The problem here is that there is considerable month-to-month variability in the unfiltered temperature data, and the coldest monthly temperature may simply reflect an unusual monthly timescale fluctuation. An alternative would be to estimate the posteruption maximum cooling from low-pass filtered data. This, however, could distort the phase of the peak cooling and markedly reduce the estimated maximum cooling. The difficulty is to find some reasonable compromise in the choice of filter.

We stress, therefore, that the automated approach used here for separation of ENSO and volcano signals, while nominally "objective," in fact relies on a number of subjective decisions, which can markedly influence results. Our procedure complements but is not superior to the "expert judgement" approach of WS.

3.2. Modification of WS Procedure

Let T_t represent either the monthly mean surface or lower tropospheric temperature, and X_t denote some specified index of ENSO variability, such as the SOI. The nominal time index is $t = 1, \ldots, n_t$, with $n_t = 252$ in most of our applications (January 1979 through December 1999). Both T_t and X_t are in the form of anomalies with respect to their climatological monthly means over the full 252-month period.

Step 1 is to select X_t, T_t, and a value of the response time τ. We consider $\tau = 30$ or 40 months, which spans the τ values that WS estimated for a climate sensitivity range of 1.5°–4.5°C.

Step 2 is to estimate T_{ref}, the preeruption "reference level" temperature. This is done for each of the two volcanic eruptions considered (Pinatubo and El Chichón). T_{ref} is the average of the unfiltered T_t values, computed over t_{base} months prior to the eruption. We considered t_{base} values of 4, 6, and 12 months.

Step 3 is to estimate the volcano parameters $\Delta T_{\max}$ and t_{ramp} from the filtered T_t data (see below). The maximum cooling $\Delta T_{\max}$ is

$$\Delta T_{\max} = T_{\text{ref}} - T_{t_{\text{ramp}}}, \tag{1}$$

where t_{ramp} is the time (in months since the eruption month) at which $\Delta T_{\max}$ occurs, with $t_{\text{ramp}} \leq 24$. This restricts the search for $\Delta T_{\max}$ and t_{ramp} to the 24 months following each volcanic eruption. Larger values of t_{ramp} are physically unrealistic.

In the presence of substantial high-frequency noise, it is suboptimal to estimate $\Delta T_{\max}$ and t_{ramp} directly from the raw, unfiltered temperature data. We experimented with a variety of filters and cutoffs, but report only on the results obtained with a five-term binomial filter. Filters with much longer cutoffs can yield excessive smoothing (and hence underestimation of $\Delta T_{\max}$).

Step 4 is to compute the volcano signal V_t, assuming linear cooling from the eruption month to t_{ramp} and exponential recovery thereafter

$$\begin{aligned} V_t &= \frac{-\Delta T_{\max}\, t}{t_{\text{ramp}}} \qquad t = t_1, \ldots, t_{\text{ramp}} \\ &= -\Delta T_{\max}\, e^{-\frac{t - t_{\text{ramp}}}{\tau}} \qquad t > t_{\text{ramp}}. \end{aligned} \tag{2}$$

Step 5 is to remove the volcano signal from the raw temperature data

$$Z_t = T_t - V_t \qquad t = t_1, \ldots, n_t. \tag{3}$$

Step 6 is to compute (as a function of the lag k between X_t and Z_t) the least squares linear regression coefficient b_k between the selected Niño index X_t and the temperature residuals Z_t

$$b_k = \frac{\sum_{t=1}^{n_t-k} [(X_t - \overline{X})(Z_{t+k} - \overline{Z})]}{\sum_{t=1}^{n_t-k} (X_t - \overline{X})^2} \qquad k = 0, \ldots, 24, \tag{4}$$

where $\overline{X}$ and $\overline{Z}$ are the time means of X_t and Z_t, respectively. Then determine the lag j that maximizes $|b_k|$, with $0 \leq j \leq 24$. (Note that r_j, the correlation between X_t and Z_t, is generally maximized at the same lag as the regression coefficient.)

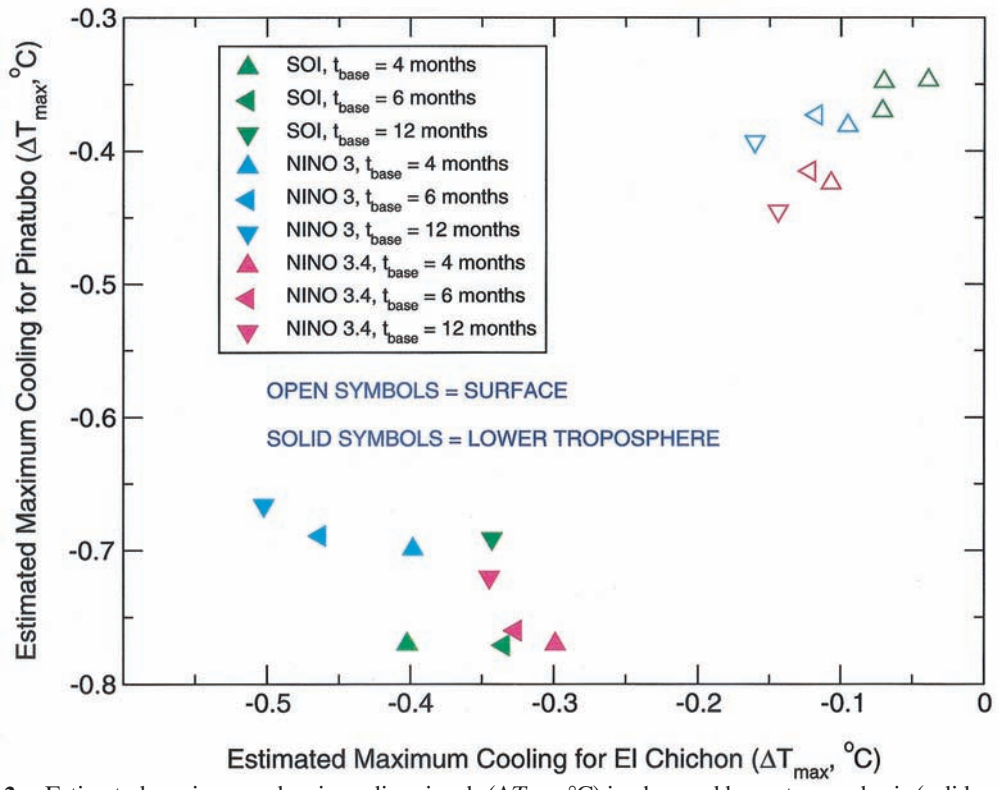

Plate 2. Estimated maximum volcanic cooling signals (ΔT_{max}, °C) in observed lower tropospheric (solid symbols) and surface temperature data (open symbols). An iterative regression-based approach (section 3.2) relying on SOI, Niño 3.4, and Niño 3 time series was employed to remove the effects of ENSO from the monthly mean surface and 2LT data. These residuals were then used to estimate ΔT_{max} for the El Chichón and Pinatubo eruptions. Values of ΔT_{max} are sensitive to t_{base} but relatively insensitive to τ. Results given here are for t_{base} choices of 4, 6, and 12 months, and for τ = 30 months. The analysis period is January 1979 through December 1999.

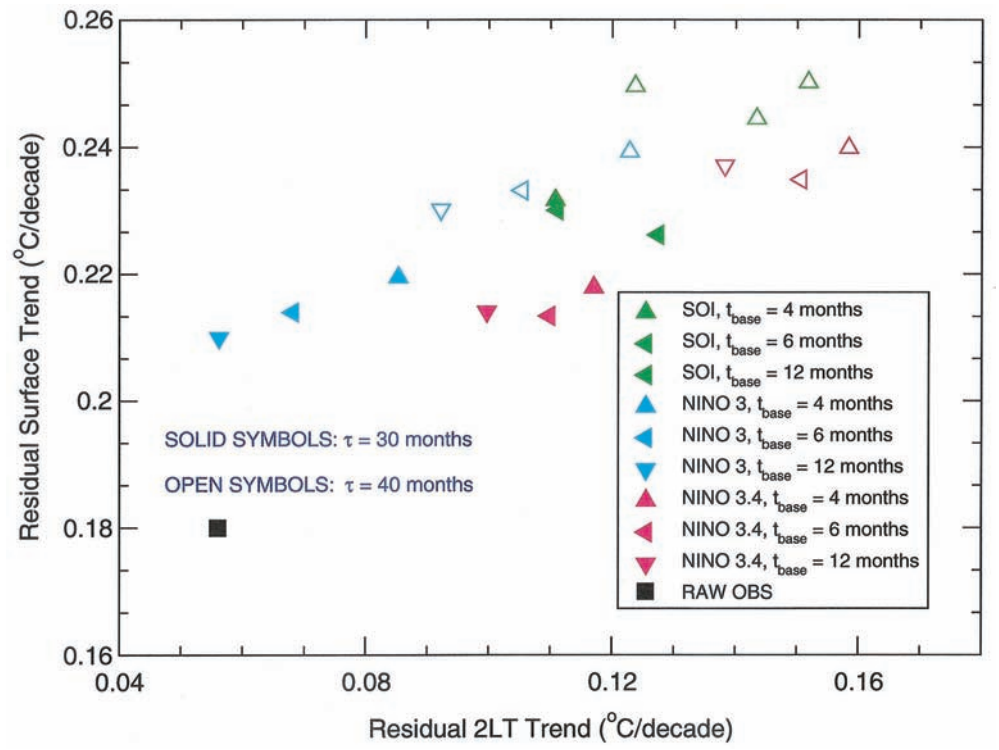

Plate 3. Uncertainties in observed residual temperature trends. Observed surface and lower tropospheric temperature data span the period January 1979 to December 1999. The large range of residual trends arises from uncertainties in the volcano parameters τ and t_{base} and from the choice of index used for removal of ENSO effects. Results for τ values of 30 and 40 months are denoted by solid and open symbols (respectively). Trends in the raw surface and 2LT data ("RAW OBS") are also shown.

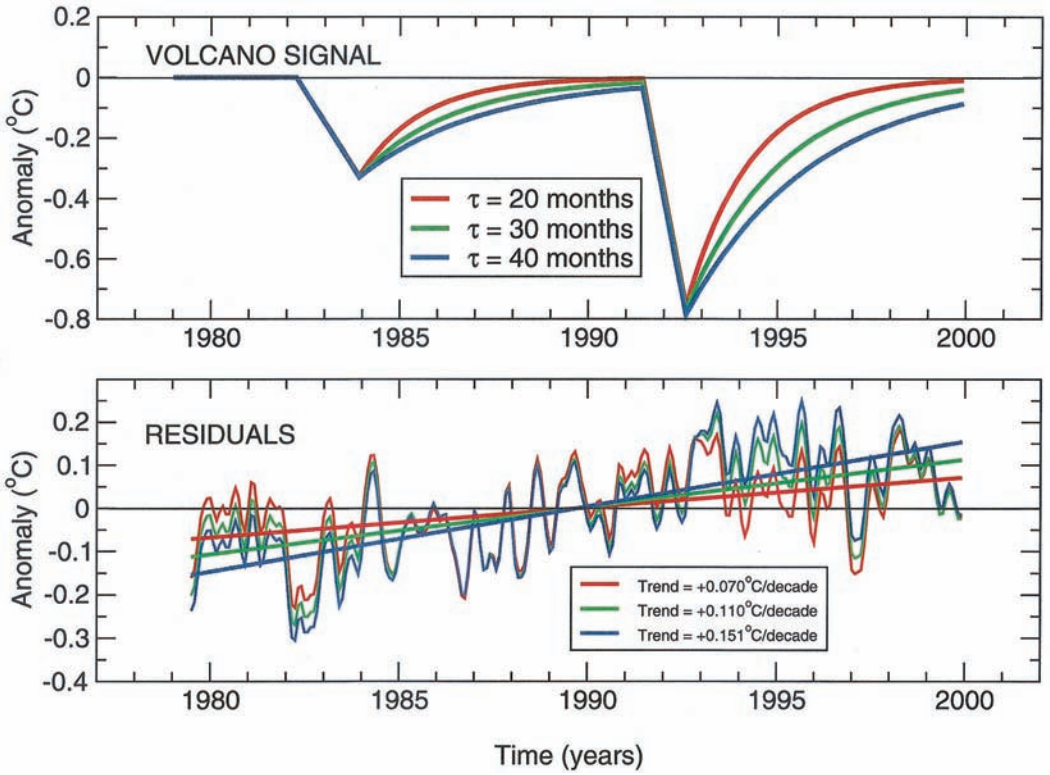

Plate 4. Effect of uncertainties in τ on observed (MSU version d) residual lower tropospheric temperature trends. The top panel shows the estimated volcano signals for three different values of the signal decay time τ (20, 30, and 40 months). Removing these signals and ENSO effects yields the residuals ε_t in the bottom panel (section 3.2). The larger the value of τ, the larger the residual trend. All results are for $t_{\text{base}} = 6$ months and Niño 3.4–based removal of ENSO effects. The analysis period is January 1979 through December 1999. Data in the bottom panel were smoothed with a five-term binomial filter.

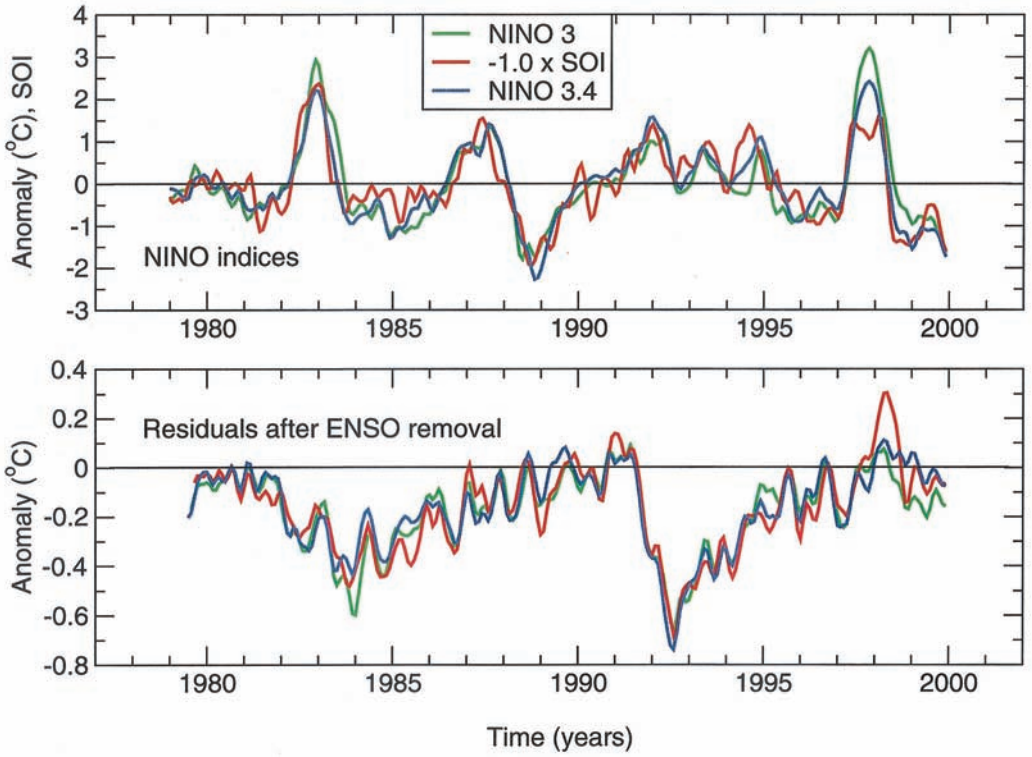

Plate 5. Sensitivity of estimated volcano signal to choice of index for removal of ENSO effects. The top panel shows the three observed ENSO indices used in the iterative method for separating volcano and ENSO signals. Removal of the estimated ENSO signals yields the residual lower tropospheric temperatures T_t^* (bottom panel). All results are for $\tau = 30$ months and $t_{\text{base}} = 6$ months. Data in both panels were smoothed with a five-term binomial filter.

Table 1. Estimated Volcanic Cooling Signals in Observed Lower Tropospheric (2LT) and Surface Temperature Data[a]

		SOI					Niño 3.4					Niño 3				
		El Chichón		Pinatubo			El Chichón		Pinatubo			El Chichón		Pinatubo		
	t_{base}	ΔT_{max}	t_{ramp}	ΔT_{max}	t_{ramp}	C/P	ΔT_{max}	t_{ramp}	ΔT_{max}	t_{ramp}	C/P	ΔT_{max}	t_{ramp}	ΔT_{max}	t_{ramp}	C/P
2LT	4	−0.402	20	−0.770	14	0.52	−0.299	20	−0.770	14	0.39	−0.398	20	−0.699	14	0.57
	6	−0.335	19	−0.771	14	0.44	−0.327	20	−0.760	14	0.43	−0.463	21	−0.689	14	0.67
	12	−0.343	19	−0.691	14	0.50	−0.345	20	−0.720	14	0.48	−0.502	21	−0.666	14	0.75
Surface	4	−0.039	14	−0.347	14	0.11	−0.107	14	−0.424	14	0.25	−0.095	12	−0.381	14	0.25
	6	−0.070	14	−0.348	14	0.20	−0.122	14	−0.415	14	0.29	−0.117	12	−0.373	14	0.31
	12	−0.071	14	−0.370	14	0.19	−0.144	14	−0.445	14	0.32	−0.160	12	−0.393	14	0.41

[a]An iterative regression-based approach (section 3.2) relying on SOI, Niño 3.4, and Niño 3 time series was employed to remove the effects of ENSO from the monthly mean surface and 2LT data. These "ENSO-removed" residuals were then used to estimate ΔT_{max}, the maximum volcano-induced cooling (°C), and t_{ramp}, the lag in months between the start of the eruption and ΔT_{max}. ΔT_{max} and t_{ramp} were determined separately for El Chichón and Pinatubo. Values of ΔT_{max} are sensitive to t_{base}, the length of averaging period used for estimating the mean preeruption temperature, but are relatively insensitive to τ, the assumed signal decay time. Results given here are for t_{base} choices of 4, 6, and 12 months and for $\tau = 30$ months. The ratio of the ΔT_{max} values for the El Chichón and Pinatubo eruptions is also given (C/P). The analysis period is January 1979 through December 1999 (252 months), and sources of observed data are described in section 2.1.

Step 7 is to compute the linear influence of X_t on Z_t

$$E_{t+j} = a_j + b_j X_t \qquad t = 1, \ldots, n_t - j, \tag{5}$$

where a_j is the regression estimate of the y axis intercept at lag j.

Step 8 is to subtract the ENSO effect from the raw temperature data

$$T_t^* = T_t - E_t \qquad t = j + 1, \ldots, n_t. \tag{6}$$

Step 9 is to return to step 2, but now computing volcano parameters T_{ref}, ΔT_{max}, and t_{ramp} from T_t^* rather than from T_t. Continue iterating until there is convergence of the estimated volcano parameters and b_j.

Convergence generally occurs within 4–5 iterations. All results shown in sections 4 and 5 are for 10 iterations. The order in which ENSO and volcano effects are removed from T_t is generally irrelevant as long as sufficient iterations are performed.

Step 10 is at the end of the iterative procedure. Final estimates of the ENSO and volcano signals are subtracted from the original temperature data, yielding residuals ε_t

$$\varepsilon_t = T_t - V_t - E_t \qquad t = j + 1, \ldots, n_t. \tag{7}$$

In the observations the linear trend in ε_t may have contributions from anthropogenic sources (changes in well-mixed greenhouse gases, stratospheric and tropospheric ozone, soot, sulfate aerosols, etc.), from changes in solar forcing, from modes of natural variability other than ENSO [*Thompson and Wallace*, 1998], and from any residual ENSO and volcanic effects that are not removed by our approach.

4. Results From Analyses of Observed Data

Plate 1 provides some insights into the operation of the iterative scheme, in this case with Niño 3.4 data as the ENSO index. The previously noted masking of El Chichón's cooling signature by the 1982/1983 El Niño event [WS; *Angell*, 2000] is clearly shown (Plates 1a and 1b). There is also overlap between a smaller El Niño event and the 1991 Pinatubo eruption. The estimated ENSO effect on lower tropospheric temperature, E_t, is simply a lagged and scaled version of the Niño 3.4 time series (Plate 1c). After removal of E_t from the raw 2LT data, the cooling signatures of El Chichón and Pinatubo are more clearly discernible (Plate 1d). The volcano signal V_t (Plate 1e) is estimated from these residuals. The final residuals after subtraction of ENSO and volcano signals (Plate 1f) display a trend of ~0.11°C/decade, twice the size of the trend in the raw 2LT data.

4.1. Estimated Volcanic Coolings and Ramp Times

Values of ΔT_{max} estimated with the iterative approach are summarized in Plate 2 and Table 1. Results are given as a function of the selected ENSO index and t_{base}. ΔT_{max} is relatively insensitive to uncertainties in τ.

Three conclusions can be drawn from Plate 2. First, irrespective of the choice of ENSO index and uncertainties in the preeruption reference temperature, the estimated maximum coolings for both El Chichón and Pinatubo are always larger in the lower troposphere than at the surface. Second, ΔT_{max} values are sensitive to t_{base} and the choice of ENSO index. The effect of varying t_{base} from 4 to 12 months is not consistent for

Table 2. The 2LT/Surface Ratio of Maximum Coolings for Observed Volcanic Signals[a]

	SOI		Niño 3.4		Niño 3	
t_{base}	El Chichón	Pinatubo	El Chichón	Pinatubo	El Chichón	Pinatubo
4	10.31	2.22	2.79	1.82	4.19	1.83
6	4.79	2.22	2.68	1.83	3.96	1.85
12	4.83	1.87	2.40	1.62	3.14	1.69

[a]Estimated ΔT_{max} values for El Chichón and Pinatubo are taken from Table 1.

El Chichón and Pinatubo, reflecting differences in noise behavior prior to the eruptions (see Plate 1d).

Third, ΔT_{max} is slightly more uncertain for El Chichón than for Pinatubo. This is due to the greater ENSO-induced masking of the El Chichón signal. In the lower troposphere, ΔT_{max} ranges from −0.66° to −0.77°C for Pinatubo and from −0.30° to −0.50°C for El Chichón. At the surface, ΔT_{max} varies from −0.35° to −0.45°C for Pinatubo and from −0.04° to −0.16°C for El Chichón.

The sensitivity of ΔT_{max} and t_{ramp} to the choice of ENSO index has several possible explanations. Different aspects of ENSO's complex space-time variability are captured by indices sampling different geographical areas (such as Niño 3 and 3.4) or different physical quantities (such as SST and sea level pressure). Each index will therefore have a different statistical relationship with global-scale atmospheric temperature. The indices may themselves show different levels of sensitivity to the cooling induced by volcanic eruptions.

The latter issue requires brief elaboration. Both Pinatubo and El Chichón probably cooled SSTs in the regions used to compute the Niño 3 and 3.4 SST indices, thereby modifying pressure gradients, circulation patterns, and hence the SOI. Model experiments with and without volcanic effects suggest that "contamination" of both SST- and pressure-based ENSO indices is possible (section 5.1). The accuracy of our results therefore depends on how well the input ENSO indices capture the "true" ENSO behavior that would have occurred in the absence of the Pinatubo and El Chichón eruptions. This contamination effect is difficult to elucidate in the observations. (Other statistical schemes, incorporating nonlinear interaction terms between volcano and ENSO signals, may be more successful in handling direct volcanic impacts on ENSO indices. Such schemes, however, introduce further unknown parameters, such as the coefficient specifying the strength of the interaction term.)

For El Chichón, estimates of the time between the start of the eruption and the maximum cooling are sensitive to the choice of ENSO index: t_{ramp} varies from 19 months (X_t = SOI) to 21 months (X_t = Niño 3) in the lower troposphere and from 12 to 14 months at the surface (X_t = Niño 3 and SOI, respectively; see Table 1). In contrast, Pinatubo t_{ramp} values are always 14 months, both at the surface and in the 2LT data. The generally longer 2LT t_{ramp} values found here for El Chichón are consistent with the observation that the ramp time for the stratospheric aerosol forcing was slightly longer for El Chichón than for Pinatubo [see *Sato et al*., 1993; WS].

4.2. Estimated El Chichón/Pinatubo Response Ratios

Christy and McNider [1994] assumed that the tropospheric ΔT_{max} values for El Chichón and Pinatubo were nonlinearly related to their respective temperature signals in the lower stratosphere. This analysis yielded a ratio between the ΔT_{max} values for El Chichón and Pinatubo (C/P ratio) of roughly 0.65. A similar ratio was assumed by *Michaels and Knappenberger* [2000]. It is instructive, therefore, to examine the estimated C/P ratios obtained here.

For the parameter space that we consider (18 combinations: 3 ENSO indices × 2 τ values × 3 t_{base} values), the C/P ratio varies by a factor of 2 in the lower troposphere (from 0.39 to 0.75; see Table 1) and by a factor of 3.7 at the surface (from 0.11 to 0.41). Both ranges include the WS "best guess" C/P ratios (0.43 for 2LT and 0.40 for the surface). The C/P ratio estimated by *Christy and McNider* [1994] is obtained only when Niño 3 SSTs are used for ENSO removal.

Note that there are large differences in observationally based estimates of the relative strengths of the El Chichón and Pinatubo eruptions. *Sato et al*. [1993] obtain a ratio of roughly 0.67 for the optical depths of El Chichón and Pinatubo, while *Andronova et al*. [1999, Figure 16a] find a ratio of adjusted net radiative forcing of ~0.44. The C/P forcing ratio used by *Hansen et al*. [1997a] is based on a modified version of the Sato et al. optical depths, with 10% higher optical depths in the 4 years after the El Chichón eruption.

Uncertainties in the C/P forcing ratios arise for several reasons. First, the optical properties of the stratospheric aerosol from the El Chichón eruption are not known as well as those from Pinatubo. While reliable solar occultation measurements for the Pinatubo aerosol are available from the Stratospheric Aerosol and Gas Experiment (SAGE II), comparable measurements do not exist for El Chichón. Second, different groups use procedures of varying sophistication for estimating optical depths from the SAGE data, and for translating these optical depths into estimates of the radiative forcing [cf. *Hansen et al*., 1997a; *Andronova et al*., 1999].

Given these forcing uncertainties, it is difficult to argue that we know a priori what the surface or tropospheric C/P temperature response ratio should be. As pointed out by *Hansen et al*. [1997a], the direct radiative effect of the stratospheric aerosol and of the aerosol-induced reduction in stratospheric ozone [*Portman et al*., 1996; *Solomon*, 1999] is only one component of the volcanic forcing (albeit probably the dominant one). A second component includes the potential dynamic effects resulting from aerosol-induced warming of the lower stratosphere, which could propagate down to the troposphere and surface [*Robock and Mao*, 1995; *Kirchner et al*., 1999; *Robock*, 2000]. This second component introduces further uncertainties in estimating C/P temperature response ratios.

4.3. Surface/2LT Amplification Factors

We next examine (for each volcano) the ratio between the maximum coolings in the lower troposphere and at the surface. The range of amplification factors is larger for El Chichón (from 2.40 to 10.31) than for Pinatubo (from 1.62 to 2.22; see Table 2). This is due to larger uncertainties in estimates of the El Chichón surface and 2LT signals (see section 4.1).

The WS "best guess" estimates of the 2LT/surface amplification factors (1.40 for El Chichón and 1.50 for Pinatubo) are smaller than those obtained here, particularly for El Chichón. The reason for this is that for both volcanic eruptions, WS obtain larger maximum surface coolings than those found here.

The fact that the volcanic signal amplification factor appears to be larger than the amplification of roughly 1.4 for high-frequency variability estimated by *Hurrell and Trenberth* [1998] and WS has several interpretations. The Hurrell and Trenberth variability ratio was computed for tropical oceans between 20°N–20°S, and is almost certainly controlled by thermodynamic processes, primarily the moist adiabatic lapse rate. It is therefore most appropriate over tropical ocean areas. The volcanic signals, however, are not confined to tropical oceans. Dynamical mechanisms can lead to warming of mid and high-latitude land areas in the winter following a volcanic eruption, particularly over Eurasia and Canada [*Robock and Mao*, 1995; *Kirchner et al*., 1999; *Robock*, 2000]. Thus amplification effects in extratropical and/or land areas are controlled by surface and tropospheric processes that may differ from those operating in

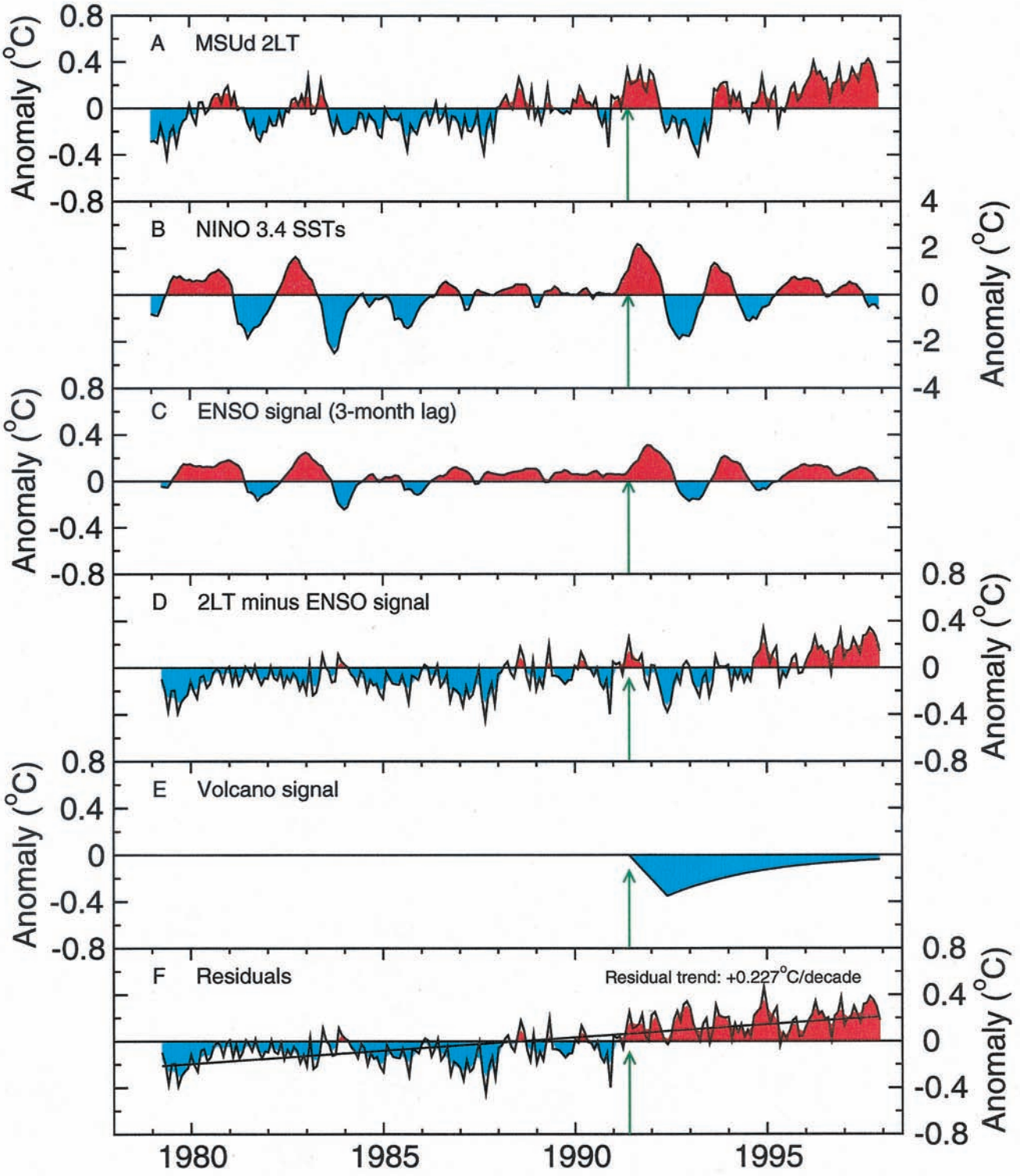

Plate 6. As for Plate 1, but for simulated lower tropospheric temperatures and Niño 3.4 time series from the GSOP experiment performed by *Bengtsson et al*. [1999]. The time period is January 1979 through December 1997 (228 months). Results are for τ = 30 months and t_{base} = 4 months.

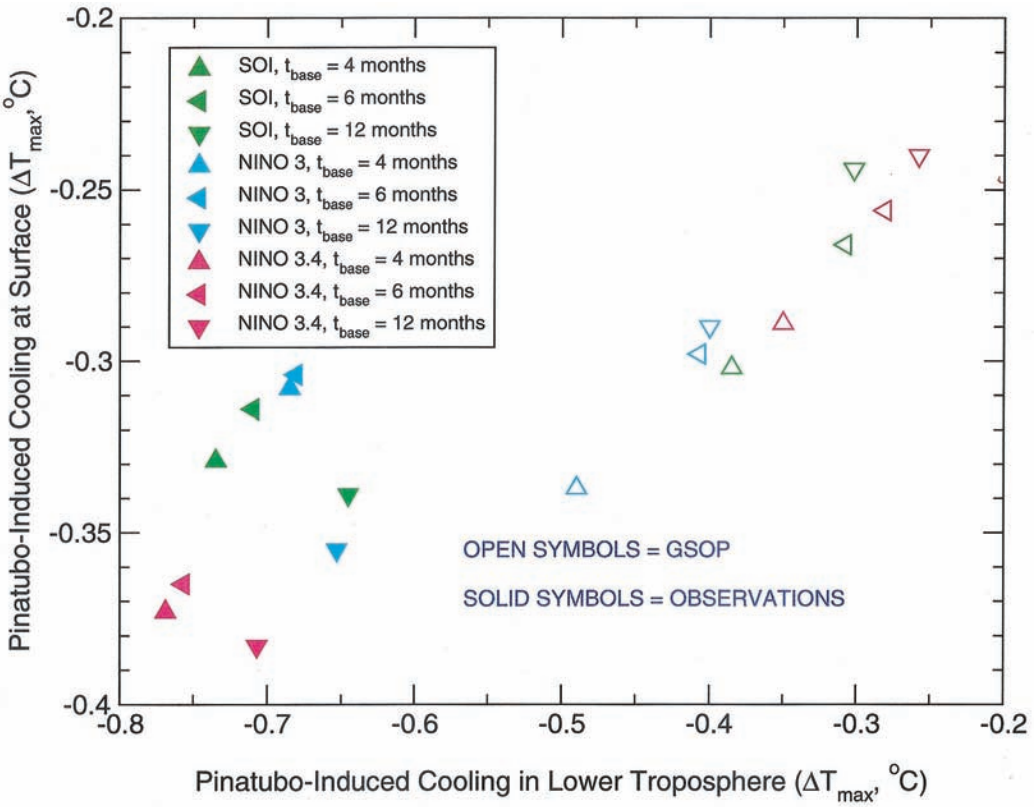

Plate 7. Observed and simulated (GSOP) values of the maximum volcano-induced surface and lower tropospheric cooling (ΔT_{max}, °C) for the Pinatubo eruption. ΔT_{max} values are sensitive to uncertainties in t_{base} and the choice of index for removal of ENSO effects, but are relatively insensitive to τ. All results shown here are for τ = 30 months and for the analysis period January 1979 through December 1997. Observations are denoted by solid symbols; GSOP results are indicated by open symbols.

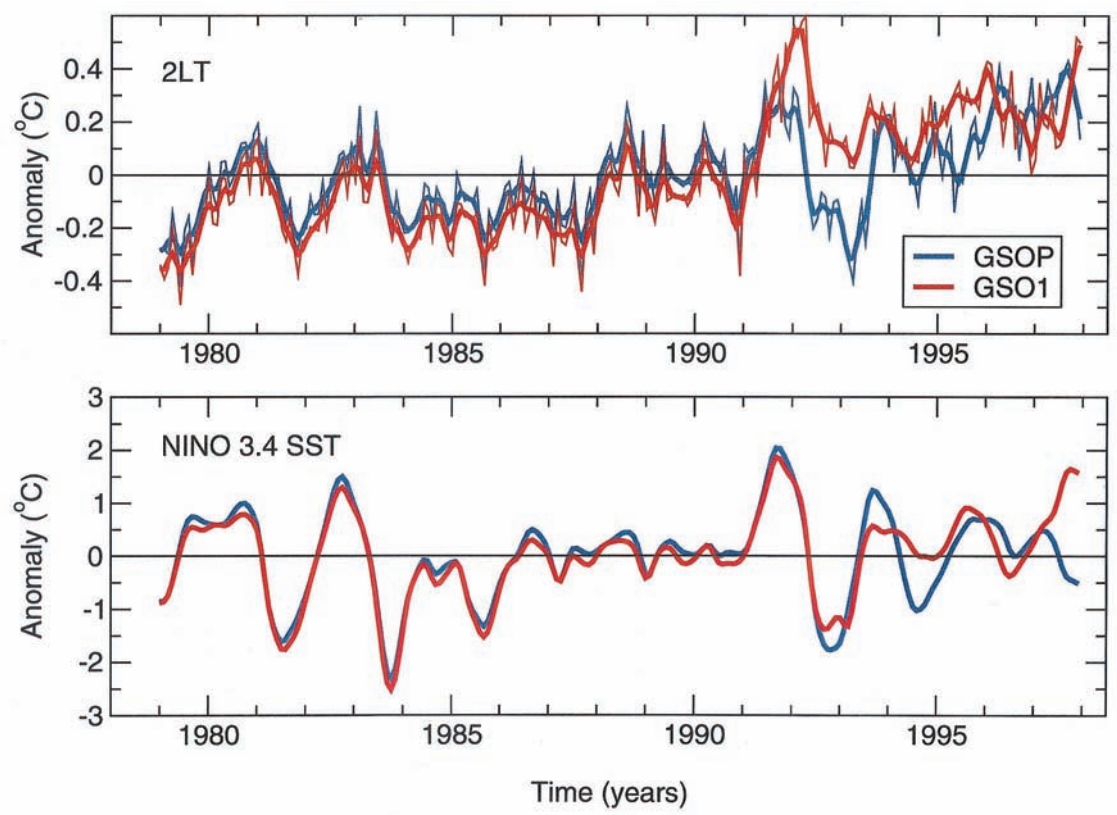

Plate 8. Influence of the Pinatubo volcanic signal on simulated Niño 3.4 SSTs in the *Bengtsson et al.* [1999] GSOP experiment. The raw lower tropospheric temperatures in GSOP and GSO1 (from which GSOP takes its initial conditions) are in the top panel (thin lines). Data smoothed with a five-term binomial filter are also plotted (thick lines). Simulated Niño 3.4 SSTs from both integrations are shown in the bottom panel. Temperatures in the two experiments are identical before June 1991; the slight offsets prior to this time are due to the fact that the climatological monthly means used for generating anomalies were computed over the entire 228-month period.

the tropics. This explains why *Wentz and Schabel* [2000] found amplification factors noticeably lower and higher than 1.4 for the northern and southern extratropics (respectively).

Furthermore, even in the tropics, there may be considerable temporal variability in the scaling between surface and tropospheric temperature anomalies [*Wentz and Schabel*, 2000]. A related issue is that different statistical measures of temperature amplification with height (ratios of 2LT/surface standard deviations, linear trends, etc.) do not yield identical scaling parameters [*Wentz and Schabel*, 2000]. One cannot therefore use the tropical 1.4 scaling of the standard deviations of tropospheric and surface temperature anomalies as a strong constraint on the possible $\Delta T_{\max}$ values for El Chichón and Pinatubo.

4.4. Residual Temperature Trends

Studies by *Christy and McNider* [1994] and *Michaels and Knappenberger* [2000] assume that ENSO and volcano signals can be easily identified and removed from observed lower tropospheric temperature data and that the trend in the residuals is known to within a few hundredths of a degree Celsius per decade. These investigations did not directly address the collinearity between ENSO and volcanic effects, nor did they consider the sensitivity of results to choice of ENSO index, uncertainties in τ, t_{base}, etc. WS provided a more realistic assessment of the uncertainties in estimates of residual surface and 2LT trends, and found large trend sensitivities to uncertainties in τ and to the choice of index used to identify ENSO effects.

Our residual trend results are generally in good agreement with the WS ranges (Plate 3). For the surface our results range from 0.210° to 0.250°C/decade, while WS obtain a range of 0.187° to 0.256°C/decade. The comparable ranges for 2LT residual trends are 0.056° to 0.158°C/decade (present study) and 0.067° to 0.158°C/decade (WS). Uncertainties in the volcanic signal decay time are a key contributor to these large ranges. Larger τ values lead to a larger volcanically induced cooling over 1979–1999, and (after this is removed) to a larger positive trend in the residual temperature data (Plate 4).

The choice of index used to remove ENSO effects also has a large influence on residual trends (Plate 3). While the three ENSO indices behave similarly over much of the record, the signatures of the two largest warm events (in 1982/1983 and 1997/1998) are quite different (Plate 5). During these events the peak warming in the Niño 3 region exceeds the warming in the Niño 3.4 area. As a result, Niño 3–based removal of ENSO effects yields a larger positive value of E_t and a larger negative value of $\Delta T_{\max}$ in 1982/1983. Values of $\Delta T_{\max}$ for Pinatubo are less sensitive to the choice of ENSO index since there is less ENSO-related masking of the Pinatubo cooling signal. The net effect is that Niño 3–based ENSO removal leads to larger C/P ratios, and hence smaller residual temperature trends (Table 1 and Plate 3). In contrast, the SOI has the smallest amplitudes for the two major warm events (Plate 5) and the smallest regression coefficients, which results in a smaller estimated cooling signal for El Chichón and larger residual trends.

In all but one case the residual trends at both the surface and for 2LT are larger than the trends in the raw data (Plate 3). This is in accord with the findings of WS. It is likely, therefore, that volcanic effects reduced trends in both surface and lower tropospheric temperatures over the 21-year period considered here.

Our range of residual trends for the surface-minus-2LT trend difference (not shown) spans 0.082° to 0.151°C/decade and is somewhat larger than that found by WS (0.092° to 0.127°C/decade). This reflects uncertainties arising from the precise ENSO/volcano removal procedures employed, particularly in quantifying the poorly known El Chichón signal. As in the work of WS, accounting for the combined effects of volcanoes and ENSO generally reduces the raw surface-minus-2LT trend difference of 0.124°C/decade. This suggests that over 1979–1999 the net effect of these two influences was to cool the lower troposphere by more than the surface. These effects alone cannot fully explain why the surface warmed relative to the lower troposphere over 1979–1999. There is still a significant difference between observed residual trends at the Earth's surface and in the lower troposphere. (This is evident from the observed results in Plate 10c: after subtraction of ENSO and volcano effects from surface and 2LT time series, the trend in the observed surface-minus-2LT difference series is significantly different from zero in all 18 parameter combinations examined.)

5. Analyses of ECHAM and PCM Model Data

In the following sections, we apply the iterative procedure described in section 3.2 to remove ENSO and volcano effects from the *Bengtsson et al*. [1999] GSOP experiment. We also remove ENSO effects from the GSO1 and GSO2 experiments. Since the latter do not incorporate volcanic aerosols, the ENSO influence on atmospheric temperature can be estimated directly from the raw surface and 2LT data, and an iterative approach is not required. We compare simulated and observed volcano parameters, residual temperature trends, and regression and correlation coefficients. To facilitate this comparison, observed results presented in section 4 are recalculated for the 19-year period 1979–1997. Finally, we analyze data from the ECHAM4/OPYC and PCM control runs in order to investigate the temporal stability of the regression relationship between ENSO indices and atmospheric temperature in the absence of external forcing.

5.1. Model Estimates of Volcano Parameters

Plate 6 shows that collinearity between ENSO and volcanic signals affects the GSOP experiment as well as the observations. In GSOP the Pinatubo-induced cooling is convolved with the signal from large El Niño and La Niña events. As in the observations (section 4.1), estimates of the Pinatubo-induced surface and 2LT maximum coolings are sensitive to t_{base} and the choice of index for removing ENSO effects (Plate 7), but are relatively insensitive to uncertainties in τ. The Pinatubo-induced cooling signal is larger in the lower troposphere than at the surface. This result holds over the entire range of parameter space explored here, both in GSOP and the observations.

Model-based estimates of the lower tropospheric cooling caused by Pinatubo range from −0.26° to −0.49°C and are consistently smaller than in the observations (−0.64° to −0.77°C; Plate 7). Simulated and observed $\Delta T_{\max}$ values are closer at the surface than in the lower troposphere, but model results (ranging from −0.24° to −0.34°C) are still generally smaller than observed estimates (−0.30° to −0.38°C).

Estimates of the observed ramp time for the Pinatubo signal (14 months in the lower troposphere and 13–14 months at the surface) show little variation over the parameter ranges considered here. Observed results fall within the t_{ramp} ranges in

GSOP (12–21 months in the lower troposphere and 11–19 months at the surface). The larger model ranges are due to the fact that ENSO-induced masking of the Pinatubo signal is a more serious problem in GSOP than in observations (compare Plates 6 and 1). Simulated ramp times are therefore more sensitive to choices made in removing ENSO effects.

The consistently smaller $\Delta T_{\max}$ values in GSOP have several possible explanations. First, *Bengtsson et al*. [1999] made linear fits to observed monthly mean latitude-height profiles of stratospheric ozone loss over November 1978 to April 1993. The estimated linear trends for each month, latitude band, and level were then prescribed over 1979 to 1997 in the GSOP experiment. The short-term (1–2 year) enhancement of polar and midlatitude ozone loss that occurred after the El Chichón and Pinatubo eruptions [see *Solomon et al*., 1996] is not captured by such linear fits. This enhanced loss was probably associated with injection of volcanic aerosols into the stratosphere and a consequent increase in the total aerosol surface area available for heterogenous chemical destruction of ozone [*Portman et al*., 1996; *Solomon*, 1999]. Its effect may have amplified the tropospheric and surface cooling responses to El Chichón and Pinatubo.

Neglect of this enhanced ozone loss in GSOP should therefore lead to an underestimate of the surface and tropospheric cooling induced by Pinatubo. It also partly explains why the Pinatubo-induced warming of the lower stratosphere in GSOP is roughly twice as large as observed (the overestimated stratospheric warming is also related to poor simulation of the equatorial cooling caused by an easterly phase of the quasi-biennial oscillation; see *Bengtsson et al*. [1999] for further details).

A second possible reason for underestimation of the Pinatubo signal in GSOP relates to the previously mentioned "contamination" of ENSO indices by volcanic effects (section 4.1). In the model world we can estimate the effect of Pinatubo on ENSO by comparing the behavior of ENSO indices in the GSOP and GSO1 experiments with and without Pinatubo aerosols.

In the ECHAM4/OPYC model the inclusion of the Pinatubo aerosol cloud may have had both direct and indirect effects on the model's ENSO cycle (Plate 8). The direct radiative effect was to cool SSTs. During the peak cooling of the first La Niña event after the eruption (in October/November 1992), GSOP SSTs in the Niño 3.4 region are roughly 0.4°C cooler than in the baseline GSO1 experiment. Because of the radiative cooling effect of Pinatubo on SSTs, we may be overestimating the true ENSO component in GSOP, thereby subtracting too large an ENSO signal (at a time close to that of the peak cooling induced by Pinatubo), and hence underestimating the "true" simulated response to Pinatubo.

The more subtle indirect volcanic effect may have been to modulate the dynamical mechanisms involved in ENSO, hence leading to a post-Pinatubo ENSO trajectory that is very different in GSOP and GSO1 (Plate 8). These differences could also be due to purely stochastic variability. Multiple realizations of the GSOP integration would be necessary to better understand possible volcanic effects on ENSO dynamics.

5.2. Regression and Correlation Coefficients

In this section we examine whether the ECHAM4/OPYC model can reliably portray the observed statistical relationships between ENSO variability and global mean temperature. Consistency between models and observations would enhance our confidence in (1) the usefulness of our iterative procedure for separation of ENSO and volcanic signals; and (2) the ability of the model to reproduce the observed regression and correlation relationships between ENSO variability and fluctuations in atmospheric temperature. Conversely, fundamental differences between observed and model-based regression and correlation coefficients (b_j and r_j; see section 3.2) could be indicative of some combination of model errors, missing forcing, errors in the observed data, or deficiencies in our method for separating ENSO and volcanic effects.

To address this issue, we estimate the sampling distributions of b_j and r_j obtained from overlapping 228-month segments of the ECHAM4/OPYC control run. We consider whether the observed regression and correlation coefficients are contained within the model-generated sampling distributions. Model regression and correlation coefficients are estimated directly from the control run data without an iterative procedure since the control run does not incorporate volcanic effects.

Observational estimates of b_j and r_j are obtained as described in section 3.2. These are given for two periods (1979–1997 and 1979–1999) to show the influence of sampling variability.

If Niño 3 or Niño 3.4 SSTs are used for ENSO removal, observed values of b_j and r_j are always contained within the model-generated sampling distributions (see Plate 9 for Niño 3.4 results). The degree of overlap between modeled and observed regression and correlation coefficients is sensitive to the exact observational period considered. In both the model and observational results, Niño SSTs are more highly correlated with lower tropospheric than with surface temperatures. Regression coefficients are also higher for the lower troposphere than for the surface.

Rather different results are obtained if ENSO removal is based on the simulated SOI. Observed regression and correlation coefficients are virtually always larger than those simulated in the ECHAM4/OPYC control run (not shown). This may in part be due to the way we have computed the SOI in the model. We use mean sea level pressure information from the grid points closest to Tahiti and Darwin, rather than from large areal averages centered on Tahiti and Darwin. This constitutes a very stringent test of model performance. We note, however, that a simulated SOI index based on the difference between large areal averages also yields a weaker than observed correlation between the SOI and surface temperature in ECHAM4/OPYC (K. Achuta Rao et al., El Niño Southern Oscillation in coupled GCMs, submitted to *Climate Dynamics*, 2001).

5.3. Residual Trends After ENSO and Volcano Removal

Residual trends for the GSOP, GSO1, and GSO2 integrations are shown in Plates 10–12 (respectively) as a function of choice of index for removing ENSO effects and uncertainties in τ and t_{base}. We also provide estimates of the statistical significance of model-versus-observed trend differences. Significance was assessed as in the work of *Santer et al*. [2000a], using

$$d = (b_o - b_m)/[s_{bo}^2 + s_{bm}^2]^{1/2}, \tag{8}$$

where b_o and b_m are the observed and modeled trends with respective standard errors s_{bo} and s_{bm}. Standard errors are adjusted for temporal autocorrelation effects [*Santer et al*., 2000b]. We assume that the normalized trend difference d has a Gaussian distribution and that values of $d > 1.96$ (2.58) indicate trend differences significant at the 5% (1%) level. In

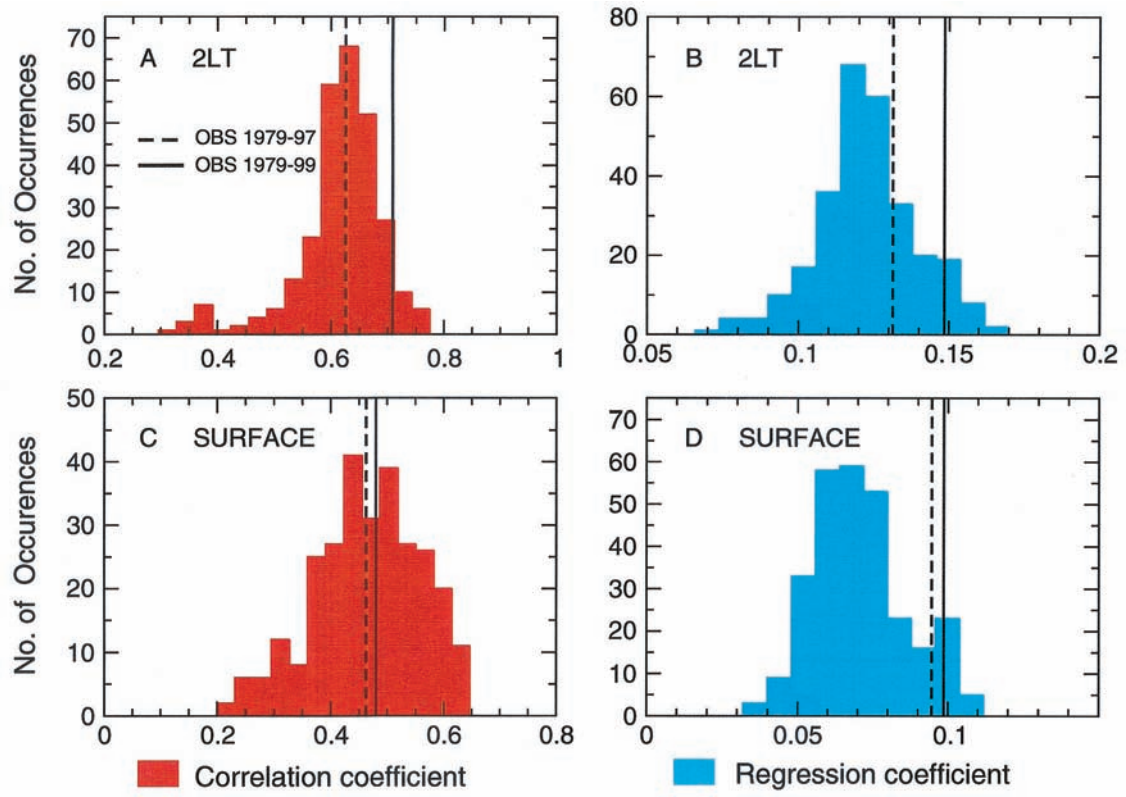

Plate 9. Observed and simulated correlation and regression coefficients (r_j and b_j) between Niño 3.4 indices and atmospheric temperatures. Model sampling distributions were computed from overlapping 228-month segments of the 300-year ECHAM4/OPYC control run, with overlap by 216 months. Observed r_j and b_j values are given for two periods (1979–1997 and 1979–1999) and are averages over the six combinations of τ and t_{base} considered here. (a and c) Correlation coefficients between Niño 3.4 SSTs and lower tropospheric (surface) temperatures. (b and d) Corresponding regression coefficient results.

testing the significance of trend differences between GSOP and observations, we also make the simplifying assumption that τ and t_{base} values are similar in the model and in the real world. This means that only trend pairs with equivalent processing choices are tested: e.g., if b_o is computed with $\tau = 30$ months, $t_{\text{base}} = 6$ months, and SOI-based removal of ENSO effects, b_m must be based on the same parameter choices. Since GSO1 and GSO2 do not include volcanic effects, their residual trends after removal of ENSO effects do not depend on τ or t_{base}. The GSO1 and GSO2 residual trends are therefore compared with b_o results for each of the six possible parameter combinations considered here (i.e., two τ values $\times$ three t_{base} values for any given ENSO index).

5.3.1. Surface trends. There is good quantitative agreement between modeled and observed surface trends in all three perturbation experiments (Plates 10–12). None of the 54 model-based residual trends is significantly different from observed values (at the 5% level) (54 = three model experiments $\times$ two τ values $\times$ three t_{base} values $\times$ three ENSO indices). The only statistically significant difference is between the "raw" GSO1 surface trend (0.256°C/decade) and the raw observed trend (0.154°C/decade; Plate 11).

For the parameter space considered here, observed residual trends range from 0.202° to 0.236°C/decade. Model ranges are 0.220° to 0.282°C/decade (GSOP), 0.210° to 0.249°C/decade (GSO1), and 0.178° to 0.182°C/decade (GSO2). The very small range in GSO2 is due to the fact that, by chance, the ENSO component of temperature variability is very small in this experiment. This results in a small sensitivity to the choice of index for removing ENSO effects.

5.3.2. Lower tropospheric trends. In all three experiments, simulated trends in "raw" lower tropospheric temperatures are significantly larger than the observed raw trend (Plates 10–12). Without removal of ENSO and volcano effects, differences between observed and simulated raw 2LT trends are 0.161°, 0.280°, and 0.214°C/decade (for GSOP, GSO1, and GSO2, respectively). With removal of these effects the differences between observed and simulated residual 2LT trends in Plates 10–12 range from 0.056° to 0.160°C/decade (GSOP), 0.043° to 0.182°C/decade (GSO1), and 0.044° to 0.127°C/decade (GSO2). In all three ECHAM4/OPYC perturbation experiments, accounting for the effects of volcanoes and ENSO helps to resolve some of the discrepancy between "raw" simulated and observed lower tropospheric temperature trends. This conclusion holds over the full range of parameter space explored here.

Decisions on the statistical significance of differences in residual trends are primarily dependent on the experiment considered, on τ, and on the choice of index used to remove ENSO effects. In GSOP, simulated and observed residual 2LT trends are significantly different at the 5% level or better in 16 out of 18 cases. Trend differences between the observations and GSO1 (GSO2) are significantly different at this level in 15 of 18 (14 of 18) cases.

Agreement between modeled and observed residual 2LT trends is poorer for shorter volcanic signal decay times and for

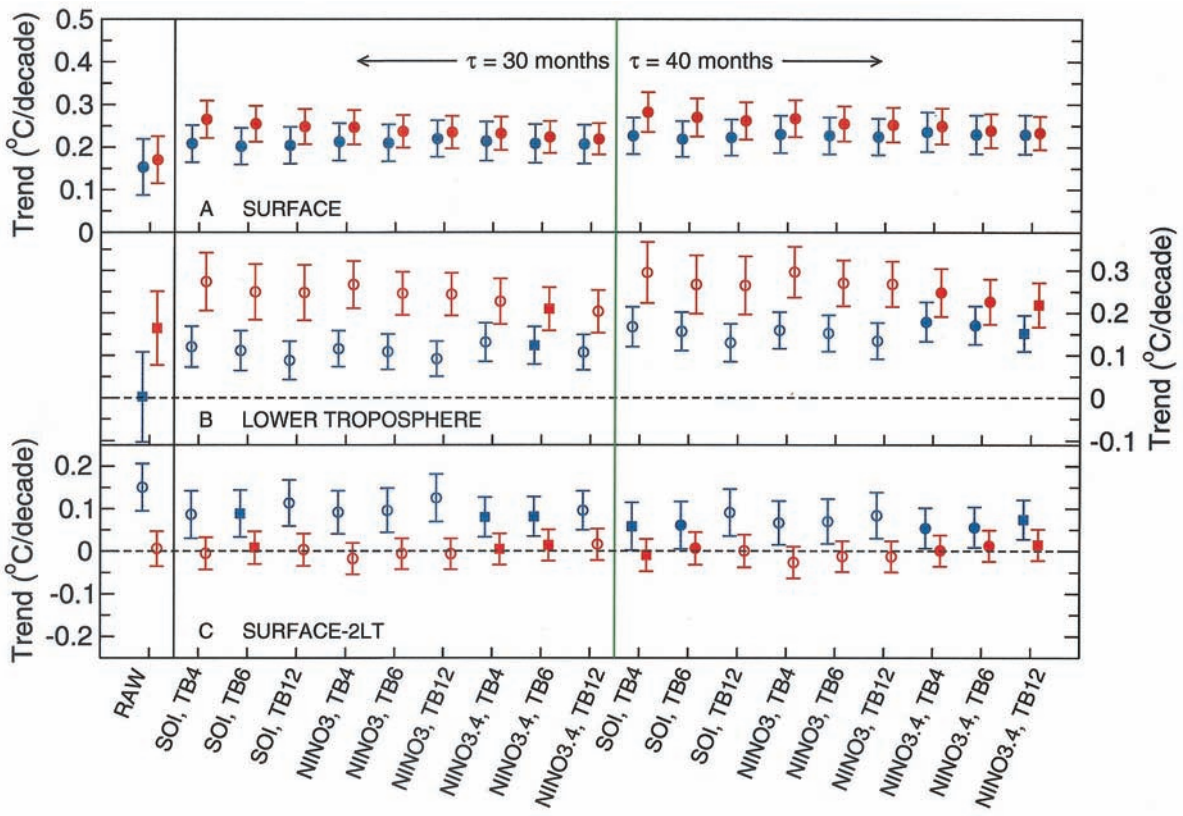

Plate 10. Comparison of modeled (GSOP) and observed temperature trends (in °C/decade) over 1979–1997 in (a) surface and (b) lower tropospheric temperature data, and in (c) the surface-minus-2LT trend difference. Model results are in red; observations are in blue. The results marked "RAW" include volcano and ENSO effects. In all other cases these effects have been removed with an iterative method, and comparisons involve residual trends. Residual trend results are sensitive to processing options: the choice of ENSO index (SOI, Niño 3, or Niño 3.4 SSTs), τ (30 or 40 months), and t_{base} (4, 6, or 12 months; TB4, TB6, TB12). Results left (right) of the green line are for τ values of 30 (40) months. Error bars indicate the 2σ standard errors, adjusted for temporal autocorrelation effects [*Santer et al.*, 2000b]. Open circles (solid squares) denote model trends that are significantly different from observations at the 1% (5%) level. Trends not significantly different at the 5% level are marked with solid circles. For further details, see section 5.3.

removal of ENSO effects with the SOI. The latter effect requires brief explanation. If Niño 3.4 SSTs are used as the basis for ENSO removal, both GSOP and GSO1 integrations have positive ENSO-induced contributions to the overall 2LT trend (0.023° and 0.056°C/decade, respectively; see Table 3). Trends in the ENSO component E_t are much smaller if the SOI is selected as the ENSO index (−0.001° and 0.009°C/decade), primarily because ECHAM4/OPYC underestimates the observed regression and correlation coefficients between the SOI and lower tropospheric temperatures (see section 5.2). SOI-based removal of ENSO effects leads to smaller subtracted ENSO components, larger residual 2LT trends in GSOP and GSO1, and hence enhances the statistical significance of model/observed residual trend differences. We have greater confidence in significance results based on use of Niño 3 and 3.4 SSTs for ENSO removal, since ECHAM4/OPYC reproduces observed statistical relationships between Niño 3/3.4 SSTs and atmospheric temperature with better fidelity than SOI/temperature relationships.

Table 3. ENSO Component of Overall Trends in Observed and Simulated Atmospheric Temperature Data[a]

	ENSO Index	GSOP	GSO1	GSO2	OBS
2LT	SOI	−0.001	0.009	−0.009	0.000
	Niño 3	0.020	0.037	−0.005	−0.012
	Niño 3.4	0.023	0.056	−0.011	0.018
Surface	SOI	−0.002	0.008	−0.003	0.008
	Niño 3	0.013	0.030	−0.001	0.009
	Niño 3.4	0.017	0.047	−0.005	0.020

[a]Results are trends (°C/decade) in E_t, the ENSO component estimated directly from raw temperature data (GSO1, GSO2), or after iterative separation of volcano and ENSO effects (GSOP, OBS). The SOI and Niño 3 and 3.4 SSTs were used to estimate ENSO effects. All calculations were performed with 228 months of data (January 1979 through December 1997). Processing choices for the iterative method were τ = 30 months and t_{base} = 6 months. GSOP and OBS results are relatively insensitive to uncertainties in τ and the preeruption reference level temperature.

The present comparison of 2LT trends is an improvement on that given by *Santer et al*. [2000a]. Here trends are compared over a common time period, explicitly accounting for incomplete volcanic forcing in GSOP and for different realizations of ENSO in models and data (see Appendix A).

5.3.3. Surface minus 2LT trend difference. The raw observational temperature data generally show larger warming at the surface than aloft over the period of the satellite record [*NRC*, 2000; *Gaffen et al*., 2000; WS]. Most model experiments are unable to reproduce this result. This discrepancy cannot be fully explained by accounting for coverage differences and/or "unforced" natural internal variability [*Santer et al*., 2000a]. Explicitly accounting for the twin effects of ENSO and recent volcanic eruptions helps to reduce the difference between the surface and lower tropospheric warming trends in the observations (WS; section 4.4). It also explains some but not all of the discrepancy between the behavior of surface and tropospheric warming rates in models and data (see WS).

Our results are generally in accord with the earlier findings of WS and *Santer et al*. [2000a]. Before removal of volcano and ENSO effects, the trend in the raw surface-minus-2LT difference series is strongly positive in the observations (+0.150°C/decade), but close to zero in GSOP (+0.006°C/decade; Plate 10) and negative in GSO1 and GSO2 (−0.027° and −0.040°C/decade; Plates 11 and 12). Thus the differences between raw results in observations and simulations are 0.144°, 0.177°, and 0.190°C/decade (for GSOP, GSO1, and GSO2, respectively).

Accounting for volcano and ENSO effects invariably reduces these discrepancies. Consider the example of GSOP (Plate 10). Simulated residual trends in the surface-minus-2LT difference series range from −0.026° to +0.016°C/decade. Ten of these results are positive in GSOP, behavior that is qualitatively similar to observations.

In the observations, trends in the residual surface-minus-2LT difference series span the range +0.054° to +0.125°C/decade. Subtracting model results from their observed "counterparts" (with identical τ, t_{base}, and ENSO index choices) yields differences between observed and simulated residual results that range from 0.043° to 0.132°C/decade. All of these values are less than the above mentioned raw observed-minus-GSOP result (0.144°C/decade).

Nevertheless, differences between modeled and observed residual surface-minus-2LT trends are still significant at the 5% level or better in 51 out of 54 cases (Plates 10–12). We conclude from this that accounting for model/data differences in volcanic and ENSO effects helps to explain some but not all of the discrepancy between modeled and observed surface and lower tropospheric warming rates over the last few decades.

5.4. ENSO-Related Trends

In the previous section our focus was on residual temperature trends after removal of ENSO and volcano effects. It is also of interest to examine the ENSO-induced trend components that we have removed, which are simply the trends in E_t (see section 3). These components are sensitive to the choice of ENSO index (see Table 3). We focus here on discussion of results obtained with Niño 3.4 SSTs for reasons discussed in section 5.2.

For the simulated 2LT data in GSOP the ENSO contribution to the overall trend is positive (0.023°C/decade) and similar to the observed value (0.018°C/decade) over 1979–1997 (Table 3). In GSO1 and GSO2 the ENSO-related trend is 0.056° and −0.011°C/decade, respectively. The latter results illustrate how a minor difference in the atmospheric initial conditions (the oceanic initial conditions were identical) can lead to entirely different evolutions of ENSO variability.

Averaging over many independent realizations of the same experiment would presumably lead to a net zero influence of simulated ENSO variability on atmospheric temperature trends (unless the character of ENSO were changed due to some imposed external forcing [see, e.g., *Knutson et al*., 1997; *Timmermann et al*., 1999]). In any individual realization, however, the ENSO contribution could be substantially different from zero, as it is here for GSOP and GSO1. The assumption by *Michaels and Knappenberger* [2000] that the ENSO component of temperature trends is negligibly small in any given model experiment is incorrect.

One way of investigating the variability in ENSO-related trends is by using overlapping 228-month segments of the 300-year ECHAM4/OPYC control run. This allowed us to derive sampling distributions of trends in both the raw lower tropospheric temperatures and in the ENSO components of the 2LT data (see Plate 13).

The mean of the sampling distribution of the ENSO-induced trend component is close to the expected value of zero. However, the ENSO contribution to 2LT trends in individual 228-month segments of the control can be as large as ±0.09°C/decade. Differences in the "spread" of the two sampling distributions in Plate 13 indicate that factors other than ENSO must also make a substantial contribution to the overall decadal timescale variability in 2LT trends. The relative contributions of ENSO and other sources of natural variability change markedly over time in the ECHAM control run (Plate 14). For example, ENSO explains very little of the variability in 2LT trends with start dates between circa 1980 and 2000, but explains virtually all of the variability in 2LT trends starting between circa 2065 and 2080.

Our results illustrate the need for ensemble calculations to obtain more reliable estimates of the simulated response to external forcing. If the ECHAM control run results are credible, statistical removal of ENSO effects facilitates the comparison of modeled and observed trends, particularly if only one or two realizations of a model experiment are available. Removal of ENSO effects alone, however, fails to account for differences between modeled and observed noise realizations that are related to other modes of variability.

5.5. Temporal Stability of Regression and Correlation Coefficients

Our iterative method for the separation of ENSO and volcanic signals assumes that the regression relationship between a selected ENSO index and atmospheric temperature is linear and relatively stable on the timescale of the analysis period. In the observations the regression coefficient between the SOI and globally averaged near-surface temperature can vary markedly on decadal and longer timescales, even after estimated volcanic effects have been removed from the temperature data [*Wigley*, 2000]. Owing to the above described uncertainties in removing volcanic effects from observational data, it is useful to examine whether substantial "unforced" temporal variability of b_j and r_j also occurs in long model control runs with no volcanic effects.

Plate 15 shows time series of b_j and r_j computed from overlapping 228-month segments of the 300-year ECHAM4/OPYC and PCM control runs. In both integrations, running regression and correlation coefficients can change by as much

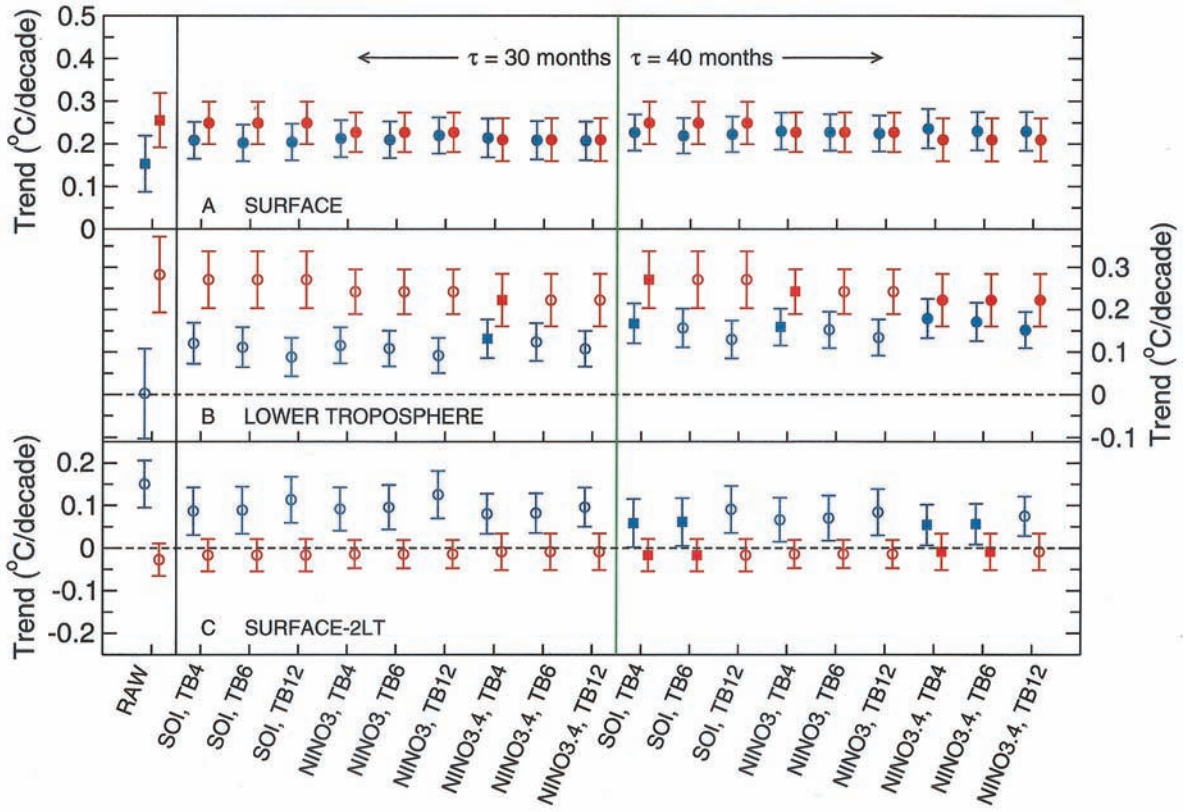

Plate 11. As for Plate 10, but for the GSO1 experiment. Unlike GSOP, residual trends in GSO1 were estimated without an iterative approach. Model values depend only on the choice of ENSO index.

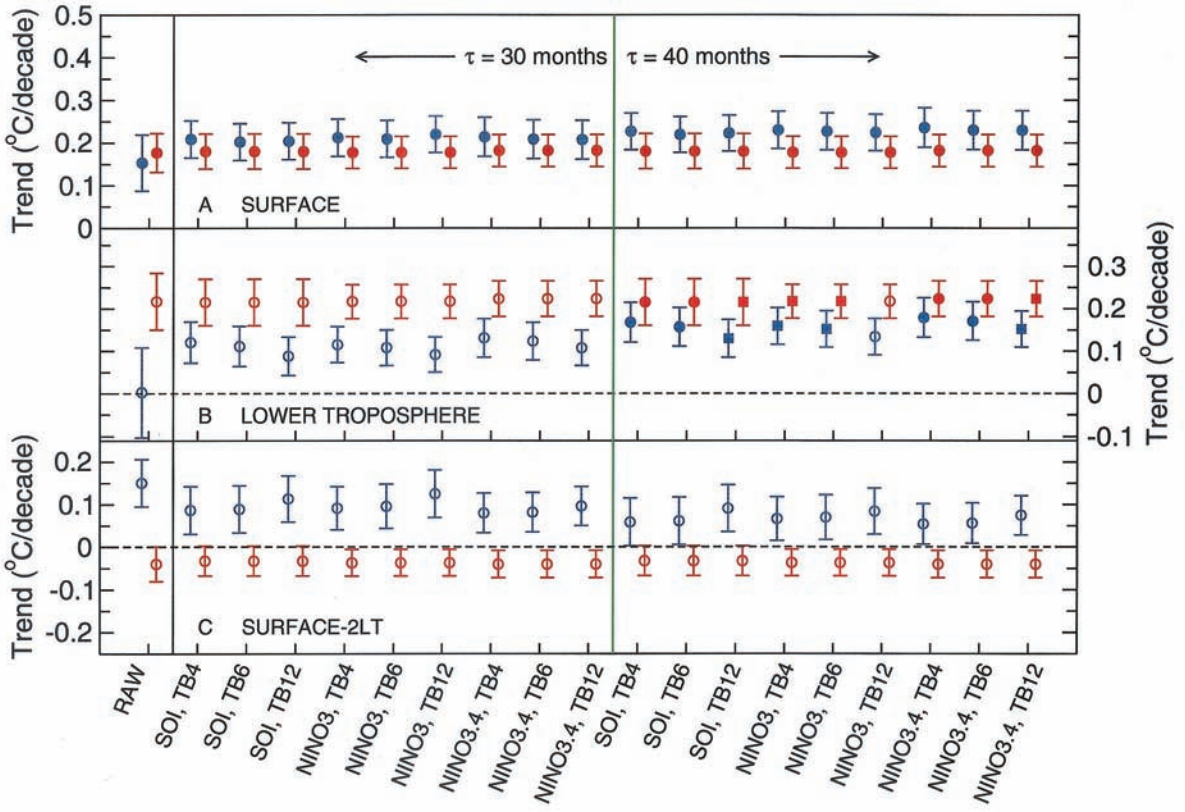

Plate 12. As for Plate 11, but for the GSO2 experiment.

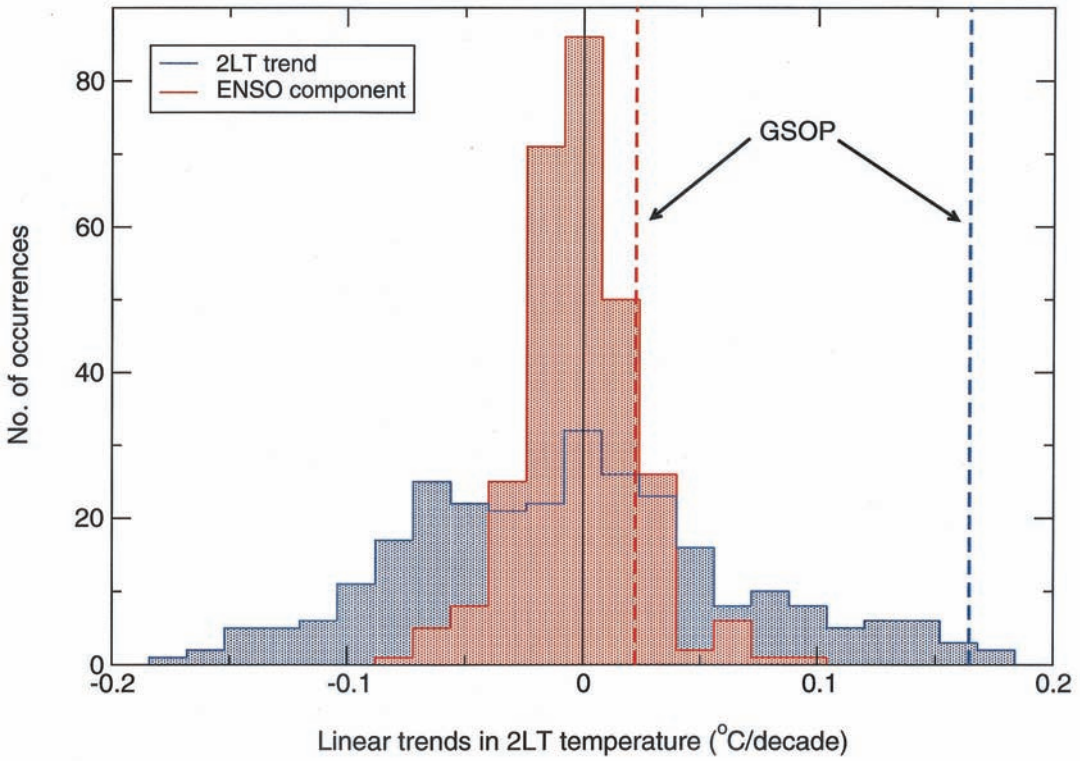

Plate 13. Sampling distribution of 19-year temperature trends in the 300-year ECHAM4/OPYC control run. Distributions are shown for 2LT trends (blue) and for the component of 2LT trends that is linearly related to simulated fluctuations in Niño 3.4 temperatures ("ENSO component"; red). All computations involved monthly mean anomaly data for overlapping 228-month segments of the control run; overlap was by 216 months (see section 5.4). Owing to the large overlap, the number of independent samples is much smaller than the number of actual samples. The dotted vertical lines are the 2LT trend and the ENSO-induced temperature component in the 19-year GSOP experiment. GSOP results are for $\tau = 30$ months, $t_{base} = 6$ months, and Niño 3.4–based ENSO removal.

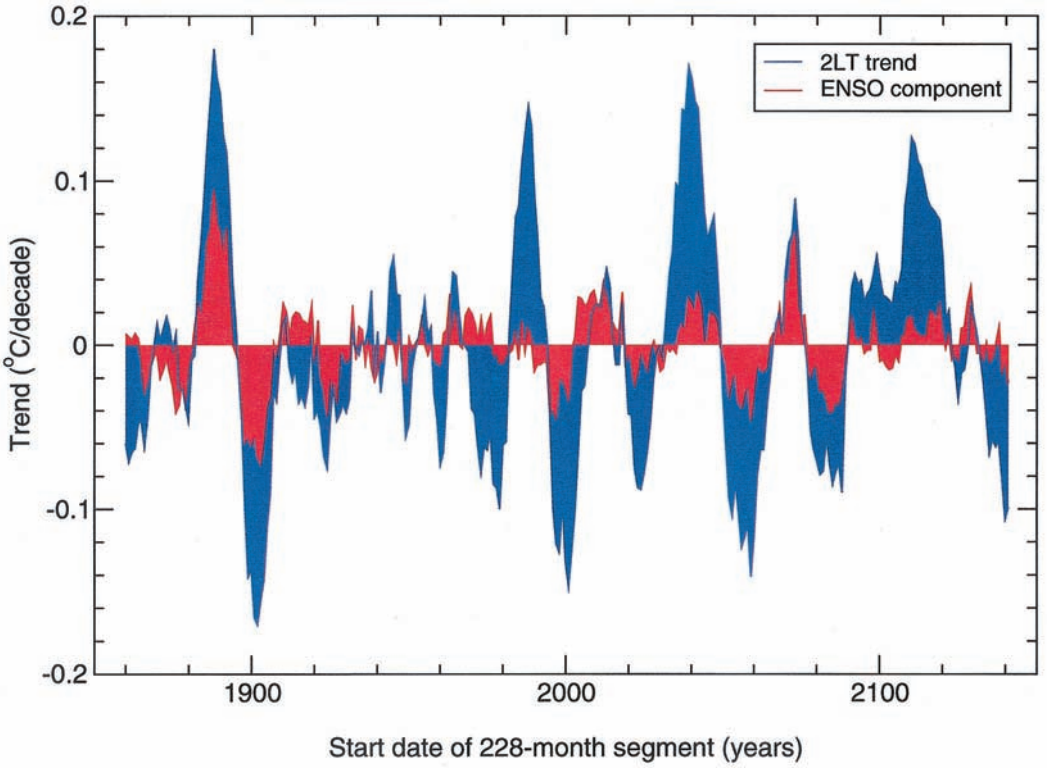

Plate 14. Time evolution of ENSO component of lower tropospheric temperatures in ECHAM4/OPYC control run. See Plate 13 for data sources. The 228-month segments overlap by 216 months and are plotted on the start date of the segment. The nominal start date is 1860.

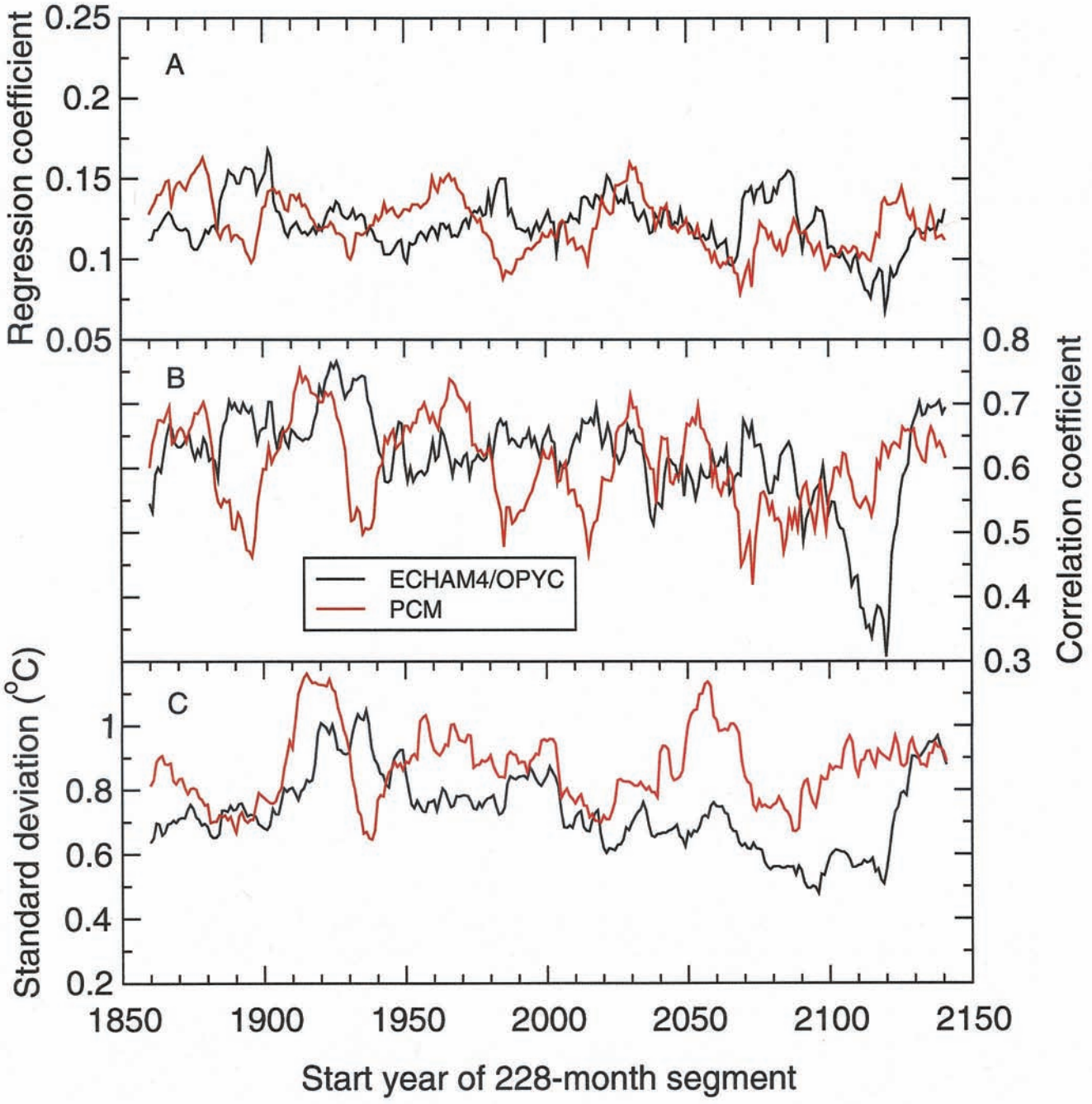

Plate 15. Regression and correlation coefficients (b_j and r_j) between simulated Niño 3.4 SSTs and lower tropospheric temperatures in the ECHAM4/OPYC and PCM control runs. Values of (a) b_j and (b) r_j are computed for successive 228-month segments of the model data, with an overlap of 216 months between segments. For each segment, j is the lag (in months) that maximizes b_j and r_j, with $j \leq 24$. (c) The 228-month running standard deviation of simulated Niño 3.4 SSTs.

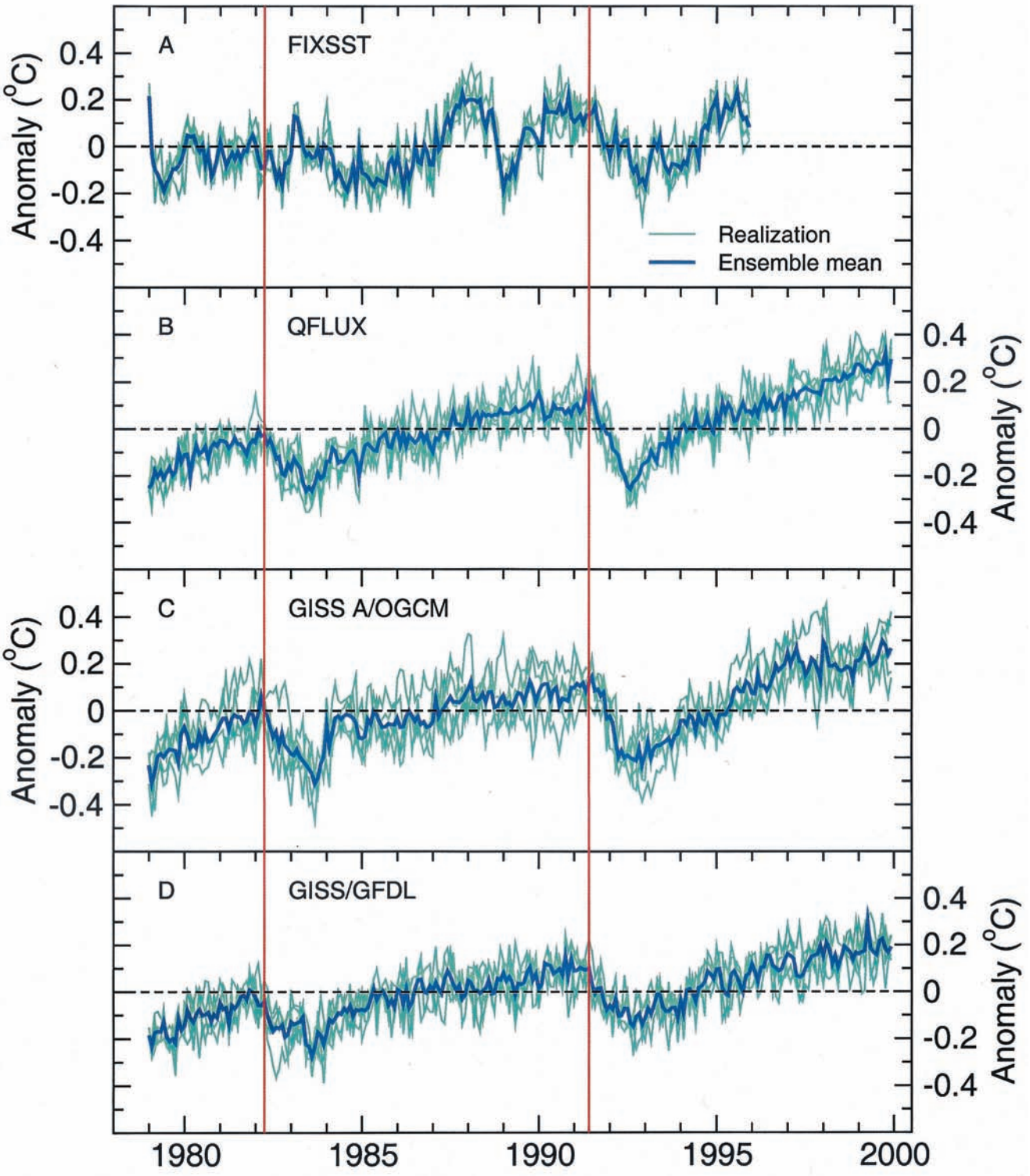

Plate 16. Near-surface temperature changes in the *Hansen et al.* [1997a] perturbation experiments with combined natural (solar, volcanic aerosols) and anthropogenic (well-mixed greenhouse gases, stratospheric and tropospheric ozone) forcings. Experiments were performed with four different experimental configurations: (a) prescribed SST experiments with the GISS AGCM, (b) fully coupled simulations with the GISS AGCM and a slab ocean, (c) the GISS A/OGCM, and (d) the GISS AGCM coupled to the GFDL OGCM. Five independent realizations of the "combined forcing" experiment were performed with each of the four model configurations. Individual realizations (in light blue) are not identified. The dark blue line is the ensemble mean result.

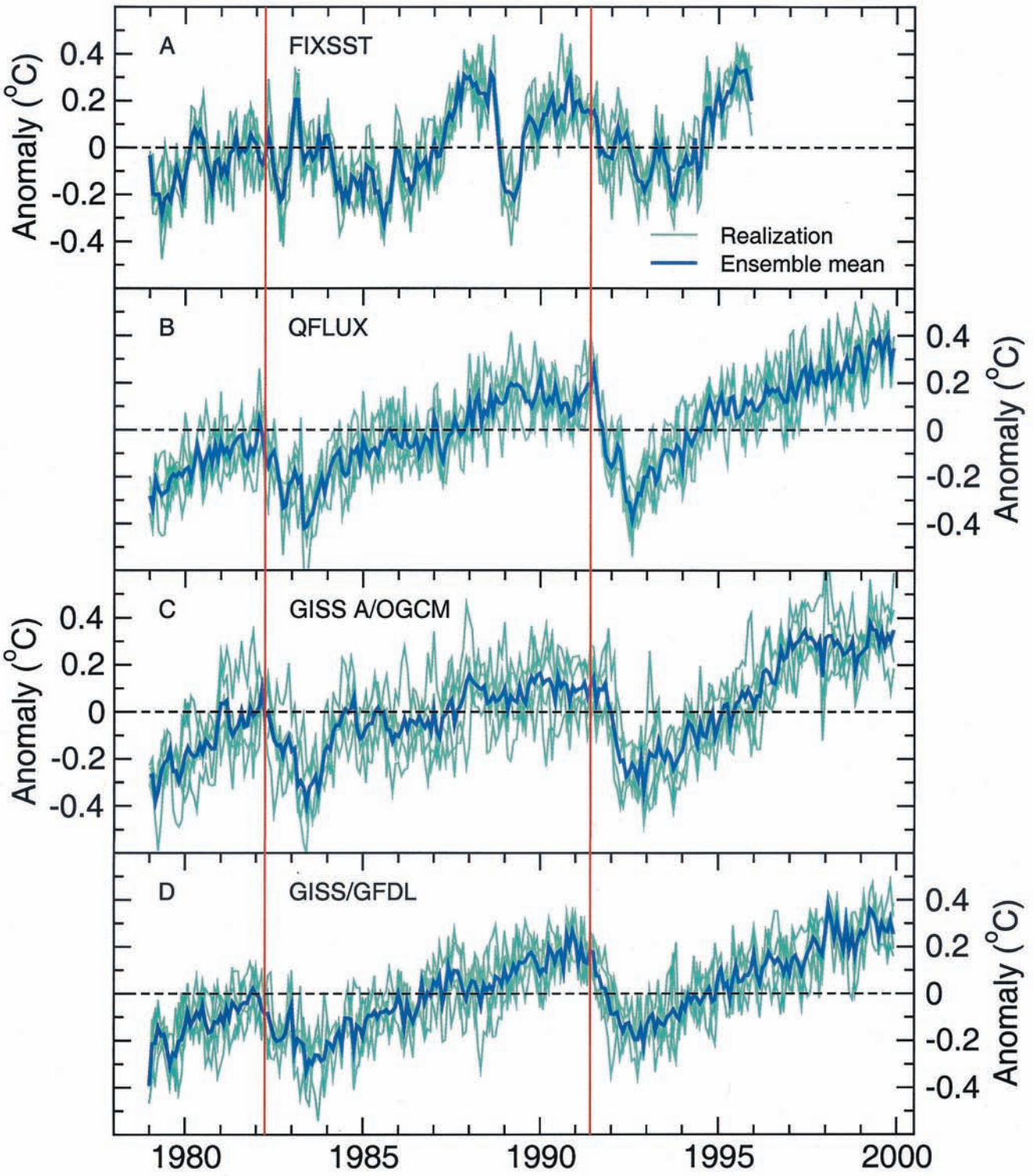

Plate 17. As for Plate 16, but for simulated lower tropospheric temperature (2LT).

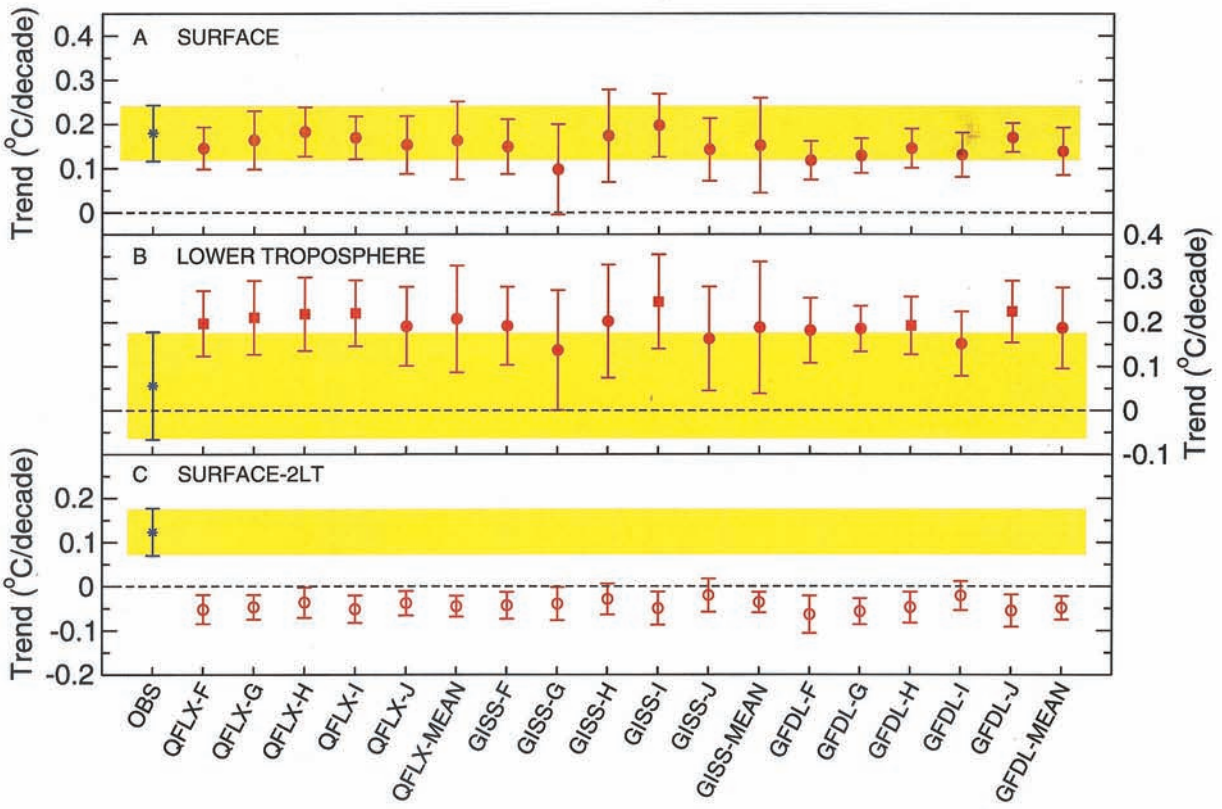

Plate 18a. Raw observed (blue) and simulated (red) trends and 2σ confidence intervals for surface and lower tropospheric temperature data (panels a and b), and for the surface-2LT trend difference (panel c). Observed and simulated trends are computed over the 252-month period January 1979 through December 1999, without subtraction of ENSO or volcanic signals. For sources of observed data, see section 2.1. Model data are from the *Hansen et al.* [1997a] experiments with combined natural and anthropogenic forcing (see Plate 16). These involve three different experimental configurations: the GISS AGCM coupled to a slab ocean ("QFLX"), the GISS A/OGCM ("GISS"), and the GISS AGCM coupled to the GFDL OGCM ("GFDL"). The letters F-J denote individual realizations of the experiment; "MEAN" is the ensemble average. Confidence intervals are adjusted for temporal autocorrelation effects [*Santer et al.*, 2000b]. The yellow shading marks the observed 2σ confidence intervals. For interpretation of significance testing results (i.e., open or solid symbols, symbol shape), refer to Plate 10.

as 25 to 30% within the space of several months. For Niño 3.4/2LT regressions the temporal variability of b_j in Plate 15a is only weakly correlated with running 228-month standard deviations of Niño 3.4 SSTs in Plate 15c (correlation coefficients are −0.008 and −0.001 for ECHAM and PCM, respectively; the corresponding results for surface data are −0.156 and 0.220). In contrast, fluctuations in r_j are more strongly correlated with variability in Niño 3.4 SSTs (0.657 and 0.634 for ECHAM and PCM 2LT data and 0.385 and 0.729 for near-surface data). On timescales significantly shorter than 228 months (for example, on the 60-month timescale that WS used to discriminate between different volcanic cooling signals, or the 66-month "volcano-free" period that *Christy and McNider* [1994] used to compute the ENSO regression coefficient), the temporal variations in b_j and r_j are much larger than in Plate 15.

If the models' portrayal of the unforced variability in b_j is realistic, the rapid and large fluctuations in b_j found here make the unravelling of ENSO and volcanic signals a difficult task. Nonlinear techniques may ultimately provide a better means of separating these signals. Such techniques, however, can introduce further uncertain parameters (see section 4.1) and will not necessarily reduce overall uncertainties in estimated volcano and ENSO signals.

6. Analyses of GISS Model Data

In sections 4 and 5 we attempted to remove ENSO and volcanic signals from observations and the ECHAM4/OPYC atmospheric temperature data. As we have shown above, these signal estimates are subject to considerable uncertainty. It is therefore useful to complement the "residual" trend comparisons in section 5.3 with comparisons of "raw" modeled and observed temperature trends. Ideally, comparisons of raw trends should be based on model experiments with relatively complete volcanic forcing and with a large enough ensemble to estimate trend uncertainties arising from natural internal climate variability.

The recent experiments by *Hansen et al.* [1997a] fulfill these requirements. They include the effects of both El Chichón and Pinatubo, and also incorporate changes in other natural and anthropogenic forcings (see section 2.2). In the following, we

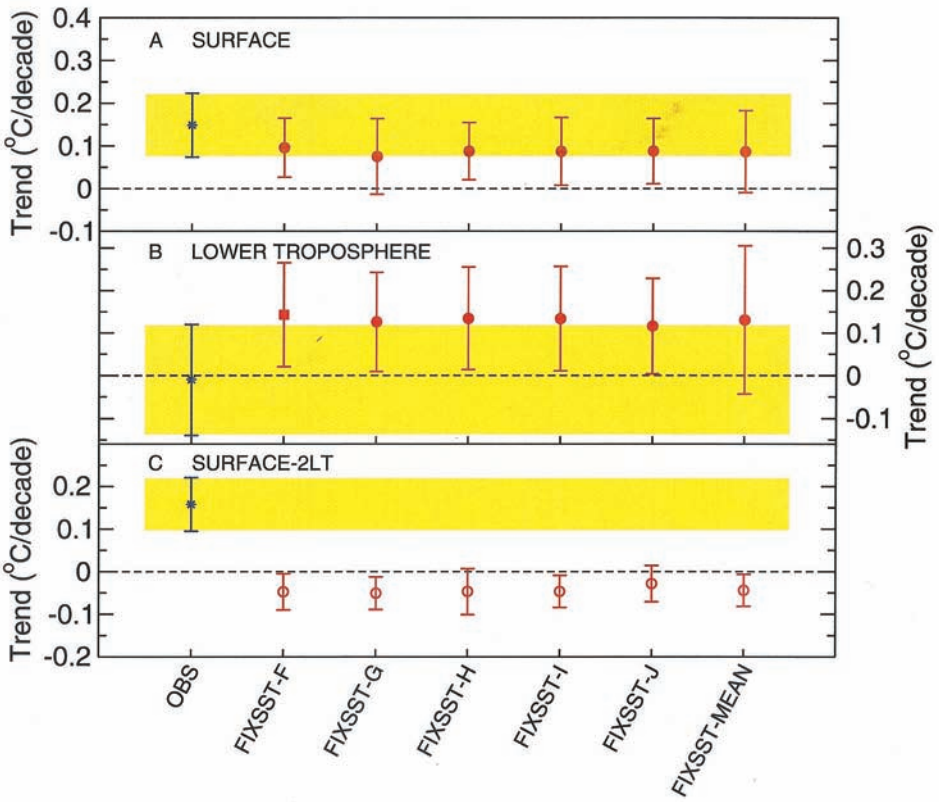

Plate 18b. As for Plate 18a, but for model results from the *Hansen et al.* [1997a] AGCM experiments with prescribed time-varying SSTs. The analysis period for both model results and observations is January 1979 to December 1995 (204 months).

compare the raw surface and tropospheric trends from these experiments with raw observed trends.

We do not examine the simulated climate response in the lower stratosphere. *Hansen et al.* [1997a] have noted that all four GISS model configurations underestimate the observed lower stratospheric warming that followed the El Chichón and Pinatubo eruptions. This is likely related to poor representation of the lower stratosphere, a consequence of the coarse vertical resolution of the SI95 AGCM.

The behavior of surface and lower tropospheric temperatures in the *Hansen et al.* [1997a] experiments is shown in Plates 16 and 17 (respectively). Recall that there are three different coupled model configurations, and one set of AGCM experiments with prescribed time-varying observed SSTs. In the latter, as in the real world, the response to El Chichón is obscured by the effects of the 1982/1983 El Niño. The temperature signatures of El Chichón and Pinatubo are more clearly evident in the coupled model experiments.

Details of the volcanic responses are model-dependent. In the GISS AGCM/slab ocean configuration the surface and 2LT responses to Pinatubo are larger and decay more rapidly (i.e., $\Delta T_{\max}$ values are larger, and τ values are smaller) than in the two A/OGCMs. The variability between ensemble members simulated by the GISS A/OGCM is noticeably larger than in the other two models, which leads to greater intraensemble variability in $\Delta T_{\max}$ and t_{ramp} estimates.

Observed and simulated "raw" least squares linear trends are shown in Plate 18a for the three sets of coupled model runs, and in Plate 18b for the fixed SST integrations. Trends are computed over different time periods (1979–1999 in Plate 18a and 1979–1995 in Plate 18b) since the fixed SST integrations are shorter than the coupled model experiments (section 2.2).

None of the 24 surface trends in the GISS experiments are significantly different from observations (at the 5% level), although 22 of 24 are smaller than observed values. Even the fixed SST experiments have smaller surface temperature trends than observed (Plate 18b). This may be partly related to coverage differences between the model and observations [*Santer et al.*, 2000a], which are not accounted for here.

Modeled lower tropospheric temperature trends are always larger than in observations, but these differences are statistically significant in only 7 out of 24 cases. Four of the seven results are for the GISS AGCM/slab ocean configuration (Plate 18a), possibly due to the fact it has smaller thermal inertia than the two A/OGCMs. The larger than observed 2LT trends in the fixed SST integrations can have various interpretations: model error, missing or inaccurately specified forcing, and/or errors in the observational data.

All 24 of the raw trends in the modeled surface-minus-2LT difference series are negative (larger warming aloft than at the surface), and significantly different from the (positive) observed results. Qualitatively and quantitatively different results are obtained if surface trends are computed from SAT over land and oceans, rather than from merged SST and SAT data. Use of SATs yields larger surface warming in all three coupled model configurations.

We do not report here on the application of our iterative volcano/ENSO removal method to the GISS data. We note, however, that some of the model configurations considered

here have reduced ENSO variability and unrealistic lags between Niño 3 or 3.4 SSTs and atmospheric temperature. These problems hamper application of our iterative approach.

7. Conclusions

We have shown that separating volcanic and ENSO signals in globally averaged observational temperature data is a difficult task. Difficulties arise primarily from collinearity between these signals over the 21-year period of the satellite temperature record. The two largest volcanic eruptions during this time (El Chichón in April 1982 and Pinatubo in June 1991) coincided with strong and weak El Niño events (see Plate 1). This collinearity complicates the estimation of poorly known volcanic parameters, such as τ, the volcanic signal decay time, T_{ref}, the preeruption reference level temperature, $\Delta T_{\max}$, the maximum cooling induced by each volcanic eruption, and t_{ramp}, the time between the start of the eruption and $\Delta T_{\max}$. Collinearity also makes it more difficult to quantify the effect of ENSO on atmospheric temperatures.

Most previous attempts to disentangle volcanic and ENSO signals in observational temperature data have either incorrectly dismissed the possibility of collinearity effects [*Christy and McNider*, 1994; *Michaels and Knappenberger*, 2000] or concluded that these are relatively unimportant [e.g., *Angell*, 2000]. A recent paper by Wigley and Santer (submitted manuscript, 2001 (WS)) showed that uncertainties inherent in separating volcano and ENSO signals are larger than hitherto assumed.

We confirmed this finding using a modified version of the iterative volcano/ENSO signal separation method used by WS. The major modification is that the volcano parameters T_{ref}, $\Delta T_{\max}$, and t_{ramp} are estimated with an automated procedure. This has the advantage of yielding reproducible results, but the disadvantage that a number of subjective decisions must be made in order to implement the approach.

Over the period 1979–1999 our estimates of the observed "residual" global mean temperature trends after removal of volcanic and ENSO effects range from 0.056° to 0.158°C/decade in the lower troposphere and from 0.210° to 0.250°C/decade at the surface. These are almost always larger than trends in the raw 2LT and surface data over the same period (0.056° and 0.180°C/decade, respectively; Plate 3). This suggests that the net effect of volcanoes and ENSO over the past 21 years has been to reduce underlying atmospheric warming trends, as noted by WS. The volcanically induced trend reduction is greater for larger values of τ (Plate 4).

The difference between global mean surface and lower tropospheric temperature trends ranges from 0.082° to 0.151°C/decade after removal of volcano and ENSO influences. Most of these values are smaller than the surface-minus-2LT trend difference in the "raw" observational data (0.124°C/decade). Our observational results and those of WS indicate that volcanic effects probably cooled the lower troposphere temperature by more than the surface. ENSO influences over 1979–1999 had the opposite effect, but were much smaller than volcanic influences during this specific period. The net result when both ENSO and volcanoes are removed is a reduction in the surface-minus-2LT trend difference by as much as a half. These effects alone, however, cannot fully explain why the surface warmed relative to the lower troposphere over 1979–1999. A significant difference still exists between observed residual trends at the Earth's surface and in the lower troposphere.

The key parameters that dictate the wide ranges for residual surface, 2LT, and surface-minus-2LT trends are the choice of ENSO index and uncertainties in the volcanic parameters τ and T_{ref}. We also find evidence to suggest that purely linear signal separation approaches (such as that used here) may be suboptimal. Our iterative approach implicitly assumes that there is a linear, stable regression relationship between fluctuations in a selected ENSO index and the variability of global mean atmospheric temperature. Analyses of temperature data from long control runs with no volcanic effects indicate that this assumption may be invalid (Plate 15). If so, uncertainties in the separation of volcanic and ENSO signals may be even larger than estimated here. A further complication is the possible existence of nonlinear interaction effects between volcanic cooling and ENSO indices.

There is a clear need to explore this signal separation problem with more sophisticated statistical approaches, such as nonlinear iterative schemes, maximum likelihood methods (Smith et al., submitted manuscript, 2001), and independent components. We have investigated the usefulness of these techniques and find that they do not automatically lead to tighter constraints on estimates of volcano and ENSO atmospheric temperature signals.

Another motivation for addressing this signal separation problem was to facilitate comparisons of modeled and observed trends on relatively short timescales (2–3 decades). Such comparisons are hampered by differences in the phasing of observed and simulated realizations of ENSO. Additionally, at least one recent model experiment (the so-called "GSOP" integration performed with the ECHAM4/OPYC model of the Max-Planck Institute for Meteorology) incorporated the effects of the Pinatubo eruption, but not of El Chichón [*Bengtsson et al.*, 1999]. Both ENSO effects and incomplete volcanic forcing can hamper reliable assessment of the true correspondence between modeled and observed temperature trends.

We used our iterative method to factor out volcanic and ENSO influences from the GSOP experiment, and then compared the simulated "residual" temperature trends with those estimated from observations. All comparisons were made over 1979–1997, the period covered by the ECHAM climate change experiments. As in the observations, there is collinearity between the volcano and ENSO signals simulated in GSOP. Coupled with uncertainties in volcano parameters and in the removal of ENSO effects, this leads to wide ranges in estimates of simulated residual trends.

We also analyzed residual temperature trends in the ECHAM GSO1 and GSO2 experiments. These do not include volcanic forcing, and only ENSO effects were removed from them.

At the surface none of the 54 pairs of simulated and observed residual temperature trends are significantly different at the 5% level. In the lower troposphere, model residual trends are larger than observed in all 54 cases and significantly different from observations in 45 of these cases. The statistical significance of residual trend differences depends on the experiment considered, on τ, and on the index used for removing ENSO effects (Plates 10–12). Simulated surface-minus-2LT trend differences are significantly smaller than observed results in 51 out of 54 cases. In GSOP the surface-minus-2LT trend difference is positive in both the raw data and (in 10 out of 18

cases) in the residual temperature data. This behavior is qualitatively similar to the observations (Plate 10).

In virtually all cases, removing the effects of volcanoes and ENSO improves the correspondence between observed temperature trends over 1979–1997 and trends simulated in the ECHAM climate change experiments. It cannot fully reconcile the observed differential warming rates in the lower troposphere and at the surface with those simulated in GSOP, GSO1, and GSO2.

Given the uncertainties inherent in the separation and removal of collinear ENSO and volcanic signals, it is desirable to augment comparisons of "residual" modeled and observed temperature trends with comparisons of "raw" trends. This requires access to model experiments that include relatively complete estimates of key natural and anthropogenic forcings. Such experiments were recently performed by *Hansen et al.* [1997a] with four different configurations of the Goddard Institute for Space Studies (GISS) coupled model. Unlike the ECHAM4/OPYC GSOP experiment, these integrations include stratospheric aerosols from both the El Chichón and Pinatubo eruptions.

Our analysis of "raw" surface, 2LT, and surface-minus-2LT trends in observations and the *Hansen et al.* [1997a] coupled experiments yields results very similar to those obtained for the ECHAM "residual" trend comparisons. None of the 24 GISS surface temperature trends over 1979–1999 are significantly different from observations. In the lower troposphere, statistically significant differences between simulated and observed trends occur in only 7 of 24 comparisons. Differences between modeled and observed trends in the surface-minus-2LT time series are highly significant in all cases.

The GISS and ECHAM results raise two questions. The first relates to how well models reproduce observed 2LT trends over the satellite era. The second question is whether models can successfully simulate the observed differential warming between the surface and lower troposphere.

Consider the issue of 2LT trends first. The GISS results, with statistically significant model-versus-observed differences in only 7 out of 24 cases, contradict recent claims [*Singer*, 1999; *Michaels and Knappenberger*, 2000] that there is a fundamental discrepancy between observed lower tropospheric temperature trends and those simulated in all climate model experiments. Comparisons of residual trends in ECHAM climate change experiments and observations are more equivocal: most (45 of 54) of the simulated 2LT trends are significantly different from observations. Unlike the GISS case, this large range is dictated by statistical uncertainties in the removal of ENSO and volcano effects, not by the availability of large ensembles of experiments.

In some cases we understand why these differences exist (section 5.3.2). In other cases we do not. Given the small number (three) of ECHAM climate change experiments available, the existence of large noise contributions from ENSO and other modes of natural variability (Plate 14), and our finding that significance decisions depend on the experiment considered, it would be important to perform multiple realizations of the GSOP experiment. Preferably, these experiments should include the forcing from El Chichón. Such ensembles would help to evaluate the true correspondence (or lack thereof) between observed 2LT trends and those simulated in ECHAM.

Now consider the issue of differential warming trends at the surface and aloft. In all but one (GSOP) of the ECHAM and GISS experiments, the recent observed warming of the surface relative to the lower troposphere is not reproduced: the raw simulations show the converse behavior. Accounting for model/data differences in volcano and ENSO effects explains some of this discrepancy, but statistically significant differences remain between simulated and observed differential warming rates.

There is a distinction, however, between formal statistical significance and physical significance. Statistical comparisons of the simulated and observed trend in the surface-minus-2LT difference series constitute a stringent test of model performance. The subtraction of variability components common to the surface and lower troposphere markedly reduces the standard errors for surface-minus-2LT trends (compare panel c and panels a and b in Plates 18a and 18b). As a result, even relatively small errors, in the model responses, the model forcings, and/or in the observations themselves, can lead to highly significant differences between modeled and observed surface-minus-2LT trends. Overestimation (underestimation) of the true surface (2LT) temperature by as little as 0.05°C/decade (a value that is within the currently estimated accuracy for the MSU 2LT data [*Christy et al.*, 1998]) would in most cases lead to nonsignificant differences between the "residual" observed surface-minus-2LT trend differentials and those simulated in the ECHAM climate change experiments (Plates 10–12). Coverage differences, which were not considered here, may also explain part of the discrepancy between models and observations (see section 2.2).

In summary, we have shown that there are significant uncertainties involved in separating volcanic and ENSO signals, both in observational and in model data. The collinearity between these signals, combined with the relatively short length of the satellite record, may make it difficult to significantly reduce these uncertainties. Confident assertions that these effects are easily separable [*Michaels and Knappenberger*, 2000] are not supported by our analysis. Finally, the failure of the MSU 2LT record to show significant warming has often been highlighted as the major remaining inconsistency between theory and data [*Singer*, 1999]. As shown here and by Wigley and Santer (submitted manuscript, 2001), accounting for the influence of recent volcanic eruptions and ENSO can leave large positive residual trends in the MSU 2LT data. Had Pinatubo and El Chichón not occurred, it is likely that the lower troposphere would have experienced more pronounced warming.

Appendix A: Differences From Previous Comparison of Modeled and Observed 2LT Trends

Our comparison of observed and simulated 2LT trends differs in three ways from previous work by *Santer et al.* [2000a]. First, Santer et al. considered "raw" simulated and observed 2LT trends only. They did not explicitly account for model/observed differences in ENSO and volcanic effects, as was done here.

Second, *Santer et al.* [2000a] calculated raw trends and performed significance testing after subsampling simulated (GSOP, GSO1, and GSO2) and observed (MSU) lower tropospheric temperature data with the incomplete coverage of the IPCC near-surface temperature data. This facilitated the interpretation of temperature trend differences between the surface and lower troposphere, but was unnecessary for comparison of modeled and observed 2LT trends (since both 2LT data

sets have complete global coverage over the satellite era). No subsampling of 2LT data was performed here.

This explains why the raw GSOP 2LT and surface trends in Plate 10 (0.164° and 0.171°C/decade, respectively) are larger than the raw trends in the subsampled GSOP data (0.149° and 0.114°C/decade [*Santer et al*., 2000a, Table 3]). As is evident in the work of *Bengtsson et al*. [1999, Plate 3], the warming at high latitudes in GSOP is largely outside the area of observed data coverage, particularly at the surface, leading to smaller 2LT and surface trends in the subsampled data. Coverage differences also explain why the surface-minus-2LT trend difference in the "raw" GSOP data is slightly positive here (0.006°C/decade; Plate 10) but negative in the subsampled GSOP data (−0.035°C/decade).

Third, significance testing by *Santer et al*. [2000a] involved observed and simulated 2LT trends computed over different periods: 1979–1997 in the case of GSOP, GSO1, and GSO2 (the duration of the experiments), and 1979–1998 for the MSU data (the length of observational record available at the time of the study). Here modeled and observed trends were computed over the same 19-year period (1979–1997).

Acknowledgments. Work at Lawrence Livermore National Laboratory was performed under the auspices of the U.S. Department of Energy, Environmental Sciences Division, under contract W-7405-ENG-48. Tom Wigley received support from the NOAA Office of Global Programs ("Climate Change Data and Detection") under grant NA87GP0105. Phil Jones was supported by the U.S. Department of Energy under grant DE-FG02-98ER62601. The authors would like to acknowledge the assistance of Reto Ruedy (Goddard Institute for Space Studies), Monika Esch (Max-Planck Institute for Meteorology), and Krishna AchutaRao and Michael Wehner (both PCMDI) in the transfer and processing of model data. Results from the ECHAM4/OPYC simulations were kindly provided by Lennart Bengtsson (MPI). Comments from Jerry Mahlman (Geophysical Fluid Dynamics Laboratory), Simon Tett (Hadley Centre), Kevin Trenberth (NCAR), and two anonymous reviewers helped to improve the paper.

References

Andronova, N. G., E. V. Rozanov, F. Yang, and M. E. Schlesinger, Radiative forcing by volcanic aerosols from 1850 to 1994, *J. Geophys. Res*., *104*, 16,807–16,826, 1999.

Angell, J. K., Impact of El Niño on the delineation of tropospheric cooling due to volcanic eruptions, *J. Geophys. Res*., *93*, 3697–3704, 1988.

Angell, J. K., Tropospheric temperature variations adjusted for El Niño, 1958–1998, *J. Geophys. Res*., *105*, 11,841–11,849, 2000.

Bengtsson, L., E. Roeckner, and M. Stendel, Why is the global warming proceeding much slower than expected?, *J. Geophys. Res*., *104*, 3865–3876, 1999.

Boville, B. A., and J. W. Hurrell, A comparison of the atmospheric circulations simulated by CCM3 and CSM1, *J. Clim*., *11*, 1327–1341, 1998.

Bryan, K., and M. D. Cox, An approximate equation of state for numerical model of ocean circulation, *J. Phys. Oceanogr*., *15*, 1255–1273, 1972.

Christy, J. R., and R. T. McNider, Satellite greenhouse signal, *Nature*, *367*, 325, 1994.

Christy, J. R., R. W. Spencer, and E. S. Lobl, Analysis of the merging procedure for the MSU daily temperature time series, *J. Clim*., *11*, 2016–2041, 1998.

Christy, J. R., R. W. Spencer, and W. D. Braswell, MSU tropospheric temperatures: Dataset construction and radiosonde comparisons, *J. Atmos. Oceanic Technol*., *17*(9), 1153–1170, 2000.

Gaffen, D. J., B. D. Santer, J. S. Boyle, J. R. Christy, N. E. Graham, and R. J. Ross, Multidecadal changes in the vertical structure of the tropical troposphere, *Science*, *287*, 1242–1245, 2000.

Gates, W. L., et al., An overview of the Atmospheric Model Intercomparison Project (AMIP), *Bull. Am. Meteorol. Soc*., *80*, 29–55, 1999.

Graf, H.-F., I. Kirchner, A. Robock, and I. Schult, Pinatubo eruption winter climate effects: Model versus observations, *Clim. Dyn*., *9*, 81–93, 1993.

Hansen, J. E., A. Lacis, R. Ruedy, and M. Sato, Potential climate impact of Mount Pinatubo eruption, *Geophys. Res. Lett*., *19*, 215–218, 1992.

Hansen, J. E., A. Lacis, R. Ruedy, M. Sato, and H. Wilson, How sensitive is the world's climate?, *Natl. Geogr. Res. Explor*., *9*, 142–158, 1993.

Hansen, J. E., et al., Forcing and chaos in interannual to decadal climate change, *J. Geophys. Res*., *102*, 25,679–25,720, 1997a.

Hansen, J. E., M. Sato, A. Lacis, and R. Ruedy, The missing climate forcing, *Philos. Trans. R. Soc. London*, *Ser. B*, *352*, 231–240, 1997b.

Hansen, J. E., M. Sato, R. Ruedy, A. Lacis, and J. Glascoe, Global climate data and models: A reconciliation, *Science*, *281*, 930–932, 1998.

Hegerl, G. C., K. Hasselmann, U. Cubasch, J. F. B. Mitchell, E. Roeckner, R. Voss, and J. Waszkewitz, Multi-fingerprint detection and attribution analysis of greenhouse gas, greenhouse gas-plus-aerosol and solar forced climate change, *Clim. Dyn*., *13*, 613–634, 1997.

Hurrell, J. W., and K. E. Trenberth, Difficulties in obtaining reliable temperature trends: Reconciling the surface and Microwave Sounding Unit records, *J. Clim*., *11*, 945–967, 1998.

Jones, P. D., Recent warming in global temperature series, *Geophys. Res. Lett*., *21*, 1149–1152, 1994a.

Jones, P. D., Hemispheric surface air temperature variations: A reanalysis and an update to 1993, *J. Clim*., *7*, 1794–1802, 1994b.

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

Kattenberg, A., F. Giorgi, H. Grassl, G. A. Meehl, J. F. B. Mitchell, R. J. Stouffer, T. Tokioka, A. J. Weaver, and T. M. L. Wigley, Climate models—Projections of future climate, in *Climate Change 1995: The Science of Climate Change*, edited by J. T. Houghton et al., pp. 285–357, Cambridge Univ. Press, New York, 1996.

Kirchner, I., G. L. Stenchikov, H.-F. Graf, A. Robock, and J. C. Antuña, Climate model simulation of winter warming and summer cooling following the 1991 Mount Pinatubo volcanic eruption, *J. Geophys. Res*., *104*, 19,039–19,055, 1999.

Knutson, T. R., S. Manabe, and D. Gu, Simulated ENSO in a global coupled ocean-atmosphere model: Multidecadal amplitude modulation and CO_2 sensitivity, *J. Clim*., *10*, 138–161, 1997.

Knutson, T. R., T. L. Delworth, K. W. Dixon, and R. J. Stouffer, Model assessment of regional surface temperature trends (1949–1997), *J. Geophys. Res*., *104*, 30,981–30,996, 1999.

Können, G. P., P. D. Jones, M. H. Kaltofen, and R. J. Allan, Pre-1866 extensions of the Southern Oscillation Index using early Indonesian and Tahitian meteorological records, *J. Clim*., *11*, 2325–2339, 1998.

Meehl, G. A., P. Gent, J. M. Arblaster, B. Otto-Bliesner, E. Brady, and A. Craig, Factors that affect amplitude of El Niño in global coupled climate models, *Clim. Dyn*., *17*, 515–526, 2000.

Michaels, P. J., and P. C. Knappenberger, Natural signals in the MSU lower tropospheric temperature record, *Geophys. Res. Lett*., *27*, 2905–2908, 2000.

Miller, R. L., and X. Jiang, Surface energy fluxes and coupled variability in the tropics of a coupled general circulation model, *J. Clim*., *9*, 1599–1620, 1996.

Mosteller, F., and J. W. Tukey, *Data Analysis and Regression*, 588 pp., Addison-Wesley-Longman, Reading, Mass., 1977.

National Research Council, Board on Atmospheric Sciences and Climate, Commission on Geosciences, Environment, and Resources, *Reconciling Observations of Global Temperature Change*, 85 pp., Natl. Acad. Press, Washington, D. C., 2000.

Oberhuber, J. M., Simulation of the Atlantic circulation with a coupled sea-ice mixed-layer isopycnal general circulation model, I, Model description, *J. Phys. Oceanogr*., *22*, 808–829, 1993.

Parker, D. E., P. D. Jones, C. K. Folland, and A. Bevan, Interdecadal changes of surface temperature since the late nineteenth century, *J. Geophys. Res*., *99*, 14,373–14,399, 1994.

Portman, R. W., S. Solomon, R. R. Garcia, L. W. Thomason, L. R. Poole, and M. P. McCormick, Role of aerosol variations in anthropogenic ozone depletion in the polar regions, *J. Geophys. Res*., *101*, 22,991–23,006, 1996.

Robock, A., Volcanic eruptions and climate, *Rev. Geophys*., *38*, 191–219, 2000.

Robock, A., and J. Mao, The volcanic signal in surface temperature observations, *J. Clim.*, *8*, 1086–1103, 1995.

Roeckner, E., J. M. Oberhuber, A. Bacher, M. Christoph, and I. Kirchner, ENSO variability and atmospheric response in a global coupled atmosphere-ocean GCM, *Clim. Dyn.*, *12*, 734–752, 1996.

Roeckner, E., L. Bengtsson, J. Feichter, J. Lelieveld, and H. Rodhe, Transient climate change simulations with a coupled atmosphere-ocean GCM including the tropospheric sulfur cycle, *J. Clim.*, *12*, 3004–3032, 1999.

Ropelewski, C. F., and P. D. Jones, An extension of the Tahiti-Darwin Southern Oscillation Index, *Mon. Weather Rev.*, *115*, 2161–2165, 1987.

Russell, G. L., J. R. Miller, and D. Rind, A coupled atmosphere-ocean model for transient climate-change studies, *Atmos. Oceans*, *33*, 683–730, 1995.

Santer, B. D., et al., A search for human influences on the thermal structure of the atmosphere, *Nature*, *382*, 39–46, 1996.

Santer, B. D., J. J. Hnilo, T. M. L. Wigley, J. S. Boyle, C. Doutriaux, M. Fiorino, D. E. Parker, and K. E. Taylor, Uncertainties in observationally based estimates of temperature change in the free atmosphere, *J. Geophys. Res.*, *104*, 6305–6333, 1999.

Santer, B. D., et al., Interpreting differential temperature trends at the surface and in the lower troposphere, *Science*, *287*, 1227–1232, 2000a.

Santer, B. D., T. M. L. Wigley, J. S. Boyle, D. J. Gaffen, J. J. Hnilo, D. Nychka, D. E. Parker, and K. E. Taylor, Statistical significance of trends and trend differences in layer-average atmospheric temperature time series, *J. Geophys. Res.*, *105*, 7337–7356, 2000b.

Sato, M., J. E. Hansen, M. P. McCormick, and J. B. Pollack, Stratospheric aerosol optical depths, 1850–1990, *J. Geophys. Res.*, *98*, 22,987–22,994, 1993.

Singer, S. F., Human contribution to climate change remains questionable, *Eos Trans. AGU*, *80*, 183–187, 1999.

Solomon, S., Stratospheric ozone depletion: A review of concepts and history, *Rev. Geophys.*, *37*, 275–316, 1999.

Solomon, S., R. W. Portman, R. R. Garcia, L. W. Thomason, L. R. Poole, and M. P. McCormick, The role of aerosol variations in anthropogenic ozone depletion at northern midlatitudes, *J. Geophys. Res.*, *101*, 6713–6727, 1996.

Spencer, R. W., and J. R. Christy, Precision and radiosonde validation of satellite grid-point temperature anomalies, part II, A tropospheric retrieval and trends during 1979–90, *J. Clim.*, *5*, 858–866, 1992.

Stott, P. A., S. F. B. Tett, G. S. Jones, M. R. Allen, J. F. B. Mitchell, and G. J. Jenkins, External control of 20th century temperature by natural and anthropogenic forcings, *Science*, *290*, 2133–2137, 2000.

Tett, S. F. B., J. F. B. Mitchell, D. E. Parker, and M. R. Allen, Human influence on the atmospheric vertical temperature structure: Detection and observations, *Science*, *274*, 1170–1173, 1996.

Tett, S. F. B., P. A. Stott, M. R. Allen, W. J. Ingram, and J. F. B. Mitchell, Causes of twentieth-century temperature change near the Earth's surface, *Nature*, *399*, 569–572, 1999.

Thompson, D. W. J., and J. M. Wallace, The Arctic Oscillation signature in the wintertime geopotential height and temperature fields, *Geophys. Res. Lett.*, *25*, 1297–1300, 1998.

Timmermann, A., J. Oberhuber, A. Bacher, M. Esch, M. Latif, and E. Roeckner, Increased El Niño frequency in a climate model forced by future greenhouse warming, *Nature*, *398*, 694–696, 1999.

Washington, W. M., et al., Parallel Climate Model (PCM) control and transient simulations, *Clim. Dyn.*, *16*, 755–774, 2000.

Wentz, F. J., and M. Schabel, Effects of orbital decay on satellite-derived lower tropospheric temperature trends, *Nature*, *394*, 661–664, 1998.

Wentz, F. J., and M. Schabel, Precise climate monitoring using complementary satellite data sets, *Nature*, *403*, 414–416, 2000.

Wigley, T. M. L., ENSO, volcanoes and record-breaking temperatures, *Geophys. Res. Lett.*, *27*, 4101–4104, 2000.

Wigley, T. M. L., and S. C. B. Raper, Implications for climate and sea level of revised IPCC emissions scenarios, *Nature*, *357*, 293–300, 1992.

Wigley, T. M. L., and M. E. Schlesinger, Analytical solution for the effect of increasing CO_2 on global mean temperature, *Nature*, *315*, 649–652, 1985.

Wigley, T. M. L., R. L. Smith, and B. D. Santer, Anthropogenic influence on the autocorrelation structure of hemispheric-mean temperatures, *Science*, *282*, 1676–1679, 1998.

J. S. Boyle, C. Doutriaux, B. D. Santer (corresponding author), and K. E. Taylor, Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, P.O. Box 808, Mail Stop L-264, Livermore, CA 94550, USA. (santer1@llnl.gov)

J. E. Hansen, NASA Goddard Institute for Space Studies, New York, NY 10025, USA.

P. D. Jones, Climatic Research Unit, University of East Anglia, Norwich NR4 7TJ, England, UK.

G. A. Meehl and T. M. L. Wigley, National Center for Atmospheric Research, Boulder, CO 80303, USA.

E. Roeckner, Max-Planck Institute for Meteorology, D-20146 Hamburg, Germany.

S. Sengupta, Engineering Directorate, Lawrence Livermore National Laboratory, Livermore, CA 94550, USA.

(Received November 27, 2000; revised April 30, 2001;
accepted May 24, 2001.)