



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

# Estimation of natural and anthropogenic contributions to twentieth century temperature change

### Citation for published version:

Tett, SFB, Jones, GS, Stott, PA, Hill, DC, Mitchell, JFB, Allen, MR, Ingram, WJ, Johns, TC, Johnson, CE, Roberts, DL, Sexton, DMH & Woodage, MJ 2002, 'Estimation of natural and anthropogenic contributions to twentieth century temperature change', *Journal of Geophysical Research: Atmospheres*, vol. 107, no. D16, ACL 10, pp. 1-24. <https://doi.org/10.1029/2000JD000028>

### Digital Object Identifier (DOI):

[10.1029/2000JD000028](https://doi.org/10.1029/2000JD000028)

### Link:

[Link to publication record in Edinburgh Research Explorer](#)

### Document Version:

Publisher's PDF, also known as Version of record

### Published In:

Journal of Geophysical Research: Atmospheres

### Publisher Rights Statement:

Published in Journal of Geophysical Research: Atmospheres by the American Geophysical Union (2002)

### General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

### Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



## Estimation of natural and anthropogenic contributions to twentieth century temperature change

Simon F. B. Tett,<sup>1</sup> Gareth S. Jones,<sup>1</sup> Peter A. Stott,<sup>1</sup> David C. Hill,<sup>2,3</sup> John F. B. Mitchell,<sup>1</sup> Myles R. Allen,<sup>2,4</sup> William J. Ingram,<sup>1</sup> Tim C. Johns,<sup>1</sup> Colin E. Johnson,<sup>1</sup> Andy Jones,<sup>1</sup> David L. Roberts,<sup>1</sup> David M. H. Sexton,<sup>1</sup> and Margaret J. Woodage<sup>1</sup>

Received 13 October 2000; revised 5 July 2001; accepted 6 September 2001; published 27 August 2002.

[1] Using a coupled atmosphere/ocean general circulation model, we have simulated the climatic response to natural and anthropogenic forcings from 1860 to 1997. The model, HadCM3, requires no flux adjustment and has an interactive sulphur cycle, a simple parameterization of the effect of aerosols on cloud albedo (first indirect effect), and a radiation scheme that allows explicit representation of well-mixed greenhouse gases. Simulations were carried out in which the model was forced with changes in natural forcings (solar irradiance and stratospheric aerosol due to explosive volcanic eruptions), well-mixed greenhouse gases alone, tropospheric anthropogenic forcings (tropospheric ozone, well-mixed greenhouse gases, and the direct and first indirect effects of sulphate aerosol), and anthropogenic forcings (tropospheric anthropogenic forcings and stratospheric ozone decline). Using an “optimal detection” methodology to examine temperature changes near the surface and throughout the free atmosphere, we find that we can detect the effects of changes in well-mixed greenhouse gases, other anthropogenic forcings (mainly the effects of sulphate aerosols on cloud albedo), and natural forcings. Thus these have all had a significant impact on temperature. We estimate the linear trend in global mean near-surface temperature from well-mixed greenhouse gases to be  $0.9 \pm 0.24$  K/century, offset by cooling from other anthropogenic forcings of  $0.4 \pm 0.26$  K/century, giving a total anthropogenic warming trend of  $0.5 \pm 0.15$  K/century. Over the entire century, natural forcings give a linear trend close to zero. We found no evidence that simulated changes in near-surface temperature due to anthropogenic forcings were in error. However, the simulated tropospheric response, since the 1960s, is  $\sim 50\%$  too large. Our analysis suggests that the early twentieth century warming can best be explained by a combination of warming due to increases in greenhouse gases and natural forcing, some cooling due to other anthropogenic forcings, and a substantial, but not implausible, contribution from internal variability. In the second half of the century we find that the warming is largely caused by changes in greenhouse gases, with changes in sulphates and, perhaps, volcanic aerosol offsetting approximately one third of the warming. Warming in the troposphere, since the 1960s, is probably mainly due to anthropogenic forcings, with a negligible contribution from natural forcings. *INDEX TERMS*: 1650 Global Change: Solar variability; 1694 Global Change: Instruments and techniques; 4215 Oceanography: General: Climate and interannual variability (3309); *KEYWORDS*: climate change, general circulation model, aerosols, greenhouse gases, anthropogenic forcings, natural forcings, detection, attribution

### 1. Introduction

[2] Several authors [e.g., *Santer et al.*, 1996a; *Hegerl et al.*, 1997; *North and Stevens*, 1998; *Tett et al.*, 1999;

<sup>1</sup>Hadley Centre for Climate Prediction and Research, Met Office, Bracknell, UK.

<sup>2</sup>Space Science and Technology Department, Rutherford Appleton Laboratory, Chilton, UK.

<sup>3</sup>Now at Clinical Trials Service Unit, University of Oxford, Oxford, UK.

<sup>4</sup>Department of Physics, Clarendon Laboratory, University of Oxford, Oxford, UK.

*Hegerl et al.*, 2000; *Stott et al.*, 2001] have carried out studies in which they claimed to have detected significant changes in temperature either at the surface or in the free atmosphere. On decadal timescales or longer they attributed changes over the last 30–50 years to anthropogenic rather than natural effects, whether externally forced or due to internal variability. Most of these studies used a variant of the optimal fingerprinting algorithm [*Hasselmann*, 1993; *North et al.*, 1995; *North and Kim*, 1995; *Hasselmann*, 1997; *Hegerl and North*, 1997; *Allen and Tett*, 1999].

[3] Previous work has found evidence that cooling from anthropogenic sulphate aerosols has offset warming from changes in greenhouse gases (e.g., *Mitchell et al.*, [1995]). *Barnett et al.* [1999] and *Hegerl and Allen* [2002] found

evidence that in a few cases the simulated linear trends in northern summer temperature due to sulphate aerosols were inconsistent with the observations. One of these cases, using a model with an interactive sulphur cycle and a physically based representation of the direct and indirect effect of aerosols, had an amplitude of the “sulphate” component that was inconsistent with observations. It is not clear whether this was due to the use of a single simulation as opposed to an ensemble, to errors in the representation of the sulphate signal, or to neglect of other forcing factors. The first possibility can be addressed by using an ensemble of simulations to produce a more representative estimate of the sulphate signal, the second by using different models to estimate the sulphate forcing, and the third by inclusion of other forcings.

[4] In this paper we address all three of these issues. We use ensembles of simulations to produce a more accurate estimate of the model’s response to forcing. We use a different model to estimate the response to external forcing and, in particular, use different and physically based estimates of the direct and indirect forcing due to sulphates. We also include estimates of the response to natural forcings (due to changes in solar irradiance and major volcanic eruptions), which have been neglected in many previous detection and attribution studies. Finally, we quantify the contribution that various combinations of forcing agents have made to twentieth century temperature change.

[5] Tett *et al.* [1999] (hereinafter referred to as T99) and Stott *et al.* [2001] (hereinafter referred to as S01) computed responses from the Atmosphere/Ocean General Circulation Model (AOGCM) HadCM2 [Johns *et al.*, 1997] to solar, volcanic, greenhouse, and direct anthropogenic sulphate forcing. They compared the responses with observations of near-surface temperature using a spatiotemporal methodology and concluded that natural causes alone could not explain observed changes in surface temperature from 1946 to 1996. HadCM2 included an ocean model with a resolution of  $2.5^\circ \times 3.75^\circ$  and needed a flux adjustment to keep the control simulation stable and to keep its climate close to the current climate. (Flux adjustments are artificial fluxes of heat and water which vary in space and throughout the seasonal cycle but are constant from year to year and in all the HadCM2 simulations.) It represented all greenhouse gases as equivalent  $\text{CO}_2$  and represented the direct effect of sulphates as changes in surface albedo.

[6] In the studies reported in this paper we use a new AOGCM, HadCM3 [Gordon *et al.*, 2000; Pope *et al.*, 2000]. HadCM3 has 19 atmospheric levels with a resolution of  $2.5^\circ \times 3.75^\circ$ , and the ocean component has 20 levels with a resolution of  $1.25^\circ \times 1.25^\circ$ . HadCM3 has a climate sensitivity of  $0.9 \text{ K/Wm}^{-2}$  corresponding to an equilibrium warming of 3.35 K for a doubling of  $\text{CO}_2$  concentrations [Williams *et al.*, 2001]. In addition to an increase in oceanic resolution, it includes many improvements on HadCM2 that have removed the need for a flux adjustment. HadCM3 represents the radiative effects of  $\text{CO}_2$ ,  $\text{N}_2\text{O}$ ,  $\text{CH}_4$ , and some of the (hydro)(chloro)fluorocarbons (H)(C)FCs individually. The direct effect of sulphate aerosol is now simulated using a fully interactive sulphur cycle scheme that models the emissions, transport, oxidation, and removal of sulphur

species. The first indirect effect of sulphate aerosol [Tomey, 1974], which was not represented at all in HadCM2, is now modeled using a relatively simple, noninteractive technique.

[7] The control simulation is stable for multicentury integrations, and the temperature variability near the surface, though not in the free atmosphere, compares well with observations [Collins *et al.*, 2001]. Simulated ENSO has greater variance than observed, but its structure and time-scale are comparable to those observed. In the free atmosphere the model has less variability in the stratosphere and over parts of the Northern Hemisphere than do the observations. HadCM2 and HadCM3 show similar global mean temperature responses to increases in greenhouse gases during the 20th and 21st centuries, but HadCM3 shows less tropical warming than HadCM2 due to changes in both the boundary layer scheme and the critical relative humidity at which clouds form [Williams *et al.*, 2001].

[8] We present an analysis based on changes in near-surface temperature change from 1897 to 1997. In order to compare results with earlier work using HadCM2, we also consider changes in near-surface temperature on 50-year timescales, as given by T99 and S01, and changes in the temperature of the free atmosphere on 35-year timescales [Tett *et al.*, 1996; Allen and Tett, 1999] (hereinafter referred to as T96 and AT99).

[9] In the rest of this paper we first describe the simulations, radiative forcings, and observations. We then describe the simulated responses and compare them with observations. Next, we describe the detection and attribution methodology. Section 4 starts with an outline of the methodology we use, with the details given in section 4.1 onward. In section 5 we show the results of the analyses, and in section 6 we give conclusions.

## 2. Simulations

[10] The control simulation for HadCM3 (CONTROL) has constant, near-preindustrial forcing, and we use the first 1100 years of the simulation in our analysis. (The concentrations (in ppbv) used for the well-mixed greenhouse gases are  $\text{CO}_2$ , 289,600;  $\text{CH}_4$ , 792.1; and  $\text{N}_2\text{O}$ , 285.1. The (H)(C)FCs all had zero concentrations.) Four ensembles with different external forcings were carried out using HadCM3. Each ensemble consisted of four simulations. The ensembles are (1) GHG, where the simulations were forced with historical changes in well-mixed greenhouse gases; (2) TROP-ANTHRO, where the simulations were forced with changes in well-mixed greenhouse gases (as GHG), anthropogenic sulphur emissions and their implied changes to cloud albedos, and tropospheric ozone; (3) ANTHRO, similar to TROP-ANTHRO except that from 1974, stratospheric ozone decline was included; and (4) NATURAL, where simulations were forced with the solar irradiance time series of Lean *et al.* [1995a] and with a time series of stratospheric aerosol due to explosive volcanic eruptions [Sato *et al.*, 1993] (both forcing time series have been extended to 1997).

[11] Four sets of initial conditions to start the GHG, ANTHRO, and NATURAL ensembles were taken from states in CONTROL separated by 100 years. Note that, for example, the first GHG and NATURAL simulations use

the same initial conditions. All simulations except TROP-ANTHRO start on 1 December 1859, and the 12 anthropogenic simulations ended on 30 November 1999. The NATURAL simulations were integrated to 30 November 1997. Initial conditions for TROP-ANTHRO were taken from ANTHRO on 1 December 1974.

## 2.1. Forcing Factors

### 2.1.1. Well-mixed greenhouse gases

[12] CO<sub>2</sub>, CH<sub>4</sub>, N<sub>2</sub>O and six (H)(C)FCs (CF<sub>2</sub>Cl<sub>2</sub>, CFC1<sub>3</sub>, CF<sub>3</sub>CFH<sub>2</sub>, CHF<sub>2</sub>Cl, CF<sub>2</sub>C1CFCl<sub>2</sub>, and C<sub>2</sub>HF<sub>5</sub>) were included, with constant mass mixing ratios everywhere. Historical values were used to 1990 [Schimel *et al.*, 1996]. From 1990 to 2000 the preliminary B2 Special Report on Emission Scenarios (SRES) scenario was used [Nakićenović *et al.*, 2000, p. 373]. (Differences between any of the SRES scenarios only occur from 2000 on, but as we use linear interpolation to obtain intermediate values, then our “historical” values will be affected by the 2000 values.) Johns *et al.* [2001] describe how emissions of greenhouse gases were converted to concentrations. Johns *et al.* [2001, Tables 1a–1d] show the greenhouse gas concentrations and sulphur emissions used in the anthropogenic simulations described in this paper.

### 2.1.2. Sulphates

[13] In the ANTHRO and TROP-ANTHRO simulations the model’s interactive sulphur cycle scheme (described by Jones *et al.* [1999]) was used to compute the distribution of anthropogenic sulphate aerosol, which was then passed to the model’s radiation scheme [Edwards and Slingo, 1996; Cusack *et al.*, 1999] for computation of its direct radiative effect. No natural emissions were included, as we assumed that the natural background of tropospheric sulphate aerosol was constant.

[14] Estimates of the anthropogenic SO<sub>2</sub> emissions were taken from Orn *et al.* [1996] for 1860–1970, from the Global Emissions Inventory Activity (GEIA) 1B data set for 1985, and from the preliminary International Panel on Climate Change SRES data sets for 1990 and 2000 [Nakićenović *et al.*, 2000] and were linearly interpolated between these times. Since the distribution of sulphate aerosol is influenced by the height at which SO<sub>2</sub> emissions occur, we assumed that a fraction of the emissions originate from elevated sources such as power station chimneys. This fraction depends on location and from 1985 onward is prescribed using the information in the GEIA 1B data set. Before 1950 it is assumed to be zero, and between 1950 and 1985 the fraction is linearly interpolated in time.

[15] CONTROL had fixed cloud droplet number concentrations, and our simulations included only anthropogenic sulphur emissions. Thus we computed the indirect effect of anthropogenic sulphates on cloud albedo using two sets of offline simulations of a modified version of HadAM3 (the atmospheric component of HadCM3). Both sets of simulations used present-day concentrations of well-mixed greenhouse gases and seasonally varying sea surface temperatures (SSTs). The first set used anthropogenic emissions of sulphur for 1860, 1900, 1950, 1975, and 2000, as well as natural emissions, to compute annual mean distributions of sulphate aerosols for these years. The second set of offline simulations was run with the radiation scheme being called twice, using aerosol distributions calculated by the previous

set of simulations as input to these two calls. These aerosol distributions change the three-dimensional distribution of cloud albedo by affecting the cloud droplet concentrations seen by the radiation scheme. The difference in cloud albedo between the two radiation calls is a measure of the indirect effect of the difference between the two aerosol distributions. This set of offline runs was used to generate a time series of three-dimensional changes in cloud albedo caused by the indirect effect. These fields were then annual averaged, linearly interpolated in time, and used in the HadCM3 simulations to modify the albedo of the clouds so as to simulate the indirect effect.

[16] In the HadCM3 simulations the radiative forcing due to the indirect effect is ~60–70% of that in the atmosphere-only simulations used to compute the albedo perturbations, because the meteorology is different in the coupled and atmospheric simulations. Cloud albedo perturbations applied to a region in HadCM3 that, unlike the atmosphere-only simulations of HadAM3, has no cloud, will clearly have no effect. In areas where the coupled simulation has cloud but the atmosphere-only simulation does not, there will again be no albedo perturbation applied, as clouds are needed in the atmosphere-only simulation to generate this perturbation.

[17] A separate study [Jones *et al.*, 1999], using HadAM3 with a new cloud microphysics parameterization [Wilson and Ballard, 1999] driven by both natural and anthropogenic sulphur emissions, suggested that the model has roughly half the near-surface concentration of anthropogenic sulphate aerosol, compared with data from a European network (European Monitoring and Evaluation Programme). This implies that the direct forcing due to anthropogenic sulphate is less than in reality. However, because the size of the indirect effect is related nonlinearly to the difference between the natural background and anthropogenically perturbed aerosol, underestimating the true aerosol concentration could increase (less natural background) or decrease (smaller increase in aerosol) the indirect forcing. More details on the parameterization of the direct and indirect effects of sulphates in HadCM3 are given by Johns *et al.* [2001].

### 2.1.3. Tropospheric ozone

[18] Three-dimensional fields of monthly-mean tropospheric ozone were computed using the STOCHEM chemical model [Collins *et al.*, 1997] for 1860, 1900, 1950, 1975, 1990, and 2000. Values of ozone between those years were interpolated by assuming linearity between increases in observed methane concentration and modeled tropospheric ozone for each month in the year. Estimates of historical anthropogenic emissions of NO<sub>x</sub>, CO, CH<sub>4</sub>, and volatile organic compounds were obtained by scaling their present-day emissions to the estimated time variation of NO<sub>x</sub> emissions of Dignon and Hameed [1990]. Biomass burning emissions were estimated by assuming that preindustrial values were 20% of present-day values and that emissions increased linearly with population. These crudely estimated emissions have both seasonal and geographic variations and are only used as input to STOCHEM to compute tropospheric ozone. Below the mean model-diagnosed tropopause the anomalies from STOCHEM’s preindustrial values were zonally averaged, interpolated to HadCM3’s levels, and then added to the HadCM3 preindustrial values.

(The tropopause was diagnosed using a simulation of HadAM3 forced with historical SSTs and ice (similar to that by Rowell [1998] using HadAM2b) for the period 1860–1997. The tropopause was diagnosed at every point and at every radiation time step, using the same lapse rate criteria used in the Met Office’s operational forecast model, which are based on the World Meteorological Organization rules for reporting observations. Ozone concentrations above and on the tropopause were set to estimated preindustrial values.

#### 2.1.4. Stratospheric ozone

[19] Randel and Wu [1999] estimated seasonally and zonally varying trends in stratospheric ozone. From 1975 to 1979 we add half these trends to the annual cycle of preindustrial ozone above the mean model-diagnosed tropopause. After 1979, when stratospheric ozone decline is believed to have accelerated, the full trends were added. Ozone mass mixing ratios below  $10^{-11}$  were set to  $10^{-11}$ .

#### 2.1.5. Volcanic aerosol

[20] The updated time series of volcanic aerosol depth given by Sato *et al.* [1993] was distributed above the model tropopause, assuming a uniform mass mixing ratio. Note that the tropopause was diagnosed as the simulations proceeded, not prescribed as for the ozone changes.

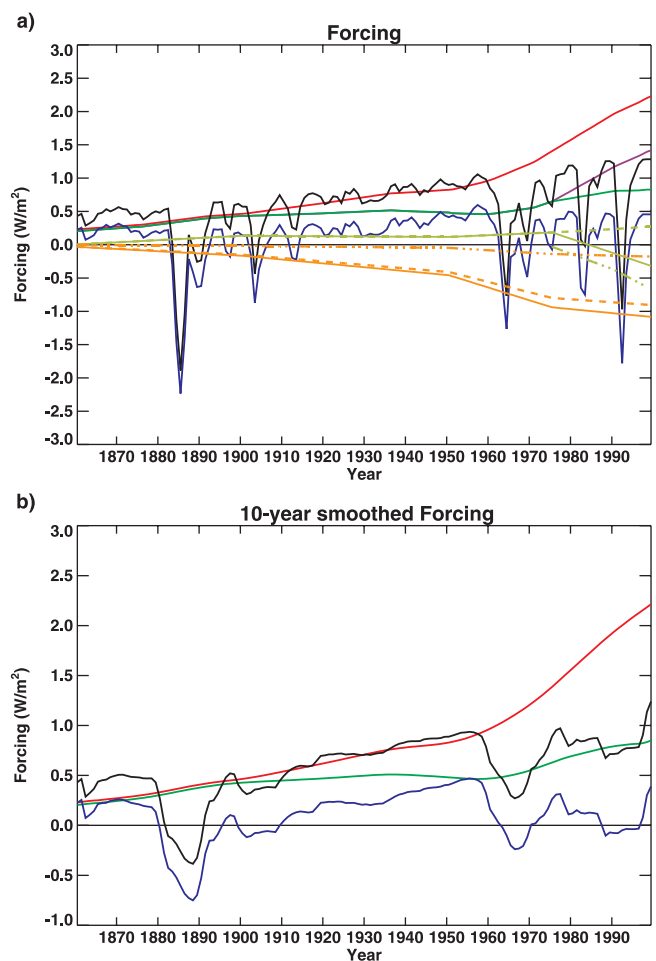
#### 2.1.6. Solar irradiance changes

[21] Changes in solar irradiance, as estimated by Lean *et al.* [1995a], were applied to the model by varying the “solar constant” in the model’s radiation scheme with the changes spread over the solar spectrum following Lean *et al.* [1995b]. This allows for the effects of ozone absorption. There is considerable uncertainty in the reconstructions of solar forcing, and we could have chosen to use other reconstructions [e.g., Hoyt and Schatten, 1993], but computational restrictions limited us to only one. We selected the reconstruction by Lean *et al.* [1995a] rather than that by Hoyt and Schatten [1993] on the toss of a coin. There has been some speculation about possible amplification of solar forcing due to changes in ozone [Haigh, 1994, 1996] or in low cloud [Svensmark and Friis-Christensen, 1997; Marsh and Svensmark, 2000b, 2000a], but we have not included these effects in our simulations.

## 2.2. Simulated Forcings

[22] Simulated forcings were computed for the various factors, using a diagnosed tropopause whose height can change (see Appendix A for details). The forcing due to greenhouse gases reaches a maximum of  $>2 \text{ W/m}^2$  by 2000 (Figure 1a). By contrast, the total anthropogenic forcing reaches a maximum of  $\sim 0.8 \text{ W/m}^2$ , while the forcing due to tropospheric anthropogenic forcings reaches a maximum value of almost  $1.5 \text{ W/m}^2$  in 2000. The difference between the two is due to a strong negative forcing from stratospheric ozone decline. Using 1998 conditions we found that the forcing due to stratospheric ozone was  $-0.5 \pm 0.1 \text{ W/m}^2$ . When we repeated these calculations using a fixed tropopause, the ozone forcing increased to  $-0.3 \text{ W/m}^2$ . These results are outside the range of Schimel *et al.* [1996], and we plan to investigate this difference in more detail in subsequent work.

[23] Natural forcings from about 1910 to 1950 show a general increase due to both an increase in solar irradiance and a lack of explosive volcanic eruptions after the 1912



**Figure 1.** (a) Annual mean radiative forcings and (b) 10-year smoothed forcings for GHG (red), TROP-ANTHRO (purple), ANTHRO (green), and NATURAL (blue). Anthropogenic forcings are shown relative to preindustrial times and natural forcing is shown relative to its time mean. Total forcing computed by summing the NATURAL and ANTHRO forcings is shown in black. Also shown in Figure 1a are the total forcing due to sulphates (solid orange line) and the direct (dot-dashed orange line) and indirect (dashed orange line) effects; total ozone changes (solid light green line), tropospheric ozone changes (dashed light green line), and stratospheric ozone changes (dot-dashed light green line), with the latter two only being shown for the period 1975–1998.

Katmai eruption. Apparent in this time series are the solar cycle and large negative excursions due to the eruptions of Agung (1963), El Chichón (1982), Pinatubo (1991), and other volcanoes. Pinatubo causes the largest negative forcing of the twentieth century with, in 1991, an annual average global mean forcing of  $-2.5 \text{ W/m}^2$ , which, when added to a solar forcing of  $\sim 0.5 \text{ W/m}^2$ , gives a total natural forcing of  $-2 \text{ W/m}^2$ . The forcing due to volcanoes in HadCM3, after stratospheric adjustment, is  $\sim 20 \text{ W/m}^2$  per unit optical depth, less than the  $30 \text{ W/m}^2$  (without adjustment) per unit optical depth quoted by Lacis *et al.* [1992]. This suggests a high degree of uncertainty in radiative forcing due to volcanic aerosol. Total natural and anthropogenic forcing

shows a complex structure with a general slow increase until the 1960s, after which total forcing is approximately constant, though punctuated by volcanic eruptions.

[24] The negative forcing due to the direct effect of sulphates is very small and is largely balanced by the small positive forcing due to tropospheric ozone changes (Figure 1a). This forcing, relative to preindustrial, is  $0.23 \text{ W/m}^2$  in 1990 and is below the lower end of the range quoted by *Stevenson et al.* [1998] of  $0.28\text{--}0.42 \text{ W/m}^2$ . Some of the difference between our results and those of *Stevenson et al.* [1998] may be due to their use of fixed dynamical heating versus our computing stratospheric adjustment from the simulated fluxes (Appendix A). There are two large negative forcings: that due to the indirect effect of sulphate aerosols and that due to stratospheric ozone. Both of these are highly uncertain [*Schimel et al.* 1996]. With the exception of ozone forcing, our computed anthropogenic forcings are all within the ranges quoted by *Schimel et al.* [1996].

[25] Ten-year smoothed natural forcings (Figure 1b) reached their maximum value in the 1950s and then fell. The 1960s are a period with small total forcing and with negative smoothed natural forcing due to two large tropical volcanic eruptions: Agung and Fernandina. Total natural and anthropogenic forcing reached a local maximum in the 1950s that, according to our calculations, was only exceeded toward the end of the twentieth century.

### 2.3. Observed Data Sets and Data Processing

[26] We compare the results of the model simulations with an updated version of the surface temperature data set of *Parker et al.* [1994] and with the HadRT2.1s radiosonde temperature data set, an updated version of that of *Parker et al.* [1997]. Radiosonde data from the Indian subcontinent ( $60^\circ\text{E}\text{--}90^\circ\text{E}$ ,  $0^\circ\text{--}30^\circ\text{N}$ ) were removed because of apparent problems with their quality, and the remaining data were corrected for known changes in instruments by comparison with colocated Microwave Sounding Unit data [*Parker et al.*, 1997].

[27] Annual averages of both the surface and radiosonde data sets were computed from monthly mean temperature anomalies. At each location we required there to be at least 8 months of observations; otherwise, we discarded the annual mean value.

[28] The annual mean surface observations were decadal averaged, with periods ending in 1997. For each decade we required that there be at least 5 years of data; otherwise the decadal mean value was discarded. In our analysis of surface temperature we consider changes on 50-year and 100-year timescales using decadal data with the 50-year or 100-year average removed. Locations in the observations at which  $<3$  ( $<5$ ) decades were present were omitted in the 50-year (100-year) analysis. These data were then filtered, using spherical harmonics, to remove scales below 5000 km (T99, S01). Harmonics were further weighted by  $1/\sqrt{2l+1}$  ( $l$  is the total spherical harmonic wave number) to give each spatial scale included equal weight [*Stott and Tett*, 1998]. Simulated data were decadal averaged, bilinearly interpolated in latitude and longitude to the observational grid. Simulated data were discarded where there were no observational data and were then processed in the same way as the observations.

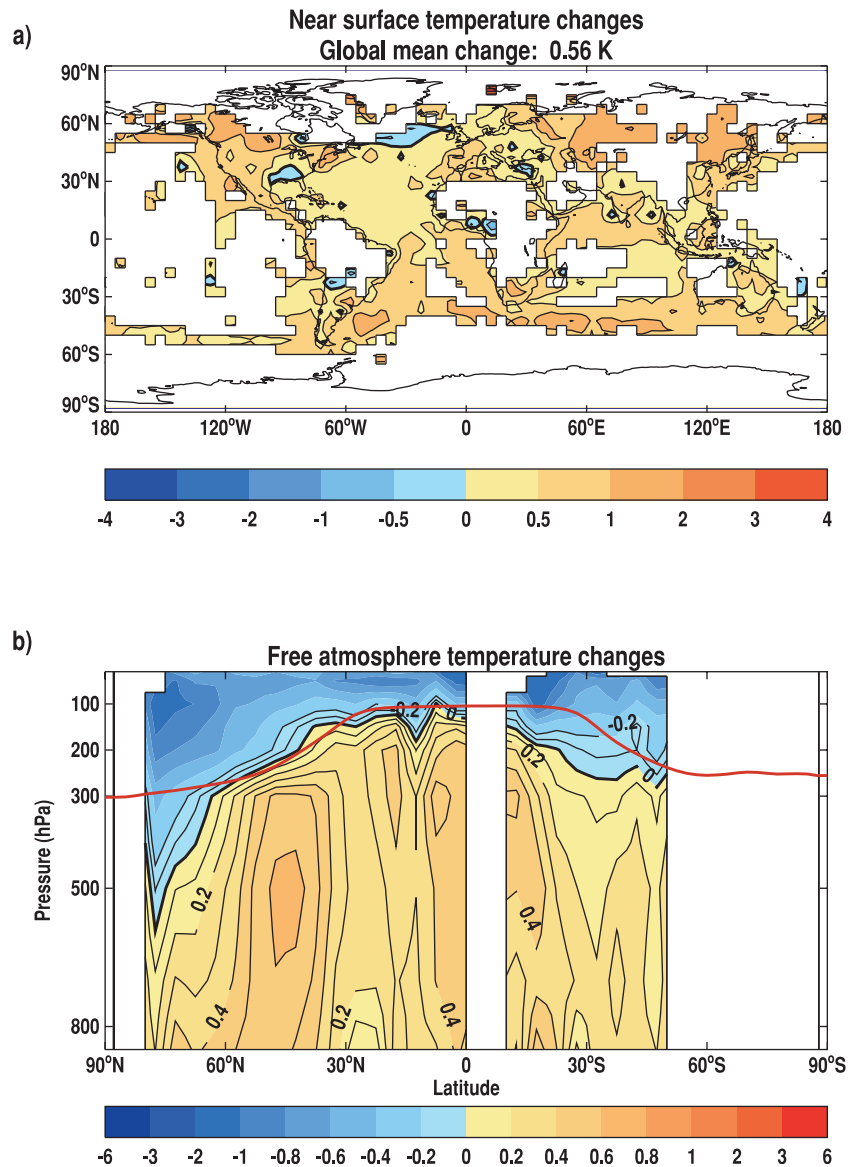
[29] When computing global mean time series, we first bilinearly interpolated (latitude and longitude) simulated annual mean near-surface temperature data to the observational grid, discarding simulated data where there were no observational data. Since the observed data are anomalies relative to 1961–1990, we computed the 1961–1990 climate mean for each simulation and for the observations, removed it, and computed global means. In order to show changes relative to the beginning of the century, we removed the global mean time average for 1881–1920 from each time series.

[30] Annual mean simulated data from throughout the atmosphere were trilinearly (pressure, longitude, and latitude) interpolated to the three-dimensional observed grid and was discarded where there were no observed data. We then processed the simulations and observations by first removing the 1971–1990 mean, zonally averaging (requiring that there be four longitudes with data present in any zonal band), and then computing the difference between 1985–1995 and 1961–1980. Unlike T96 and AT99, simulated data had the observational mask applied and the 1971–1990 normal removed before zonal averaging. This change in processing had little impact on the signals and tended to reduce slightly the variability of the annual average zonal mean temperatures [*Collins et al.*, 2001].

[31] Changes in surface temperature observed over the century show warming (Figure 2a) over most of the world with, in general, land warming more than the ocean, central Eurasia and Canada warming most, and cooling occurring in parts of the North Atlantic to the south of Greenland and Iceland. The free-atmosphere changes show cooling (Figure 2b) in the stratosphere and show warming in the troposphere. The cooling extends down to 500 hPa above the Arctic, far below the reanalysis tropopause. The tropospheric warming is uneven, with a maximum warming of 0.6 K occurring at  $\sim 50^\circ\text{N}$  and with almost no warming at  $30^\circ\text{N}$ . Differences between the observations shown here and that of T96 (see their Figure 2d) are due to the continued development of the radiosonde data set and due to removal of data from the Indian subcontinent.

### 3. Model and Observed Temperature Responses

[32] First, we consider the annually and globally averaged temperature response to each set of forcings (i.e., averaged over each set of ensembles) (Figure 3). (Global mean, in this context, means the area-weighted average over all locations where there are data. Recall that simulated data were discarded where there were no observations.) All ensemble averages differ from the observations. From the 1920s until the 1950s, GHG warms less than the observations. From the 1940s onward it begins to warm, and by the end of the twentieth century it has warmed more, over the century, than the observations. Addition of sulphates and ozone to GHG, giving ANTHRO, delays the simulated warming until the 1960s. From then until the end of the century, ANTHRO, TROP-ANTHRO, and the observations warm at approximately the same rate. The small differences between ANTHRO and TROP-ANTHRO suggest that stratospheric ozone changes have little impact on near-surface temperature despite the large differences in radiative forcing (Figure 1). We believe that this small response is due to the



**Figure 2.** Observed changes in (a) near-surface temperature (1977–1997 minus 1881–1920) and (b) in zonal mean temperature (1985–1995 minus 1961–80). In Figure 2a a contour interval of 1 K is used from  $-4$  to 4 K, with additional contours at  $\pm 0.5$  K. In Figure 2b contour interval of 0.1 K is used, with every second contour labeled from  $-0.2$  to 1 K. Colors change every 0.2 K from  $-1$  to 1 K, with additional changes at  $\pm 6$ ,  $\pm 3$ , and  $\pm 2$  K. Red line denotes the zonal mean position of the tropopause from the NCEP/NCAR reanalysis for the period 1985–1995 using data provided by the NOAA-CIRES Climate Diagnostics Center, Boulder, Colorado (available from <http://www.cdc.noaa.gov/>).

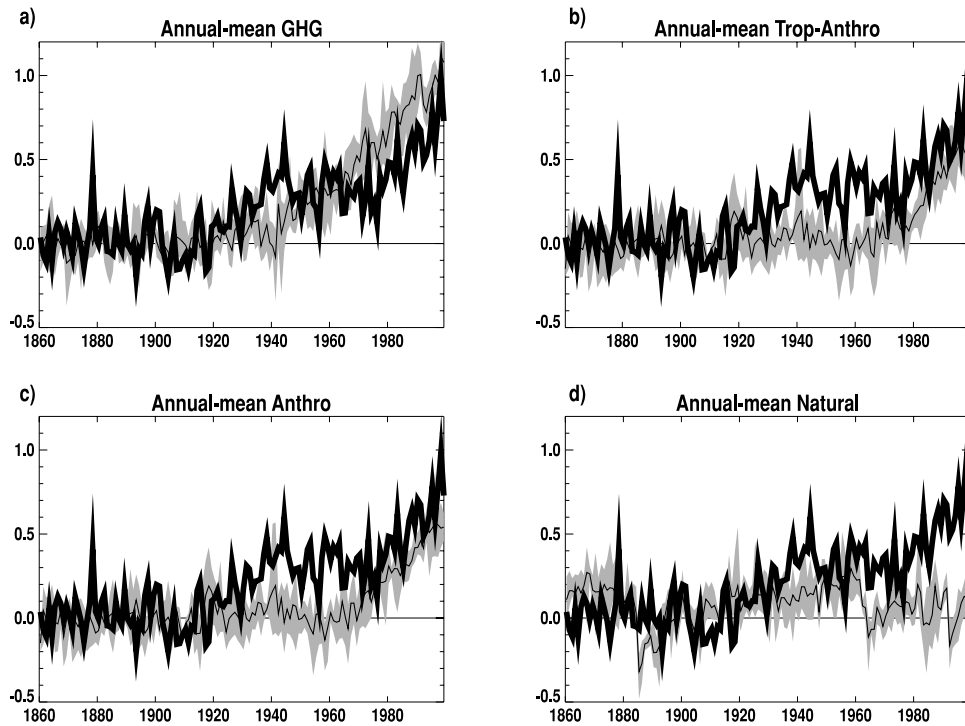
stratospheric ozone forcing being concentrated over Antarctica and due to the forcing changing the temperatures of the southern ocean (a region of deep mixing) rather than the land. This is to be the subject of a separate investigation.

[33] Simulated natural forcings produce a general warming from the 1910s until the eruption of Agung in 1963. After this the observations warm while the subsequent eruptions of El Chichón and Pinatubo cool NATURAL.

[34] The patterns of simulated response from the twentieth century are shown in Figure 4. All three anthropogenic ensembles (GHG, TROP-ANTHRO, and ANTHRO) produce more warming over land than over the sea. GHG has the most warming of these ensembles and warms more than

the observations. In the GHG ensemble the Arctic warms most, while the North Atlantic and large regions of ocean in the Southern Hemisphere warm considerably less than the global average (Figure 4a). ANTHRO and TROP-ANTHRO are in reasonable agreement with the observations (Figure 2a), and both warm less than GHG, especially in the midlatitudes of the Northern Hemisphere, where the sulphate cooling will be large. NATURAL shows no distinctive signal, probably because there is little change in natural forcing between the start and end of the century (Figure 1).

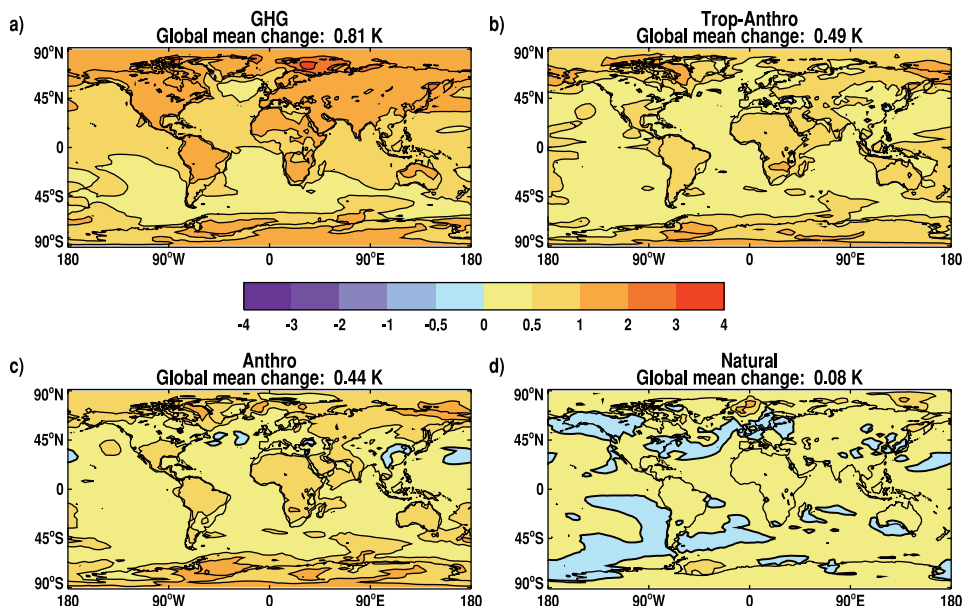
[35] The simulated anthropogenic changes are, in most parts of the world, outside the range of simulated internal



**Figure 3.** Near-surface changes in global mean temperature, relative to the 1881–1920 mean for the observations (thick black lines) and the ensemble mean of the (a) GHG, (b) TROP-ANTHRO, (c) ANTHRO, and (d) NATURAL simulations (thin black lines). Maximum and minimum ranges from the individual simulations are shaded.

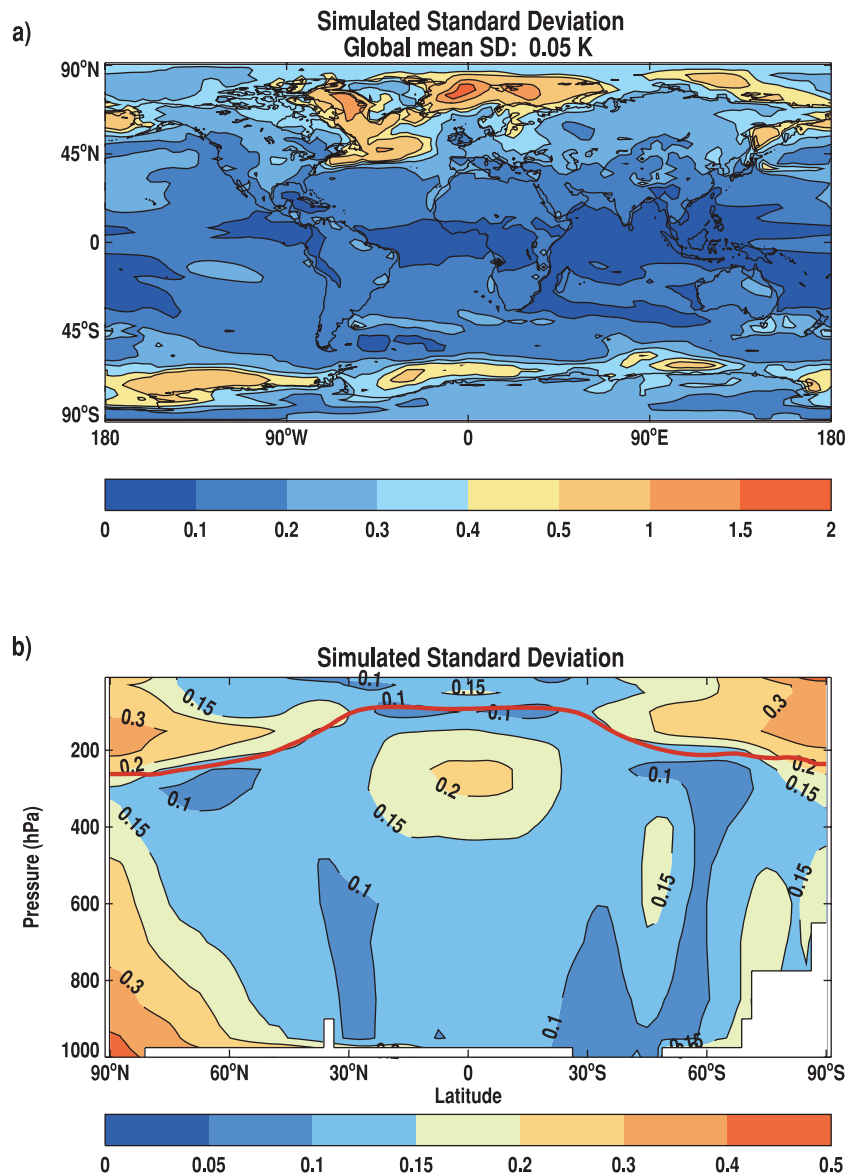
variability (compare changes shown in Figure 4 with the standard deviations shown in Figure 5a), while the response to natural forcings is, in most regions, within the range. On the global scale, NATURAL is also outside the range of

expected variability. On century timescales the regions of highest variability are the North Atlantic and near the sea-ice edge, while lowest variability occurs over the tropical Atlantic and Indian Oceans.



**Figure 4.** Temperature difference (in K) between the 20-year average 1977–1997 and the 40-year average 1881–1920 for the four ensembles, (a) GHG, (b) TROP-ANTHRO, (c) ANTHRO, and (d) NATURAL. Note that these plots show the raw model data (i.e., without the observed mask). All other details are as in Figure 2a.



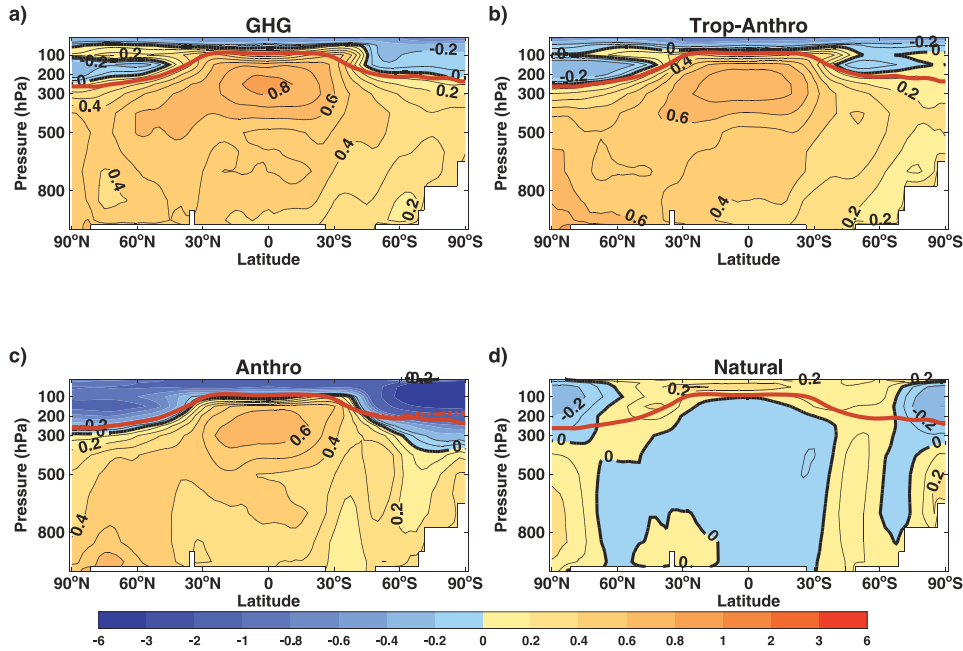


**Figure 5.** (a) Standard deviation of surface temperature changes computed from CONTROL for data shown in Figure 4. A contour interval of 0.1 is used from 0 to 0.5. From 0.5 till 2.0 an interval of 0.5 is used. Also shown is the standard deviation of the global means. (b) Simulated standard deviation of zonal mean temperature changes computed from CONTROL for data shown in Figure 6. Contours are drawn at 0, 0.05, 0.1, 0.15, 0.2, 0.3, 0.4, and 0.5. Red line shows the position of the zonal mean tropopause. Note when comparing changes shown in Figures 4 and 6 with these standard deviations that, because of use of four-member ensembles, uncertainties are halved; the two-sigma uncertainty for the ensemble will be one control standard deviation.

[36] We now examine temperature changes throughout the atmosphere between the decade 1985–1995 and the 20-year period 1961–1980. All three anthropogenic ensembles have similar warming in the troposphere and greatest warming in the upper tropical troposphere and warm more in the Northern Hemisphere than in the Southern Hemisphere (Figure 6). The upper tropical troposphere and Southern Hemisphere warm more in GHG than in TROP-ANTHRO, while high northern latitudes warm less. The latter could be due to the effects of tropospheric ozone or to internal climate variability. Neither simulation cools the stratosphere or upper troposphere as much as do the

observations (Figure 2b). Inclusion of stratospheric ozone decline in ANTHRO produces large stratospheric cooling (of up to 6 K over Antarctica), especially in high latitudes, which brings this ensemble into better agreement with the observations (Figure 6c). Unlike the anthropogenic simulations, NATURAL warms in the tropical stratosphere, probably due to the 1991 Pinatubo eruption, but has little temperature response in the troposphere.

[37] The anthropogenic simulations are, over most of the free atmosphere, outside the range of internal variability. (Compare Figure 6 with Figure 5b). However, in the troposphere the response to natural forcings is within the range of



**Figure 6.** Differences (in K) between 1985–1995 and 1961–1980 for the four ensembles, (a) GHG, (b) TROP-ANTHRO, (c) ANTHRO, and (d) NATURAL. Solid red lines show the position of the mean tropopause in CONTROL, while the dash-dotted red line in Figure 6c shows the mean position of the tropopause in an atmosphere-only simulation with 1990 stratospheric ozone. Maximum difference between the two lines is  $\sim 50$  hPa. All other details are as in Figure 2b.

internal temperature variability, though in the tropical stratosphere it is not. Variability in the free atmosphere on multidecadal timescales is small throughout most of the troposphere and equatorial stratosphere (Figure 5b). Greatest simulated variability occurs in the polar stratosphere, near the polar surface, and in the upper tropical troposphere.

[38] The boundary between cooling and warming is close to the tropopause in all ensembles except over Antarctica in ANTHRO (Figure 6). In this ensemble the cooling over Antarctica extends down to 500 hPa, and the tropopause rises, its pressure falling by 50 hPa. The data over Antarctica are insufficient to tell if this occurred in reality. However, the observed Arctic cooling down to 500 hPa is not present in any of the ensembles.

[39] Qualitative comparison of our ensembles with the observations suggests that ANTHRO is the most similar to the observations (compare Figure 6c with Figure 2b). Since all the anthropogenic ensembles are quite similar in the troposphere, it appears that, since the 1960s, increases in greenhouse gases and stratospheric ozone decline are the most important contributors to temperature changes in the free atmosphere.

#### 4. Detection and Attribution Methodology

[40] One of the main problems in attributing climate change to possible causes arises from the difficulties in estimating the radiative forcing and climate response due to different forcings. In particular, there are large uncertainties in the overall magnitude of the climate response to a given forcing due, for example, to uncertainties in climate sensitivity or in the rate of ocean heat uptake [Kattenberg *et al.*, 1996]. The size of the forcing associated with many of the

factors other than well-mixed greenhouse gases, notably aerosols, is also uncertain [Shine *et al.*, 1995]. To reduce the impact of these uncertainties, we use a methodology first proposed by Hasselmann [1979], which has been shown to be a form of multivariate regression (AT99). Both assume that the observations ( $\mathbf{y}$ ) may be represented as a linear sum of simulated signals ( $\mathbf{X}$ ) and internal climate variability ( $\mathbf{u}$ ),

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{u}, \quad (1)$$

where  $\beta_i$  is the scaling factor, or amplitude, that we apply to the space-time ensemble average signal ( $\mathbf{x}_i$ ), corresponding to forcing  $i$ , to obtain the best fit to the observations. In this paper the signals are ensemble averages from the simulations described in section 2. Any errors in the magnitude of the forcing and climate responses are allowed for through scaling the model responses ( $\mathbf{x}_i$ ) by the signal amplitudes  $\beta_i$ . Errors in the patterns of forcing and response are not taken into account by this procedure. The values of  $\beta$  that give the best fit (the best-estimate value  $\tilde{\beta}$ ) to observations, using the standard linear regression approach, are those of (AT99):

$$\tilde{\beta} = (\mathbf{X}^T \mathbf{C}_N^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{C}_N^{-1} \mathbf{y} = \mathbf{F}^T \mathbf{y}, \quad (2)$$

where  $\mathbf{C}_N$  is the covariance matrix of internal variability ( $\mathcal{E}(\mathbf{u}\mathbf{u}^T)$ ) estimated, in our case, from simulations of coupled-atmosphere ocean global climate models. We do not normally have enough data to accurately estimate the inverse covariance matrix ( $\mathbf{C}_N^{-1}$ ), so we estimate its inverse from a truncated representation of it based on its leading eigenvectors. Simulated and observed data are also filtered by projection onto these eigenvectors.

[41] Both the observations and signals include internal climate variability (noise), which leads to uncertainty in  $\beta$ . We estimate uncertainty ranges (the 5–95% range unless stated otherwise) in  $\beta$  using its covariance matrix [AT99; Mardia et al., 1979],

$$\tilde{\mathbf{V}}(\tilde{\beta}) = \mathbf{F}^T \mathbf{C}_{N_2} \mathbf{F}, \quad (3)$$

where  $\mathbf{C}_{N_2}$  is an estimate of  $\mathcal{E}(\mathbf{u}\mathbf{u}^T)$  using data that are statistically independent of those used to estimate  $\mathbf{C}_N$ .

[42] We perform two related tests:

1. “Detection” tests the null hypothesis that the observed response to a particular forcing or combination of forcings is zero. We do this by computing the two-tailed uncertainty range about  $\tilde{\beta}$  using  $\tilde{\mathbf{V}}(\tilde{\beta})$  and testing whether it includes zero. Rejection of this null and a positive value of  $\beta_i$  implies detection.

2. “Amplitude consistency” tests the null hypothesis that the amplitude of the observed response is consistent with the amplitude of the simulated response. We do this by computing the two-tailed uncertainty range about  $\tilde{\beta}$  using  $\tilde{\mathbf{V}}(\tilde{\beta})$  and testing whether it includes unity. In this test we inflate  $\tilde{\mathbf{V}}(\tilde{\beta})_{ij}$  by a factor of  $\sqrt{(1 + 1/m_i)}\sqrt{(1 + 1/m_j)}$ , where  $m_i$  and  $m_j$  are the ensemble sizes, to compensate for sampling noise in the signals. Failure of this test means that the simulated signal amplitude is inconsistent with the observations. When we report consistency with unity, we mean that it is neither greater than nor less than unity at a given confidence level.

Unless otherwise stated, results are reported as significant if the relevant null hypothesis can be rejected at the 5% level. All reported uncertainty ranges are 5–95%.

[43] The best estimate of the temperature trend (or of any other linear diagnostic such as change in global mean temperature), due to a forcing factor, is the product of the signal amplitude and the trend computed from the appropriate ensemble average. The covariance matrix used to compute uncertainties is computed by multiplying  $\tilde{\mathbf{V}}(\tilde{\beta})_{ij}$ , inflated by  $\sqrt{(1 + 1/m_i)(1 + 1/m_j)}$  to approximately compensate for signal noise, by the trends of the  $i$ th and  $j$ th ensembles.

[44] Covariance matrices are estimated from intraensemble variability (i.e., variability within the ensemble) and from CONTROL. To obtain these estimates, we process data in exactly the same manner as we do the observations and simulations giving the  $\mathbf{u}$  in equation (1). In all our analyses, realizations of  $\mathbf{u}$  were overlapped by 10 years. When computing covariance matrices from intraensemble variance, we remove the ensemble average and scale each realization by a factor of  $\sqrt{(m - 1)/m}$ , where  $m$  is the number of ensemble members.

[45] In section 5 we analyze changes in near-surface temperature on 100-year timescales (century) and on 50-year timescales (50-year), and we analyze changes in zonal mean temperature throughout the atmosphere (free atmosphere). The two near-surface analyses examine changes in time and in space, while the free-atmosphere analysis looks at spatial changes over a 35-year period (section 2.3).

[46] For both the 50-year and the free-atmosphere analysis we use intraensemble variability from the GHG, ANTHRO, and NATURAL ensembles to estimate  $\mathbf{C}_N$ , and we use data from CONTROL to estimate  $\mathbf{C}_{N_2}$ . Any

significant differences between  $\mathbf{C}_N$  and  $\mathbf{C}_{N_2}$  would reduce the power of the optimization algorithm (i.e., increase uncertainty ranges) but would not introduce a bias in the estimated signal amplitudes.

[47] For the century analysis we believe that nine realizations of century timescale variability from the intraensemble variability of HadCM3 are not enough to generate a sufficiently reliable estimate of  $\mathbf{C}_N$ . Therefore we use control and intraensemble variability from five ensembles of HadCM2 (S01) to estimate  $\mathbf{C}_N$ , while  $\mathbf{C}_{N_2}$  is estimated using HadCM3 CONTROL and using intraensemble variability from the GHG, ANTHRO, and NATURAL ensembles. Note that using an incorrect estimate of internal variability reduces the power of our tests but should not bias their results.

#### 4.1. Consistency

[48] We test that the best-estimate combination of signals is consistent with our linear statistical model (equation (1)) by computing the residual sum of squares,

$$R^2 = \sum_{i=1}^{\kappa} \frac{\left(y_i - \sum_{j=1}^n X_{ij}\tilde{\beta}_j\right)^2}{C_{N_{2i}}}, \quad (4)$$

where  $i$  is an index over the ranked eigenvectors of  $\mathbf{C}_N$ ,  $j$  is an index over signals, and  $\kappa$  is the number of eigenvectors used to filter signals and observations (see section 4.3 for details). In the case of noise-free signals,  $R^2$  has a distribution that lies between  $(\chi^2(\kappa - n))/\kappa$  and  $F(\kappa - n, \nu_2)$ , where  $\nu_2$  is the degrees of freedom (DOF) of  $\mathbf{C}_{N_2}$ . We use the F distribution at the 90%, rather than the 95%, level to test for consistency. As an ad hoc correction for noise in the signals, we scale  $R^2$  by  $1/(1 + s)$  and assume that it still has the same distribution, where  $s$  is

$$s = \sum_{i=1}^n (\tilde{\beta}_i/m_i)^2$$

and  $m_i$  is the number of ensemble members in the  $i$ th ensemble. The justification for this ad hoc scaling is that the expected difference between the observations and the best-estimate response would be larger by a factor of  $\sqrt{1 + s}$  due to the noise in the simulations. In the case of signals (and observations) with high signal-to-noise ratio, we verified this scaling by Monte Carlo tests.

#### 4.2. Estimated DOF for Covariance Matrices

[49] In order to compute uncertainties and truncations, we need an estimate of the DOF of the covariance matrices that we compute. These matrices are computed from various different data sets, and their DOF is the sum of the DOF of the individual data sets. For CONTROL the estimated DOF, assuming maximally overlapped data, is the number of nonoverlapping realizations multiplied by 1.5 (Allen and Smith [1996] and S01) and rounded down to the nearest integer. For each ensemble the estimated DOF is the number of nonoverlapping segments in a single simulation multiplied, again, by 1.5, rounded down to the nearest integer, and then multiplied by  $m - 1$  (to account for removal of the mean). The estimated DOF for the two covariance matrices used in our analysis are shown in Table 1. Note that the estimated DOF of  $\tilde{\mathbf{V}}(\tilde{\beta})$  is that of  $\mathbf{C}_{N_2}$ .

**Table 1.** Signal Properties<sup>a</sup>

Case	Period	Trunc.	%Var.	GHG	T-A	ANTHRO	NATURAL	$\nu_1$	$\nu_2$
Surface (Century)	1897–1997	40	96.4	5.85	3.77	3.35	1.17 <sup>b</sup>	40	25
	<i>1897–1997</i>	20	91.3	7.45	4.50	3.93	1.43	40	25
Exchange	<i>1897–1997</i>	25	95.3	6.97	–	4.37	1.30	25	40
	1897–1997	12	93.8	9.23	–	5.73	1.75	25	40
No weight	1897–1997	40	83.2	5.06	–	2.72	1.16 <sup>b</sup>	40	25
	1897–1997	20	65.9	6.17	–	3.01	1.36	40	25
Index	1897–1997	18 <sup>c</sup>	98.0	6.42	–	3.75	1.39	40	25
	<i>1897–1997</i>	9	67.0	7.82	–	3.78	1.43	40	25
90 years	1897–1997	36 <sup>c</sup>	94.9	4.47	–	2.33	1.14 <sup>b</sup>	42	27
	1897–1987	19	91.7	5.72	–	2.77	1.35	42	27
	<i>1907–1997</i>	36 <sup>c</sup>	96.0	5.60	–	3.32	1.01 <sup>b</sup>	42	27
	<i>1907–1997</i>	19	91.7	6.93	–	3.86	1.17 <sup>b</sup>	42	27
Surface (50–year)	1897–1947	27	95.6	1.54	–	1.35	1.37	27	33
	1907–1957	27	95.1	1.92	–	1.29	1.44	27	33
	1917–1967	27	94.1	2.21	–	1.23 <sup>b</sup>	1.39	27	33
	<i>1927–1977</i>	27	91.8	3.01	–	1.70	1.59	27	33
	<i>1937–1987</i>	27	86.3	3.97	2.76	2.34	1.57	27	33
	1947–1997	27	92.9	4.79	4.37	3.66	1.60	27	33
Free Atmosphere	1961–1995	7 <sup>c</sup>	48.0	–	6.11	5.90	0.97 <sup>b</sup>	36	42

<sup>a</sup> Shown for each analysis are the truncation (Trunc.) used and the fraction of the observed variance (after processing) after filtering in the truncated eigenvector space (%Var.). By processing, we mean, for example, projection onto spherical harmonics and weighting by  $\sqrt{(1/2l+1)}$  for the surface analyses and zonal meaning and mass weighting for the free-atmosphere analysis. Italics denote cases where tests for signal degeneracy suggest that the three-signal combination is degenerate. Columns 5–8 show the signal-to-noise ratio (SNR) (see section 4.6 for details) of the simulated signals. TROP-ANTHRO (T-A) is identical to ANTHRO before 1975. Therefore results from TROP-ANTHRO are not shown for those 50-years analyses before 1937–1977 and for the sensitivity analyses. Shown in the two right-hand columns are the estimated degrees of freedom of  $C_N$  ( $\nu_1$ ) and  $C_{N_2}$  ( $\nu_2$ ).

<sup>b</sup> These are SNR values where the value is not significantly different, at the 90% level, from unity (that expected by chance), suggesting significant noise contamination of that simulated signal.

<sup>c</sup> These are cases in which the truncation used is less than the largest possible.

[50] The estimated DOF for the century analysis (see Table 1) may be overoptimistic, as the individual HadCM2 ensemble members were all initialized from the same 1700-year control. Furthermore, the last three simulations of each of the two solar ensembles were initialized by applying small random perturbations to the first solar simulation in each ensemble. Similarly, the three HadCM3 ensembles were all initialized from the same HadCM3 control. Thus the 100-year segments may not be completely independent of one another. Uncertainty in the DOF of  $C_{N_2}$  is relatively unimportant: halving the DOF used in our statistical tests increases the uncertainty ranges by 4%. The estimated DOF of  $C_N$  is used to determine the maximum allowable truncation (see section 4.3), so we explore the sensitivity of our results to truncation.

#### 4.3. Truncation

[51] If  $C_N$  is an order  $n \times n$  matrix, then, where possible, we perform all analysis at the smaller of its DOF and  $n$ . If the consistency test passes, all further analysis is carried out at this truncation ( $\kappa$ ). All data are then filtered by projection onto the leading  $\kappa$  eigenvectors of  $C_N$ . If the test fails at this truncation, then we carry out the analysis at the largest truncation at which the test passes and explore the reasons for the test failure. Our estimated DOF are somewhat arbitrary, as are the criteria we use to determine truncation. Therefore we explore the sensitivity of our results to truncation.

#### 4.4. Degeneracy

[52] We used the same three empirical tests as T99 and S01 to test for signal degeneracy or collinearity [see *Mardia*

*et al.*, 1975, pp. 243–248]. These three tests give empirical estimates of the number of independent “factors” in the signal combination. We conclude that a signal combination is likely to be degenerate if the maximum value of those three tests is less than the number of signals being considered.

[53] If two signals are degenerate, the usual consequence is that uncertainty ranges are large. Then the best-estimate amplitudes are not likely to be close to the true ones. It is also likely that neither signal is individually detectable, since a range of linear combinations are equally consistent with the data, including those that assign zero amplitude to one signal or the other. However, specific combinations of these signals may easily be detectable and may have smaller uncertainty ranges.

#### 4.5. Transformations

[54] We assume that the three anthropogenic signals, GHG, ANTHRO, and TROP-ANTHRO, are linear combinations of the following physically based signals: (1)  $G$ , response to well-mixed greenhouse gases alone; (2)  $O_T$ , response to tropospheric ozone changes; (3)  $O_S$ , response to stratospheric ozone decline; (4)  $O$ , response to both stratospheric and tropospheric ozone changes; and (5)  $S$ , response to sulphates (indirect and direct), namely,

$$\begin{aligned} \text{GHG} &= G, \\ \text{ANTHRO} &= G + S + O = GSO, \\ \text{TROP-ANTHRO} &= G + S + O_T = GSO_T. \end{aligned}$$

[55] The amplitudes and covariance matrices of these physically based signals are given by a linear transformation of the original amplitudes and of  $\tilde{V}(\hat{\beta})$ ; see Appendix B for

details. For example, suppose we model the observations as a linear superposition of the GHG and ANTHRO simulations,

$$\mathbf{y} = \mathbf{x}_{\text{GHG}}\tilde{\beta}_{\text{GHG}} + \mathbf{x}_{\text{ANTHRO}}\tilde{\beta}_{\text{ANTHRO}}.$$

$\tilde{\beta}_{\text{GHG}}$  in this equation is not simply the estimated amplitude of the greenhouse response. It is the additional greenhouse response we need to add to the best fit ANTHRO simulation to obtain the best overall fit to the observations. In this case the amplitude of the greenhouse and “other anthropogenic” signals is

$$\begin{aligned}\tilde{\beta}_G &= \tilde{\beta}_{\text{ANTHRO}} + \tilde{\beta}_{\text{GHG}} \\ \tilde{\beta}_{\text{SO}} &= \tilde{\beta}_{\text{ANTHRO}}.\end{aligned}$$

In this example the variance in  $\tilde{\beta}_G$  is equal to the sum of the variances in  $\tilde{\beta}_{\text{GHG}}$  and  $\tilde{\beta}_{\text{ANTHRO}}$ .

#### 4.6. Signal-to-Noise

[56] Amplitude uncertainty ranges, and particularly the upper bound, estimated from signals with a low signal-to-noise ratio (SNR) are likely to be incorrect [Allen and Stott, 2002]. We use the following summary statistic for the  $j$ th signal to give us some guidance when this may be occurring:

$$(\text{SNR})^2 = \frac{m_j}{\kappa} \sum_{i=1}^{\kappa} \frac{X_{ij}^2}{C_{N_{2ii}}},$$

where  $\kappa$  is the truncation. When the “signal”  $\mathbf{x}_j$  is pure Gaussian noise,  $(\text{SNR})^2$  has an expected value of 1 and is distributed similarly to  $R^2$  from section 4.1 (between  $(\chi^2(\kappa))/\kappa$  and  $F(\kappa, \nu_2)$ ). We use an F test at the 90% level to determine if there is significant noise contamination.

## 5. Detection and Attribution of Observed Temperature Changes

### 5.1. Changes in Near-Surface Temperature on Century Timescales

[57] We now examine changes from 1897 to 1997 using information about both temporal and spatial changes in near-surface temperature. We make two further simplifying steps. First, since TROP-ANTHRO and ANTHRO are identical for all but the last 2 decades of the twentieth century (when the difference in surface temperature response is small), we use only ANTHRO in this analysis. Second, we make a simple linear transformation of the amplitudes of GHG and ANTHRO to obtain amplitudes of  $G$  (greenhouse gases) and  $SO$  (sulphates and ozone) as described in section 4 and Appendix B. Tests for degeneracy (section 4.4) suggest that these patterns are different enough that we can reliably estimate the amplitude of  $G$ ,  $SO$ , and NATURAL (response to solar irradiance and volcanic aerosols) signals simultaneously (Table 1).

[58] We first check that the reduced space in which we carry out the detection procedure provides an adequate representation of the observed changes. When filtered onto the leading eigenvectors of the covariance matrix (see

section 4), the observations contain >96% of the variance (Table 1). The best-estimate linear combination of the signals is consistent with the observations as shown by the weighted sum-of-squares of the residuals (Figure 7a and section 4.1) at all truncations. Thus the representation is adequate.

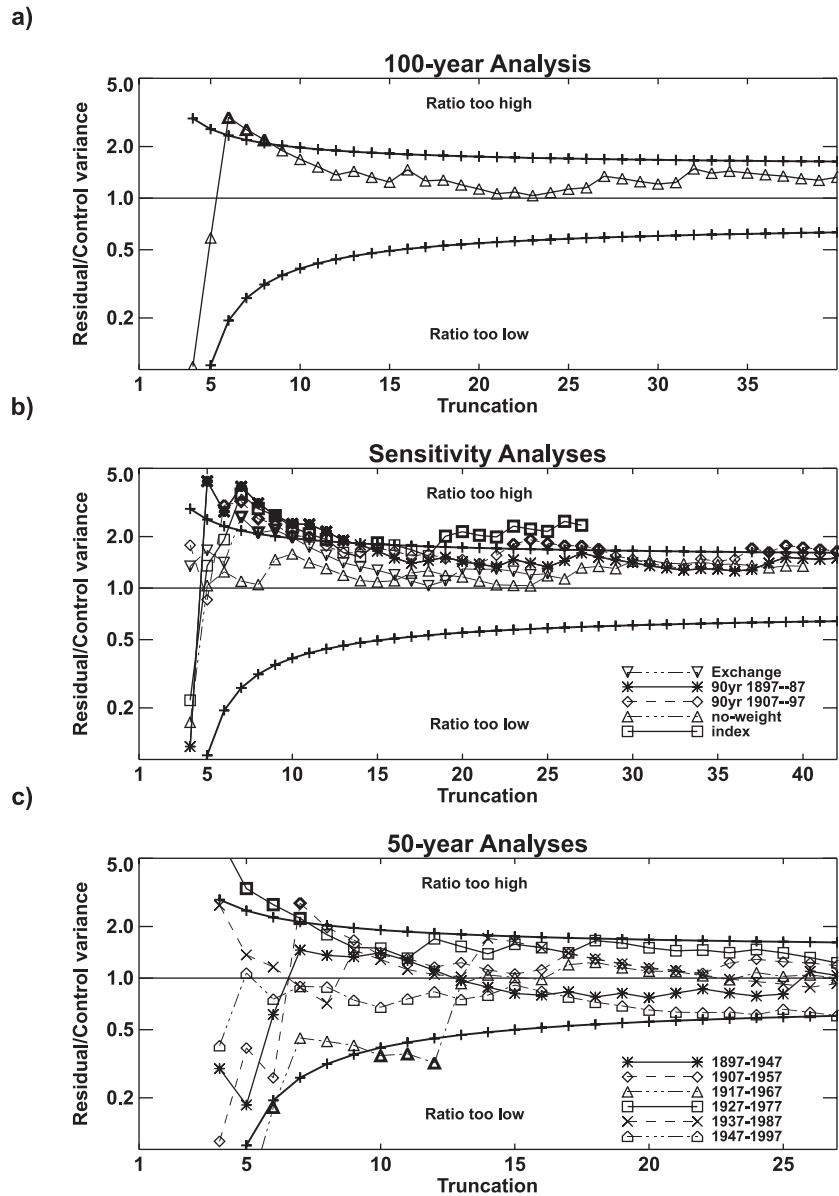
[59] All three signals are detected (Figure 8, left), demonstrating that all have had a statistically significant impact on changes in near-surface temperature over the twentieth century. Furthermore, the amplitudes are all consistent with unity; the model is consistent with observations on decadal timescales and on continental to global spatial scales.

[60] Signal-to-noise ratio is large for the anthropogenic signals but is small for NATURAL (first line of Table 1), suggesting that it is significantly noise contaminated. Noise contamination of the signals biases the best estimate toward zero. Hence our detection of NATURAL is probably robust, though its estimated amplitude ranges, and in particular the upper range, are sensitive to this noise contamination [Allen and Stott, 2002].

[61] We reconstruct the global mean temperature changes from the best-estimate signal amplitudes and simulated responses (Figure 9). From the 1900s to the 1960s, well-mixed greenhouse gases and other anthropogenic effects (largely the indirect effect of sulphate aerosols) almost balance, giving a total anthropogenic warming of  $\sim 0.1$  K. Thereafter, anthropogenic effects warm the planet by  $\sim 0.5$  K. From the 1950s onward, natural and anthropogenic non-greenhouse gas forcings each cause a cooling of  $\sim 0.1$  K. Together, they offset  $\sim 0.2$  K of the estimated 0.6 K warming due to greenhouse gases over the same period.

[62] While Figure 9 shows the best-estimate combination of signals, it is even more important to consider uncertainty ranges. These are most easily summarized in terms of linear trends (Figure 10) over selected periods (the entire century, 1897–1947, and 1947–1997). The uncertainty ranges in the trends were computed by taking the amplitude ranges from the century analysis and applying them to the simulated trends over the three periods; see section 4 for details. Over the twentieth century, anthropogenic forcings cause a warming trend of  $0.5 \pm 0.15$  K/century. The trend due to greenhouse gases is  $0.9 \pm 0.24$  K/century, while the remaining anthropogenic factors cool at a rate of  $0.4 \pm 0.26$  K/century. The uncertainty in the total anthropogenic warming trend is less than the uncertainties in the individual trends, as they are correlated with one another; see below. Over the century, natural forcings contribute little to the observed trend. Our analysis considers only uncertainty in the amplitude of the simulated response and neglects uncertainty in the time dependence of the forcing and in the spatial patterns of response, as well as neglecting uncertainties in the observations. However, our best estimates are consistent with the observations. Furthermore, in a single ensemble of simulations forced with both natural and anthropogenic forcings, changes in simulated near-surface temperature are consistent with those observed [Stott *et al.*, 2000], suggesting that those uncertainties may not be too great.

[63] During the first half of the century, greenhouse gases and natural forcings cause warming trends of  $\sim 0.2$ – $0.3$  K/century, while other anthropogenic factors produce negligible cooling trends (Figure 10). Over the last half of the

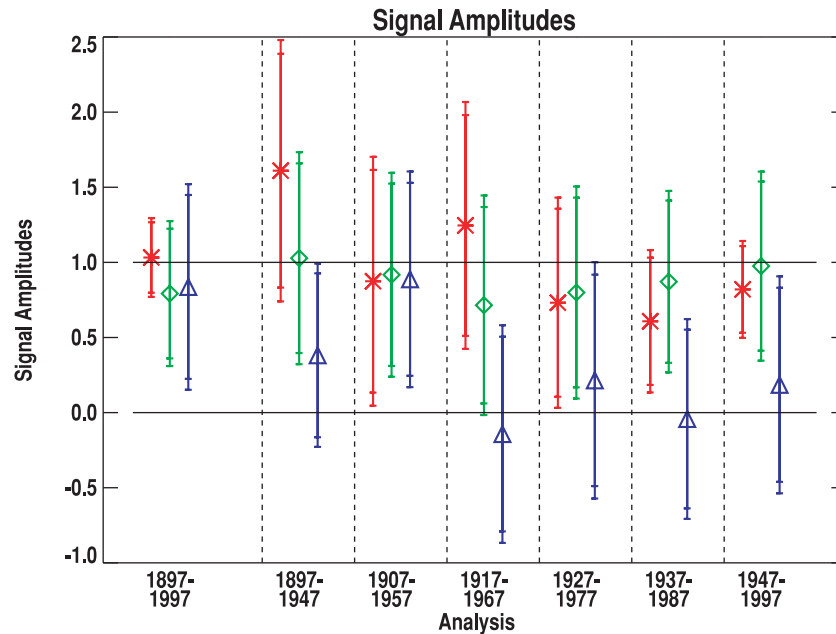


**Figure 7.** Ratio of the residual to control variance using a logarithmic scale (solid line with triangles) for (a) the century analysis, (b) five sensitivity studies, and (c) all six 50-year analyses of surface temperature. Also shown are the 10–90% values of the ratio under the null hypothesis that CONTROL and residual variances are the same (solid lines with pluses). Note that CONTROL variance has been inflated (see section 4.1 for details). Bold symbols show values outside these limits.

century, greenhouse gases warm the climate at a rate of  $1.7 \pm 0.43$  K/century, with natural forcings (largely volcanic aerosol) and other anthropogenic factors (mainly the indirect effect of sulphate aerosols) both causing an estimated cooling trend of  $\sim 0.3 \pm 0.2$  K/century. Thus, since 1947, changes in aerosol concentrations (anthropogenic and natural) have offset about a third of the greenhouse gas warming.

[64] Uncertainties in signal amplitudes are correlated, as the signals are not orthogonal. The joint confidence regions allow us to examine these correlated uncertainties. We find that all three simulated signals are consistent with the observations (the signal amplitudes are simultaneously consistent with unity; that is, the point (1,1,1) is within

the three-dimensional uncertainty ellipsoid), as are any combination of two signals (that is, all the solid ellipses in Figure 11 include the point (1,1)). The uncertainty ellipse for the two anthropogenic signals has a strong tilt, showing that the amplitudes of these signals are highly correlated (Figure 11a). Thus estimated large amplitudes of  $G$  are consistent with estimated large amplitudes of  $SO$ ; that is, the observations require a larger greenhouse gas warming to accompany a stronger cooling from sulphates. Over the century, there is little tilt between the natural and either of the anthropogenic signals (Figures 11b and 11c). Thus errors in the amplitude of the natural signal have little impact on the estimated amplitude of the two anthropogenic signals. Consequently, the uncertainties in the linear trends



**Figure 8.** Amplitudes and uncertainty ranges for 100-year analysis and for all 50-year analyses for *G* (red error bar with asterisk), *SO* (green error bar with diamond), and NATURAL (blue error bar with triangle). Error bars show the 5–95% uncertainty ranges for detection (inner) and amplitude consistency (outer). Best-estimate signal amplitude is shown as a symbol at the center of the bar. Where the inner bars do not include zero (lower line), the signal is detected. Where the outer bars do not cross 1 (upper line), the amplitude of the simulated signal is inconsistent with the observations.

(Figure 10) due to NATURAL are independent from those due to *SO* and *G*.

[65] One “technical” issue in optimal detection is the eigenvector truncation used (see section 4.3). Our results are insensitive to truncation for both detection (the shaded inner regions in Figure 12 (top row) do not include zero) and “amplitude consistency” (the shading plus thick lines in Figure 12 include 1).

[66] If we omit the effect of stratospheric ozone decline, by replacing ANTHRO with TROP-ANTHRO, we find little change in the residuals and find only small changes in the amplitudes. There is a slight reduction in the cooling attributed to sulphate aerosols and tropospheric ozone from the 1950s, which is compensated for by a small increase in naturally forced cooling (not shown).

## 5.2. Sensitivity to Processing and Variability Estimates

[67] In this section we explore the sensitivity of results from the previous analysis to details of the data processing and to increases in the magnitude of the simulated climate variability. We consider the following cases: (1) No-weight, where we did not apply the weighting of  $1/\sqrt{2l+1}$  to the spherical harmonics (see section 4 for details); (2) Exchange, in which we used  $C_{N_2}$  to optimize and used  $C_N$  to compute uncertainties; that is, we used HadCM3 data to optimize and used HadCM2 data to compute uncertainties; (3) Index, where, rather than projecting simulated and observed data onto spherical harmonics, three indices were computed: the global average, the land temperature, and the Northern Hemisphere minus the Southern Hemisphere; and (4) 90-year, where, rather than doing the analysis for the century, we carried out the

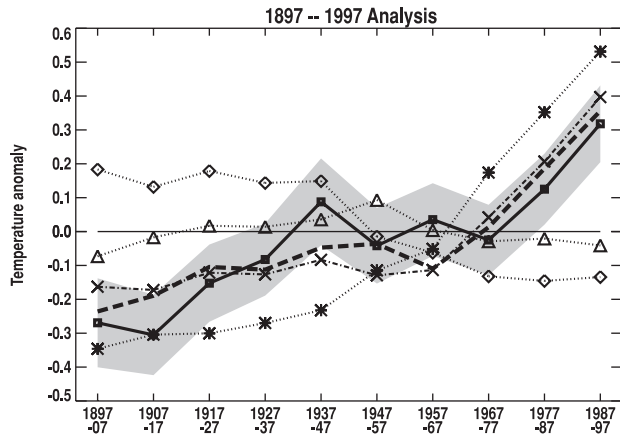
analysis on two 90-year segments (1897–1987 and 1907–1997).

[68] In the 90-year 1907–1997 and Index sensitivity studies we find that the simulations and observations are inconsistent (section 4.1) at the largest truncations that we consider (Figure 7b). Therefore we truncate at the largest truncations that are consistent with the observations (Table 1). We carry out both 90-year analyses at the truncation determined by the 1907–1997 case.

[69] We repeat these analyses and the century case at half the largest truncation to see if our results are insensitive to truncation. Thus, including the “normal” data processing at truncation 20, we examine a total of 11 sensitivity studies, giving 12 cases in all. At these truncations the filtered observations contain at least 80% of the observed variance (Table 1), except in two cases.

[70] The SNR for the anthropogenic signals is always  $>2$ , suggesting little noise contamination (Table 1). By contrast, SNR for NATURAL is close to 1 and in five cases is not significantly different from that expected by chance. There is evidence of signal degeneracy (see section 4.4) in five cases (Table 2), meaning that results in those cases may be sensitive to small changes in the signals. We find that (1) *G* is detected in all cases (Table 2); (2) *SO* is detected in all but two cases, both at half the maximum truncation; and (3) NATURAL is detected in all but two cases.

[71] All amplitudes are consistent with unity (not shown). Therefore we conclude that our detection and amplitude consistency results are robust to changes in both processing and truncation. The best-estimate amplitudes are, however, more sensitive to these choices, with  $\beta_{\text{NATURAL}}$  extending



**Figure 9.** Reconstruction of global mean temperature variations for 1897–1997. Observations (solid line with squares), best-estimate changes (thick dashed line), and best-estimate contributions from  $G$  (dotted line with asterisks),  $SO$  (dotted line with diamonds), and NATURAL (dotted line with triangles). Also shown is the best-estimate total anthropogenic contribution (dot-dashed line with crosses). All time series were reconstructed from data in which the 100-year mean had first been removed. Shaded region centered on the observations shows the uncertainty range due to internal variability (two-sigma decadal variability computed from  $C_{N_2}$ ).

from 0.60 to 1.11,  $\tilde{\beta}_G$  varying from 0.77 to 1.10, and  $\tilde{\beta}_{SO}$  ranging from 0.49 to 0.91 (Table 2).

[72] Our claims of signal detection all rely on simulated internal climate variability. We compute how much the model variability needs to be increased to prevent the detection of the signals in all the cases considered above. The amplitude of the simulated variability needs to be inflated by 2.2–4.8 to nullify our detection of greenhouse gases (Table 3) (an increase in variance of 5–23). However, detection of the  $SO$  and NATURAL signals is much less robust. Here, an increase in variability by  $\sim 40\%$  (i.e., doubling the variance) is enough to stop detection of  $SO$  and NATURAL in half the cases considered.

### 5.3. Surface Temperature Changes on 50-Year Timescales

[73] We now examine changes on 50-year timescales to allow comparison with the HadCM2 results of T99 and S01. Six 50-year periods, each of five decadal means, are considered: 1897–1947, 1907–1957, ..., 1947–1997. At least 85% of the observed variance (Table 1) is captured in these periods. Unlike the century analysis, NATURAL is generally not significantly noise contaminated (Table 1), though the signal-to-noise ratio is below 1.5 for NATURAL and ANTHRO before 1937. NATURAL may be less noise contaminated than in the century analysis because the truncation is smaller as more noisy components of the signal are discarded. (Compare the SNR of the century analysis NATURAL signal at truncation 20 with that at truncation 40 in Table 1). We use the same signal combination ( $G$ ,  $SO$ , and NATURAL) as used in the century analysis and find evidence of signal degeneracy during 1927–1977 and during 1937–1987 (Table 1). The residuals

are consistent with the variance computed from CONTROL at almost all truncations and for all periods (Figure 7c).

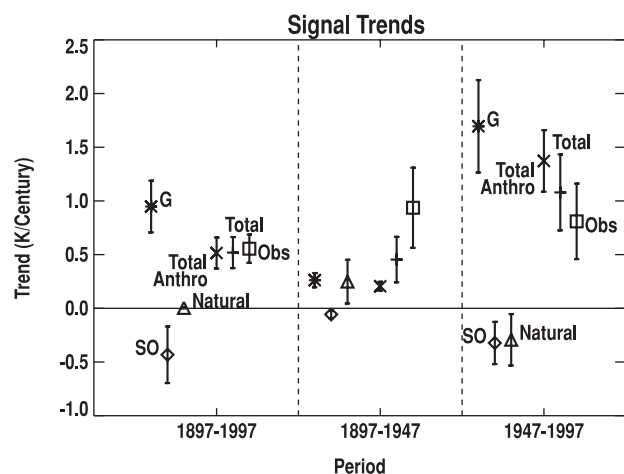
[74] Both of the anthropogenic signals ( $G$  and  $SO$ ) are detected in all six 50-year periods, with amplitudes consistent with unity (Figure 8). Natural effects on climate are only detected during the 1907–1957 period (Figure 8), whereas the amplitudes are consistent with unity only in 1897–1947, 1907–1957, and 1927–1977.

[75] We wish to compare our results with T99 and S01, include the period when NATURAL is detected, and also examine both periods of warming during the twentieth century. Thus we consider in more detail the 1907–1957 and 1947–1997 periods.

[76] We first of all consider how robust our results are to changes in truncation. Detection of both the anthropogenic signals during these periods, unlike the natural signal, is largely robust to truncation (Figure 12). All signal amplitudes are consistent with unity, except during 1947–1997 for NATURAL at all truncations and for  $G$  for truncations below 13.

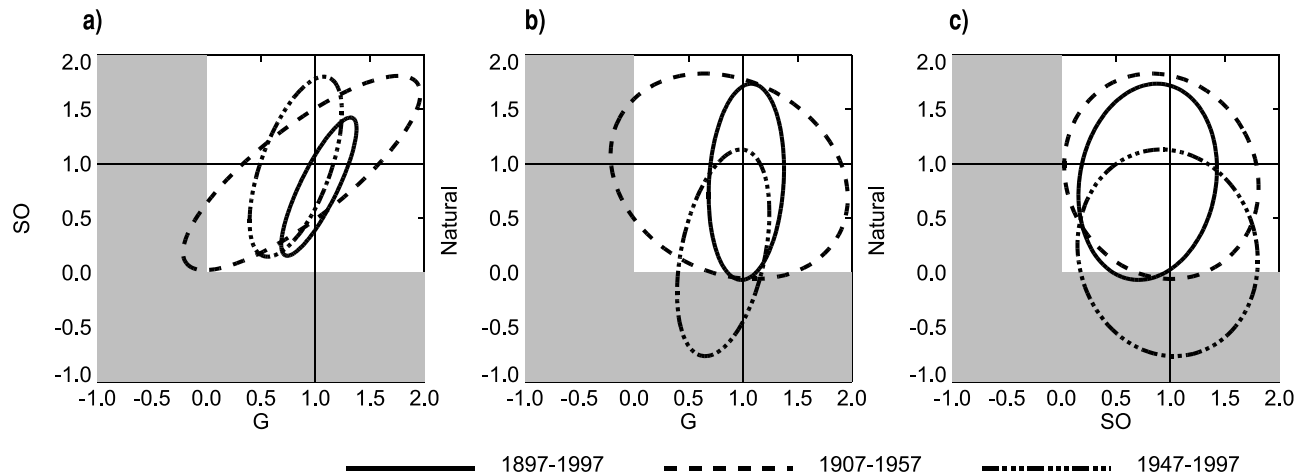
[77] Best-estimate global mean temperature changes and trends (section 4) are proportional to the amplitudes shown in Figure 8. Thus we can compare best-estimate changes and trends from the 50-year and century analyses by comparing their amplitudes. The 50-year analyses produce smaller natural changes than the century analysis, except in the 1907–1957 period. Cooling from sulphates and ozone is about the same in both cases, while greenhouse gas warming is less in the 50-year analyses from 1927 onward (Figure 8). Thus total anthropogenic changes and trends are generally smaller in the 50-year analyses than in the century analysis.

[78] Only in the 1907–1957 analysis do natural forcings make a substantial contribution to temperature trends (Figure 13). In this period the temperature trend due to anthropogenic forcings is close to zero. In this period, amplitudes (Figure 8), and hence temperature trends (Figure 13),



**Figure 10.** Best-estimate linear trend and uncertainty ranges (in K/century). Symbols are as in Figure 8, with the addition of total anthropogenic trend (crosses), total trend (pluses), and observed trends (squares). Symbols show best-estimate trend, while error bars show the 5–95% uncertainty range inflated to allow for four-member ensembles.



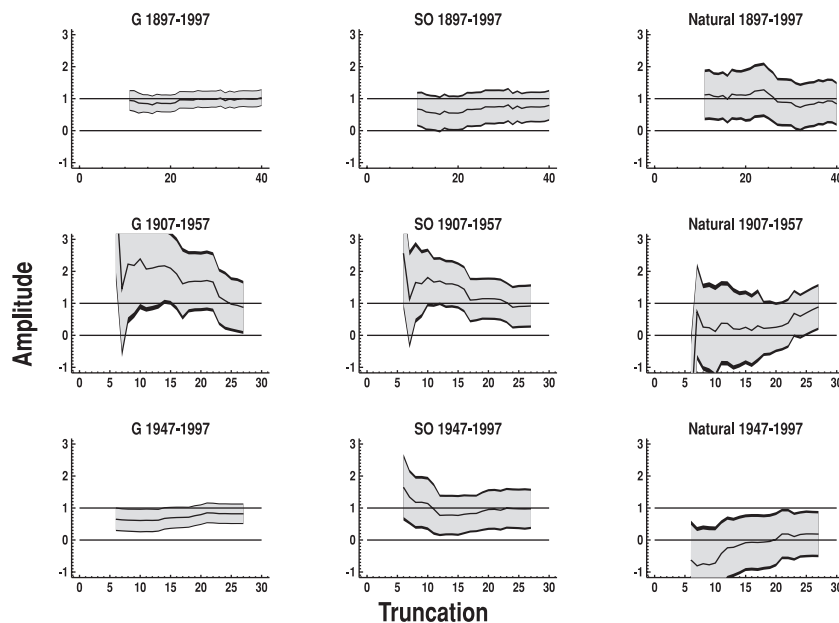


**Figure 11.** The 90% joint confidence regions for the 1897–1997 (solid line), 1907–1957 (dashed line) and 1947–1997 (dot-dashed line) from *G SO NATURAL* analysis. Shown are the two-dimensional confidence regions for (a) *G SO*, (b) *G NATURAL*, and (c) *SO NATURAL*. In all plots, shading denotes regions where signal amplitudes are less than zero.

of all three signals are also very similar to the century analysis. Note that the trends and uncertainties shown in Figure 10 are computed from scaling factors using the century analysis (section 5.1), while those shown in Figure 13 are computed using a 50-year analysis.

[79] In the 1907–1957 analysis the difference between the best estimate and the observed trend (residual) is the largest of all the periods considered (Figure 13). The residual is still consistent with our estimated internal climate variability and so could be due to internal climate variability alone. It could also, partly or wholly, be due to observational error, error in the forcing time series, some other forcing not considered in our analyses, model error, or noise

in the signals. T99 found a large residual in their GS SOL analysis (see T99, Figure 2b) in the 1906–1956 period, suggesting that this result is robust to using the solar time series of *Hoyt and Schatten* [1993], neglecting the effect of volcanos and the use of a different model. *Hegerl et al.* [2001] found that observational error was much smaller than internal variability. This suggests that the large residual is probably due to internal climate variability. *Delworth and Knutson* [2000] found that one simulation from an ensemble of anthropogenically forced simulations was similar to the observations of twentieth century near-surface temperature change. However, like us, they found that the ensemble was inconsistent with the observed changes in the early century.



**Figure 12.** Best-estimate amplitudes (thin line), 5–95% detection uncertainties (shading) and 5–95% amplitude consistency uncertainties (shading plus thick lines) are shown for (top row) the 1897–1997 analysis and two 50-year analyses, (middle row) 1907–1957 and (bottom row) 1947–1997 for (left column) *G*, (center column) *SO*, and (right column) *NATURAL* as a function of truncation.

**Table 2.** Sensitivity Studies<sup>a</sup>

Case	Truncation	Signal Amplitude		
		<i>G</i>	<i>SO</i>	NATURAL
Century	40	1.03 <sup>b</sup>	0.79 <sup>b</sup>	0.84 <sup>b</sup>
Century	20	0.85 <sup>b</sup>	0.55 <sup>b</sup>	1.11 <sup>b</sup>
<i>Exchange</i>	25	1.01 <sup>b</sup>	0.87 <sup>b</sup>	0.62 <sup>b</sup>
<i>Exchange</i>	12	1.10 <sup>b</sup>	0.91 <sup>b</sup>	0.91 <sup>b</sup>
No weight	40	0.96 <sup>b</sup>	0.67 <sup>b</sup>	0.87 <sup>b</sup>
No weight	20	0.79 <sup>b</sup>	0.49	1.03 <sup>b</sup>
Index	18	0.96 <sup>b</sup>	0.83 <sup>b</sup>	0.60
<i>Index</i>	9	0.77 <sup>b</sup>	0.58 <sup>b</sup>	0.70
90 years, 1897–1987	36	0.95 <sup>b</sup>	0.71 <sup>b</sup>	1.05 <sup>b</sup>
90 years, 1897–1987	19	0.79 <sup>b</sup>	0.50	1.07 <sup>b</sup>
<i>90 years, 1907–1997</i>	36	1.03 <sup>b</sup>	0.89 <sup>b</sup>	0.70 <sup>b</sup>
<i>90 years, 1907–1997</i>	19	0.86 <sup>b</sup>	0.63 <sup>b</sup>	0.97 <sup>b</sup>

<sup>a</sup> Best-estimate signal amplitudes for the base analysis (century) and sensitivity studies are shown. Italics denote cases where tests for signal degeneracy suggest that the three-signal combination is degenerate. The degrees of freedom used in the tests are given in Table 1.

<sup>b</sup> These are detectable signals.

[80] Thus the 1907–1957 warming is best explained by a combination of natural forcings (an increase in solar irradiance, a lack of large volcanic forcing, and a recovery from earlier volcanic forcing), near-zero response to total anthropogenic forcing, and a large warming from internal climate variability. If correct, this suggests that a large part of the early century warming is due to a combination of natural forcing and natural internal variability. In other words, it is naturally caused. In our simulations, sulphates offset most of the greenhouse warming prior to the 1960s. If this were not the case, then we would be likely to have smaller residuals and thus estimated a smaller contribution from internal climate variability to the early century warming.

[81] The model is consistent with the observations in the two 50-year periods that we have chosen to focus on (1907–1957 and 1947–1997). In these periods the uncertainty ellipses for the two anthropogenic amplitudes are strongly tilted, showing that their amplitudes are highly correlated (Figure 11a). Amplitudes of the natural and anthropogenic signals are less correlated (Figures 11b and 11c). In 1947–1997 the tilt is such that a larger amplitude of

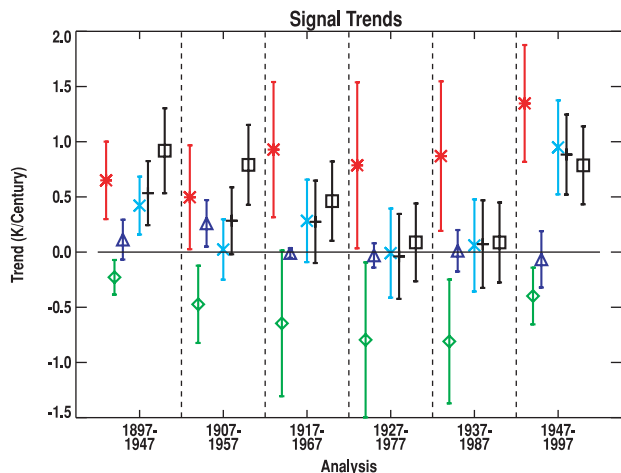
the *G* signal requires a larger amplitude of the NATURAL signal, whereas the ellipse is weakly tilted in the opposite direction in the 1907–1957 period. The natural and anthropogenic amplitudes are less correlated in the century analysis than in either of the 50-year analyses. Therefore the former analysis is better at discriminating between natural and anthropogenic forcings than is the latter. All three signals are simultaneously consistent with the observations (the point (1,1,1) is inside the three-dimensional ellipsoid centered on  $\tilde{\beta}$ ) in all periods except 1917–1967 (not shown).

[82] We can compare our results with those of T99 and S01, though our experimental design differs from theirs. For example, we included the effects of ozone, while they did not. Unlike T99 and S01, we detect anthropogenic influences in all 50-year periods considered. Our detection of a combined solar and volcanic effect on climate during 1907–1957 corresponds to their detection of a solar influence during 1906–1956. There are differences in the warming during this period (compare our Figure 14a with Figure 1b of T99), some of which may be due to use of the

**Table 3.** Ratio of Signal Amplitudes to Uncertainty Range<sup>a</sup>

Case	Truncation	Amplitude/Uncertainty		
		<i>G</i>	<i>SO</i>	NATURAL
Century	40	4.40	1.84	1.37
Century	20	3.40	1.11	1.49
<i>Exchange</i>	25	4.77	2.16	1.29
<i>Exchange</i>	12	4.80	2.07	1.47
No weight	40	3.61	1.44	1.52
No weight	20	2.73	0.98	1.45
Index	18	3.84	1.82	0.95
<i>Index</i>	9	2.26	1.19	0.78
90 years, 1897–1987	36	3.07	1.54	1.93
90 years, 1897–1987	19	2.47	0.91	1.51
<i>90 years, 1907–1997</i>	36	4.22	1.93	1.04
<i>90 years, 1907–1997</i>	19	3.28	1.21	1.05

<sup>a</sup> Shown, for the sensitivity studies and base century analysis, are the ratio of the best-estimate signal amplitudes to half the uncertainty range. Inflating the simulated variability by this factor (scaling  $C_{N_2}$  by its squared) makes the signal amplitude consistent with zero at the 5% level. Where the factor is greater than unity, it is the minimum amount needed to inflate the simulated variability so that the signal is no longer detected. Italics denote cases where tests for signal degeneracy suggest that the three-signal combination is degenerate.



**Figure 13.** Best-estimate linear trend and uncertainty ranges (in K/century) for 50-year timescale analysis. Colors and symbols are as in Figure 8, with the addition of total anthropogenic trend (pale blue crosses), total trend (black pluses), and observed trends (black squares).

solar forcing of *Lean et al.* [1995a] rather than that of *Hoyt and Schatten* [1993]. Our total anthropogenic changes for 1947–1997 are similar to those of T99, but with less sulphate cooling and less greenhouse warming than T99; compare our Figure 14b with Figure 1c of T99.

[83] The linear trend and uncertainty range for each signal are comparable with those computed by T99 (compare Figure 2 of T99 with our Figure 13). As in T99, the total anthropogenic warming trend is only greatly different from zero in the 1947–1997 period. There are greater greenhouse warming and sulphate cooling trends in our analysis than in T99 (compare our Figure 13 with Figure 2a of T99) in all but the 1947–1997 period. Thus, while the total anthropogenic warming estimated here (using HadCM3) is similar to

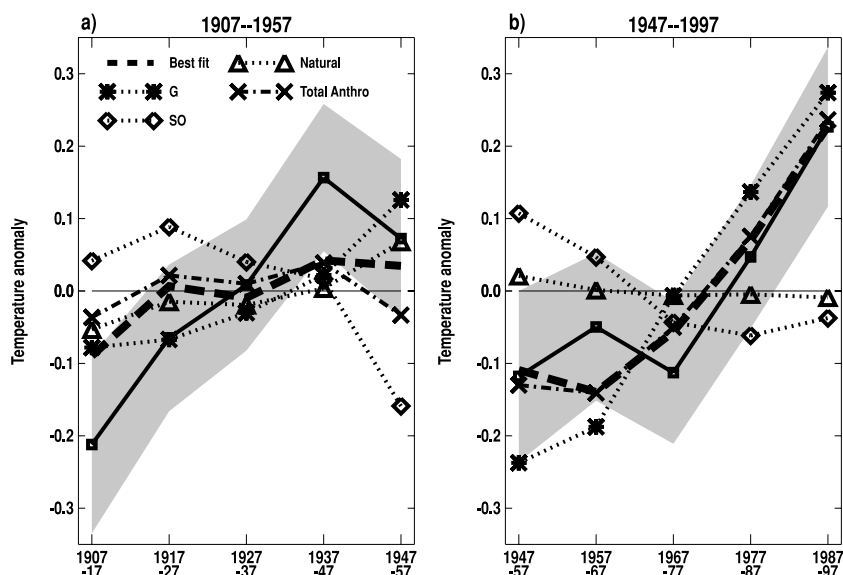
that of T99 and S01 (using HadCM2), the partitioning into warming from greenhouse gases and cooling from other anthropogenic forcings is different.

[84] Finally, as in the earlier century analyses, we omit the effect of stratospheric ozone decline and repeat our analysis. We find that the residuals are similar, except during 1947–1997 when the fit to observations is too good for truncations greater than 17, suggesting that the model may have too much internal variability. Though the same signals are detected, the amplitude of G is significantly smaller than unity in the 1937–1987 and 1947–1997 analyses, meaning that the simulated response is significantly too large. We also find that anthropogenic aerosols and tropospheric ozone offset less greenhouse warming in 1947–1997 than in our original 50-year analysis. Since the near-surface temperature responses in ANTHRO and TROP-ANTHRO are similar, then some of our results may be sensitive to relatively small amounts of noise in the signals. Alternatively, they may be sensitive to the highly uncertain ozone forcing.

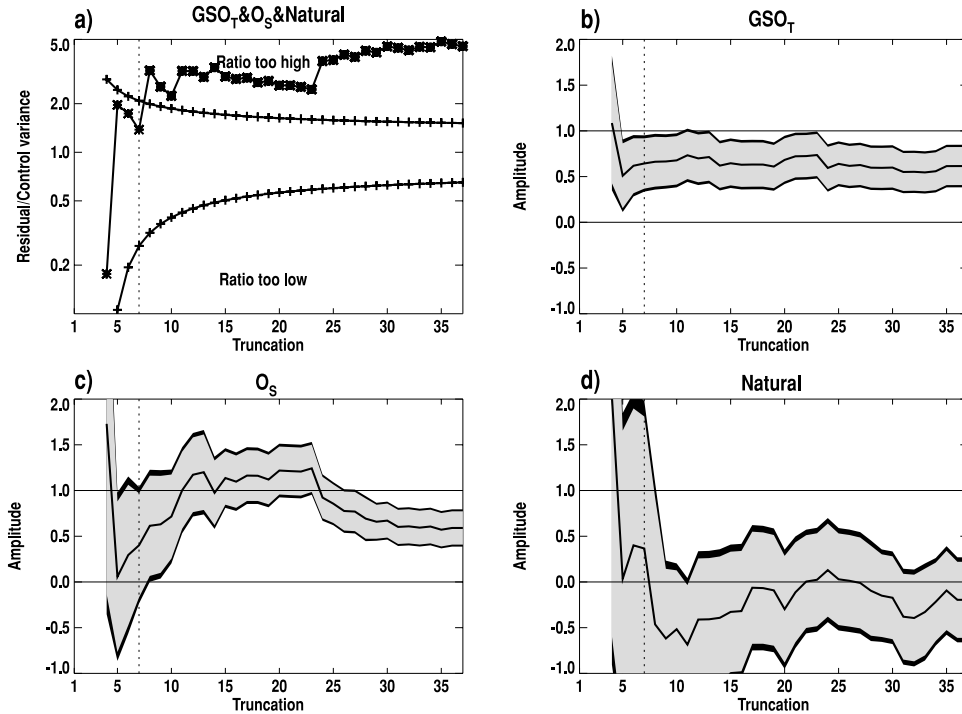
### 5.4. Free-Atmosphere Changes

[85] Several earlier detection and attribution studies have focused on the changes in the zonally averaged temperature of the free atmosphere [*Santer et al.*, 1996b, 1996a; *Tett et al.*, 1996; AT99]. Although it turns out that the stratospheric changes are not particularly well represented in this study by the truncated eigenvectors of the appropriate covariance matrix, we have included this analysis to show how the new simulations compare with earlier work. In particular, we wish to see if the conclusions in the earlier study still hold when the response to natural forcings is taken into account. We examine the difference between the 10-year zonal mean from 1986–1995 and the 20-year zonal mean for 1961–1980, as given by AT99.

[86] Earlier, we showed that the changes in the free atmosphere simulated by TROP-ANTHRO and GHG are



**Figure 14.** Best-estimate reconstruction of temperature variations for (a) 1907–1957 and (b) 1947–1997. Symbols are as in Figure 9, but are reconstructed from data from which the 50-year mean had first been removed.



**Figure 15.** Sensitivity to truncation for free atmosphere analysis. (a) Ratio of the residual to the CONTROL variance (solid line with asterisks), using a logarithmic scale. Other details of plot are as Figure 7. Note that CONTROL variance has been inflated (see section 4 for details). Vertical dotted line shows truncation 7, the largest truncation for which the residual and CONTROL variance are consistent. Shown as a function of truncation are the best-estimate amplitudes (thin lines), 5–95% detection uncertainties (shading), and 5–95% amplitude consistency uncertainties (shading plus thick lines) for (b)  $GSO_{T_7}$ , (c)  $O_S$ , and (d) NATURAL.

similar. We therefore do not use GHG in this analysis, examining combinations of TROP-ANTHRO, ANTHRO, and NATURAL. This assumes that the relative amplitudes of the  $G$  and  $SO_T$  responses are as in TROP-ANTHRO. To separate the impact of stratospheric ozone decline from all other anthropogenic effects, we transform the amplitude of the TROP-ANTHRO and ANTHRO signals to give amplitudes of  $GSO_T$  (all anthropogenic forcings except stratospheric ozone decline) and  $O_S$  (stratospheric ozone decline on climate); see section 4.5 and Appendix B for details.

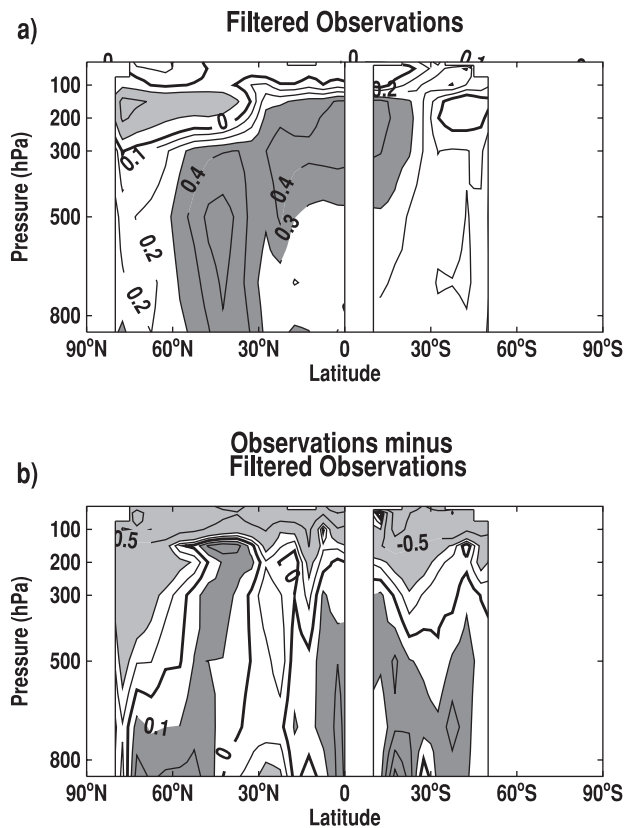
[87] In the three-signal case the maximum truncation of  $C_N$  is seven. For truncations beyond this, the ratio of the residual to control variance is 3–5 times too large (Figure 15a). At truncation 7 the filtered observations contain 48% of the observed mass-weighted variance (Table 11) compared to 71% at truncation 36 (the truncation we believe is the largest we could reasonably consider, given the estimated DOF of  $C_N$ ; see Table 1). The SNR for the two anthropogenic signals is reasonably high (Table 1), while the SNR for the natural signal is  $<1$ .

[88] The  $GSO_T$ ,  $O_S$  and NATURAL case has residual variance consistent with CONTROL for all truncations less than or equal to seven (Figure 15a). At these truncations,  $O_S$  and NATURAL are consistent with unity and zero; that is, they are not detected, but the simulated amplitudes could be correct (Figures 15c and 15d).  $GSO_T$  is detected but is inconsistent with unity (Figure 15b). Its best-

estimate value is 0.65, suggesting that the simulated tropospheric response is  $\sim 50\%$  stronger than the observed response.

[89] Our failure to detect NATURAL does not rule out the possibility of a statistically significant natural influence on climate, because the simulated signal is noise contaminated and so could be substantially in error. Furthermore, there remains the possibility that natural effects may have an influence on shorter timescales, for example, the stratospheric warming associated with volcanoes and possible links between changes in the upper tropospheric circulation and the solar cycle [e.g., Salby and Callaghan, 2000; Hill et al., 2001].

[90] Above truncation 7 the residual variance is  $\sim 3$ –5 times larger than that of CONTROL (Figure 15a), and we now consider why this might be. The observations filtered by these leading seven eigenvectors do capture the gross features of the tropospheric warming (Figure 16a). However, at this truncation the filtered observations do not show the observed stratospheric cooling (Figure 2b) as seen more clearly in the difference between the raw and the filtered observations (Figure 16b). The raw observations are cooler in the stratosphere and are  $\sim 0.1$  K warmer throughout large regions of the troposphere than are the filtered observations. Therefore our failure at truncations greater than seven is probably due to the simulated stratospheric variability being too small, though gross signal error cannot be ruled out. At truncation 7 the best-estimate



**Figure 16.** (a) Observed changes in zonal mean temperature filtered by projection onto the leading seven eigenvectors of  $C_N$ . A contour interval of 0.1 K is used between 0 K (bold contour) and 0.6 K, with additional contours at  $-0.2$ ,  $-0.5$ , and  $-1$  K and with dark shading above 0.3 K and light shading below  $-0.2$  K. (b) Raw observations minus data from Figure 16a (i.e., what the filtering removes). A contour interval of 0.1 K is used between  $-0.2$  and 0.6 K, with additional contours at  $-0.5$  and 1 K. The zero contour is bold, with light shading for differences below  $-0.2$  K and with dark shading for differences above 0.1 K.

warming from  $GSO_T$  is similar to the filtered observations (Figure 16a) in the troposphere.

## 6. Summary and Conclusions

[91] We have presented results from a set of simulations of HadCM3. It has a physically based interactive sulphur cycle, a simple parameterization of the first indirect effect of sulphate aerosols [Twomey, 1974], and a more complete and accurate radiation scheme than its predecessor, HadCM2, allowing explicit representation of well-mixed greenhouse gases. HadCM3 has higher resolution in the ocean than HadCM2, and additional changes were made to the atmospheric component of the model. These changes have removed the need for flux adjustments to keep the model stable for multicentury integrations.

[92] We forced the model with “historical” changes in greenhouse gas concentrations, sulphate emissions, tropospheric and stratospheric ozone, and solar irradiance changes, and changes in volcanic stratospheric aerosol in

four ensembles each of four simulations. Total simulated anthropogenic forcing is almost constant from 1980 onward due to a strong negative forcing from stratospheric ozone decline. Despite this, ensembles with and without ozone decline warm at similar rates. This negative forcing due to ozone is outside the range quoted by Schimel *et al.* [1996] and is partly due to changes in tropopause height. Therefore, we plan to investigate both forcing and response in more detail in a subsequent publication. Other anthropogenic forcings are within the range quoted by Shine *et al.* [1995].

[93] We found that the effects of well-mixed greenhouse gases, other anthropogenic effects (largely the indirect effect of sulphate aerosols), and natural causes (solar irradiance changes and volcanic eruptions) could be detected in the record of surface temperature change during the entire twentieth century. The best fit combination of simulations was consistent with observations during the century and in all 50-year periods we considered. We detected the responses to both well-mixed greenhouse gases and other anthropogenic forcings in all six 50-year periods we investigated. We also detected the response to natural forcings in the 1907–1957 period, but this was not robust to some technical details of the analysis.

[94] We found that the early twentieth century warming can be explained by a response to natural forcings, a large warming, relative to other factors, from internal climate variability, with the effect of greenhouse gases largely being balanced by other anthropogenic forcings. During 1907–1957 we found that there was negligible net anthropogenic warming, with the effect of greenhouse gases largely being balanced by other anthropogenic forcings. Therefore, in this period, the warming was largely naturally caused. Reconstructions of temperature changes, using proxy indicators, of the last 500–1000 years [Crowley, 2000; Mann *et al.*, 1998] suggest that the observed warming in this period is unusually rapid. If our analyses are correct, in attributing it largely to natural causes, this was an unusual natural event. We believe that further investigation of this period is needed.

[95] The late century warming was largely explained by greenhouse gases offset by the effect of volcanic aerosol and the indirect effect of anthropogenic aerosols. Over the entire century, natural forcings make no net contribution as they warm early in the century and cool from the 1960s on. Greenhouse gases warm at a rate of  $0.9 \pm 0.24$  K/century, while other anthropogenic forcings cool at a rate of  $0.4 \pm 0.26$  K/century, giving a total anthropogenic warming of  $0.5 \pm 0.15$  K/century.

[96] On 50-year timescales our results are generally similar to that of Tett *et al.* [1999], with similar total anthropogenic warming. We find more warming from greenhouse gases and more cooling from sulphates and ozone than do Tett *et al.* [1999] in all periods except the 1947–1997 period, when we find less sulphate cooling. Thus the total anthropogenic warming is robust to using HadCM3 rather than HadCM2, but the contributions from different factors are less so.

[97] We detected the effect of other anthropogenic forcings on the radiosonde record of temperature change in the free atmosphere from 1961 to 1995, but with a simulated tropospheric response  $\sim 50\%$  too large. We found no evi-

dence of a climatic effect from stratospheric ozone decline nor a natural effect on the free troposphere. Analysis on shorter timescales might detect the influence of volcanic eruptions and the solar cycle.

[98] The most crucial caveat in our work is that the variability we use to compute uncertainty limits is derived from simulations. Analysis of the free atmosphere suggests that the simulated stratospheric variance is too small by as much as a factor of 5. *Collins et al.* [2002] compared the variability of simulated summer near-surface temperatures from CONTROL with a proxy temperature data set from circa 1400 to 1950. These results suggest that the internal variance of HadCM3 is 2–3 times smaller than the variance estimated from the proxy data, but at least some of the differences may be due to neglect of naturally forced climate variability. After inflating the simulated variance by a factor of 5, we still detected the effect of greenhouse gases, though not other factors.

[99] Before 1979, there is little direct measurement of the changes in solar irradiance and thus considerable uncertainty in its time series. For example, we could have used the time series of *Hoyt and Schatten* [1993] rather than *Lean et al.* [1995a]. These time series are based on different assumptions about what determines solar irradiance change on decadal to century timescales. There is also some uncertainty in the forcing from explosive volcanic eruptions. *Lacis et al.* [1992] quote a forcing from volcanoes of  $30 \text{ W/m}^2$  (without stratospheric adjustment) per unit aerosol optical depth. We find a forcing of  $20 \text{ W/m}^2$  per unit aerosol optical depth once we include stratospheric adjustment. In the century analysis we found that the simulated and observed responses to natural forcings agreed, but in several 50-year analyses they did not agree. Since we only carried out simulations with total natural forcing, we were not able to explore differential error in the solar and volcanic forcings.

[100] European surface observations indicate that the model has about half the anthropogenic sulphate aerosol concentrations observed. Nonsulphate aerosols such as black carbon have not been taken into account. Since black carbon exerts a positive forcing and there should be a strong correlation between the spatial and temporal distributions of sulphur and black carbon emissions from fossil fuel combustion, this may mitigate the effect of the underestimated direct sulphate forcing. Furthermore, the bulk of the negative radiative forcing (offsetting the effect of the well-mixed greenhouse gases) is due to the first indirect effect of sulphate aerosol on cloud albedo, the magnitude of which is extremely uncertain [*Schimel et al.*, 1996], as is the impact of underestimating anthropogenic sulphate aerosol concentrations on it. We have not included the second indirect effect, which increases cloud lifetime [*Albrecht*, 1989], which could be of similar importance to the first indirect effect.

[101] In our simulations, stratospheric ozone decline produced a strong negative forcing but a weak near-surface temperature response. If we neglect this forcing, we find that the simulated response to greenhouse gases is significantly overestimated in the 1937–1987 and 1947–1997 periods.

[102] We have not considered the effects of other forcings, such as changes in land-surface properties and

mineral dust, which could have affected climate. Nor have we considered the effect of observational error on our results, which may be significant for the radiosonde data [*Gaffen et al.*, 2000]. Finally, we have not explicitly considered the effect of noise in the signals. In the century analysis the natural signal has a low signal-to-noise ratio, so that its estimated amplitude is biased toward zero and the computed uncertainty ranges are probably too small. Work is in progress to investigate the effects of such contamination. Nevertheless, our results strongly suggest that anthropogenic forcings have been the dominant cause of temperature changes over the last 30–50 years.

## Appendix A.

### Appendix A.1. Computations of Radiative Forcings

[103] Radiative forcing at the tropopause varies because of changes in the composition of radiatively active substances such as  $\text{CO}_2$  and aerosols and also because of changes in the climate of the stratosphere [*Schimel et al.*, 1996; *Hansen et al.*, 1997]. In this appendix we derive an expression that allows us to calculate it and then show how the forcing was computed for each component.

[104] Radiative forcing ( $\Delta F$ ) is defined as

$$\Delta F = F(S_1, R_1) - F(S_0, R_0), \quad (\text{A1})$$

where  $F$  is the net flux across the tropopause,  $S$  is the stratospheric climate, and  $R$  is the composition of the radiatively active substances. States are labeled 1 (perturbed state, for example, current concentrations of  $\text{CH}_4$ ) and 0 (reference state against which forcing is computed, for example, preindustrial concentrations of  $\text{CH}_4$ ).

[105] We can rewrite (A1) as

$$\Delta F = \Delta_S F(R_0) + \Delta_R F(S_1), \quad (\text{A2})$$

where we define

$$\Delta_S F(R_0) = F(S_1, R_0) - F(S_0, R_0)$$

and

$$\Delta_R F(S_1) = F(S_1, R_1) - F(S_1, R_0).$$

[106] From the perturbed simulations, we diagnosed the instantaneous forcing ( $\Delta_R F(S_1)$ ) by calling the radiation scheme twice. In one call the changes in forcing agents were applied ( $R_1$ ), and in the other call the forcing agent was kept at its preindustrial composition ( $R_0$ ). After both calls the increments from the first call were then applied to update the model state with radiative diagnostics stored from both. The instantaneous forcing was then computed as the difference in total flux at the tropopause (diagnosed by the model at each point and time step) between the two calls. This differs from *Schimel et al.* [1996], who compute instantaneous forcing from  $\Delta_R F(S_0)$ .

[107] We computed the adjustment of the forcing ( $\Delta_S F(R_0)$ ) as the change in downward flux ( $\Delta_S F^{\downarrow}(R_0)$ ) at the tropopause, with any change in the upward flux being

considered part of the climate system's response, not its forcing. Then we have

$$\begin{aligned}\Delta_S F^\downarrow(R_0) &= F^\downarrow(S_1, R_0) && -F^\downarrow(S_0, R_0) \\ &= F^\downarrow(S_1, R_1) - \Delta_R F^\downarrow(S_1) && -F^\downarrow(S_0, R_0)\end{aligned}$$

[108] This then gives the total radiative forcing:

$$\Delta F = \overbrace{F^\downarrow(S_1, R_1) - \Delta_R F^\downarrow(S_1) - F^\downarrow(S_0, R_0)}^{\text{Adjustment}} + \underbrace{\Delta_R F^\downarrow(S_1)}_{\text{Instantaneous}} \quad (\text{A3})$$

[109] We compute total radiative forcing from this equation. The adjustment to the forcing (the change in flux due to changes in the stratospheric temperature and the change in height of the tropopause) was computed using the first three terms. The first term was diagnosed in the main simulation, the second term was diagnosed by calling the radiation code twice in the same simulation, and the third term was diagnosed from a reference simulation. Note that the tropopause height, which has adjusted to the forcing factors, will be different from that in the reference simulation. We computed the instantaneous forcing as outlined earlier. Note that (A3) could be rewritten as

$$\Delta F = F^\downarrow(S_1, R_1) - F^\downarrow(S_0, R_0) + \Delta_R F^\downarrow(S_1), \quad (\text{A4})$$

that is, the difference in downward flux between the forced and control simulations plus the instantaneous change in upward flux.

[110] Variations in tropopause height are not normally considered in radiative forcing calculations. We believe that this effect should be included to the extent that the height of the tropopause changes due to changes in the stratospheric climate. Tropopause height can also vary systematically because of changes in the troposphere and is thus part of the climate system's response. Most of our computations of radiative forcing use experiments with fixed SSTs, so, to first order, any changes in the tropopause are due to changes in stratospheric climate (or noise).

## Appendix A.2. Diagnosis of Radiative Forcings

### Appendix A.2.1. Greenhouse gases

[111] A 15-month simulation of HadAM3 using climatological SSTs was carried out with twice the preindustrial values of  $\text{CO}_2$ , and the total forcing was diagnosed from the last 12 months of that simulation. The reference state used current concentrations of  $\text{CO}_2$ . The forcing was then scaled by  $\log[\text{CO}_2]$  to obtain the time-dependent forcing. For  $\text{N}_2\text{O}$  and  $\text{CH}_4$ , single time step simulations with each individually and with both were carried out. The forcing for each was independently scaled by the square root of the concentration, and the overlap factor was computed according to *Shine et al.* [1990], scaled to match the simulation in which both gases were included. The forcing from HCFCs is calculated from the *Schimel et al.* [1996] values, rescaled to give agreement with instantaneous forcing diagnosed

from the full model and then to allow for a small stratospheric adjustment.

### Appendix A.2.2. Sulphates

[112] Single-year reruns of sections of the first HadCM3 ANTHRO simulation were used to diagnose forcing due to both the direct and indirect effects. These reruns were carried out for the years 1860, 1900, 1950, 1975, and 2000. Three calls were made to the radiation code: The first call had the direct effect of sulphates removed, and in the second call the cloud albedo perturbation was not applied. The third call was used to evolve the model simulation as in the standard HadCM3 run and so had both effects included. Forcings were then computed from the differences between the first and third calls (direct forcing) and between the second and third calls (indirect forcing) and were linearly interpolated in time. No account was taken of stratospheric adjustment in these calculations.

### Appendix A.2.3. Ozone

[113] Seasonally varying ozone for the years 1860, 1900, 1950, 1975, 1990, and 2000 used in ANTHRO simulations were used to force several simulations of HadAM3. Each simulation used seasonally varying climatological SSTs and the ozone values (both tropospheric and stratospheric) for one of the years and was integrated for 3 years. All other climate forcings were set to the CONTROL values. Data were discarded from the first year of each integration to allow the stratosphere to adjust, and forcings were computed as earlier. The stratospheric adjustment was computed by differencing the average downward tropopause fluxes from a 10-year control simulation using the same SSTs but using preindustrial ozone values.

[114] Similar computations were done for tropospheric and stratospheric only ozone changes for 1975, 1990, and 1998 conditions. Forcings were then linearly interpolated in time. In 1998 we found global averages of the instantaneous forcing due to stratospheric ozone to be  $0.04 \text{ W/m}^2$ , the adjustment forcing to be  $-0.57 \text{ W/m}^2$ , and the total forcing to be  $-0.53 \text{ W/m}^2$ . If the calculations are done with a fixed tropopause, then the instantaneous forcing is  $0.10 \text{ W/m}^2$  and the adjustment is  $-0.41 \text{ W/m}^2$ , giving a total forcing of  $-0.31 \text{ W/m}^2$ .

### Appendix A.2.4. Natural

[115] Forcing was computed by setting the reference values of solar irradiance, its distribution, and volcanic aerosol across the solar spectrum to their control values and calling the radiation code once every 15 hours throughout the coupled simulations. Sampling the forcing every 15 hours gives good coverage of the diurnal cycle over a month. In these simulations, there may be some feedbacks on the stratosphere, and thus on the adjusted fluxes, from changes in tropospheric temperatures, but, as the near-surface temperature changes are generally small (Figure 3d) we neglect them.

## Appendix B. Transformations

[116] We use the linear transformation,  $\mathbf{A}$ , to transform  $\mathbf{X}$  to  $\mathbf{X}'$ .

[117] Given

$$\begin{aligned}\mathbf{X}' &= \mathbf{X}\mathbf{A} \\ \mathbf{X}'\tilde{\beta}' &= \mathbf{X}\tilde{\beta} = \mathbf{y},\end{aligned}$$

then

$$\mathbf{X}(\mathbf{A}\tilde{\beta}') = \mathbf{X}\tilde{\beta};$$

thus

$$\tilde{\beta}' = \mathbf{A}^{-1}\tilde{\beta}. \quad (\text{B1})$$

[118] For example,

$$\begin{pmatrix} (G, SO, \text{NATURAL}) \\ (\text{GHG}, \text{ANTHRO}, \text{NATURAL}) \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

then

$$\begin{pmatrix} \tilde{\beta}_G \\ \tilde{\beta}_{SO} \\ \tilde{\beta}_{\text{NATURAL}} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \tilde{\beta}_{\text{GHG}} \\ \tilde{\beta}_{\text{ANTHRO}} \\ \tilde{\beta}_{\text{NATURAL}} \end{pmatrix}$$

and similarly for the other transformations we use.

[119] To obtain the transformation for  $\tilde{\mathbf{V}}(\tilde{\beta})$ , premultiply equation (2) by  $\mathbf{A}^{-1}$ , giving

$$\begin{aligned} \mathbf{A}^{-1}\tilde{\beta} &= \mathbf{A}^{-1}\mathbf{F}^T\mathbf{y} \\ \tilde{\beta}' &= \mathbf{A}^{-1}\mathbf{F}^T\mathbf{y}; \end{aligned}$$

thus

$$\begin{aligned} \tilde{\mathbf{V}}(\tilde{\beta}') &= \mathbf{F}'^T\mathbf{C}_{N_2}\mathbf{F}' \\ &= \mathbf{A}^{-1}\mathbf{F}^T\mathbf{C}_{N_2}\mathbf{F}\mathbf{A}^{-1T} \\ &= \mathbf{A}^{-1}\tilde{\mathbf{V}}(\tilde{\beta})\mathbf{A}^{-1T}. \end{aligned} \quad (\text{B2})$$

## Notation

$\tilde{\beta}$	Best-estimate scaling of simulated signals.
$\mathbf{C}_N$	covariance matrix used for optimization.
$\mathbf{C}_{N_2}$	covariance matrix used to estimate uncertainties.
$\nu_1$	estimated DOF of $\mathbf{C}_N$ .
$\nu_2$	estimated DOF of $\mathbf{C}_{N_2}$ .
$\mathbf{X}$	matrix of simulated signals; each column is a signal.
$\mathbf{y}$	observations.
$\tilde{\mathbf{V}}(\tilde{\beta})$	covariance of $\tilde{\beta}$ .
$\kappa$	truncation applied to $\mathbf{C}_N$ .
$m_i$	size of $i$ th ensemble.
GHG	simulated response to well-mixed greenhouse gases.
ANTHRO	simulated response to greenhouse gases, sulphates, and ozone.
TROP-ANTHRO	as in ANTHRO but without stratospheric ozone decline.
NATURAL	simulated response to solar irradiance and volcanic aerosol forcings.

CONTROL control simulation.

$G$  response to well-mixed greenhouse gases.

$S$  response to direct and first indirect effect of sulphates.

$O_T$  response to tropospheric ozone.

$O_S$  response to stratospheric decline.

[120] **Acknowledgments.** Financial support to carry out the simulations and fund S. F. B. T., G. S. J., P. A. S., D. C. H., A. J., C. E. J., D. L. R., D. M. H. S., and M. J. W. was provided by UK Department of Environment, Food and Rural Affairs contract PEC/D 7/12/37. J. F. B. M., W. J. L., and T. C. J. were all supported by the UK Government Meteorological Research contract. M. R. A. was supported by a Research Fellowship from the UK Natural Environment Research Council. Supplementary support was provided by European Commission contract ENV4-CT97-0501 (QUARCC). The help and encouragement of Geoff Jenkins during the work reported here is gratefully acknowledged, as is the contribution of the many colleagues who developed HadCM3. Comments from three anonymous referees helped to improve the paper.

## References

- Albrecht, B. A., Aerosols, cloud microphysics and fractional cloudiness, *Science*, *245*, 1227–1230, 1989.
- Allen, M. R., and L. A. Smith, Monte Carlo SSA: Detecting irregular oscillations in the presence of coloured noise, *J. Clim.*, *9*, 3373–3404, 1996.
- Allen, M. R., and P. A. Stott, Estimating signal amplitudes in optimal fingerprinting, I, Theory, *Clim. Dyn.*, in press, 2002.
- Allen, M. R., and S. F. B. Tett, Checking for model consistency in optimal fingerprinting, *Clim. Dyn.*, *15*, 419–434, 1999.
- Barnett, T. P., et al., Detection and attribution of recent climate change: A status report, *Bull. Am. Meteorol. Soc.*, *80*, 2631–2659, 1999.
- Collins, M., S. F. B. Tett, and C. Cooper, The internal climate variability of HadCM3, a version of the Hadley Centre coupled model without flux adjustments, *Clim. Dyn.*, *17*, 61–81, 2001.
- Collins, M., T. J. Osborn, S. F. B. Tett, K. R. Briffa, and F. H. Schweingruber, A comparison of the variability of a climate model with a network of tree-ring densities, *J. Clim.*, *15*, 1497–1515, 2002.
- Collins, W., D. Stevenson, C. Johnson, and R. Derwent, Tropospheric ozone in a global-scale three-dimensional Lagrangian model and its response to  $\text{NO}_x$  emission controls, *J. Atmos. Chem.*, *26*, 223–274, 1997.
- Crowley, T., Causes of climate change over the past 1000 years, *Science*, *289*, 270–277, 2000.
- Cusack, S., J. M. Edwards, and J. M. Crowther, Investigating  $k$  distribution methods for parameterizing gaseous absorption in the Hadley Centre climate model, *J. Geophys. Res.*, *104*, 2051–2057, 1999.
- Delworth, T. L., and T. R. Knutson, Simulation of early 20th century global warming, *Science*, *287*, 2246–2250, 2000.
- Dignon, J., and S. Hameed, Global emissions of nitrogen and sulfur oxides from 1860 to 1980, *J. Air Pollut. Control Assoc.*, *39*, 180–186, 1990.
- Edwards, J. M., and A. Slingo, Studies with a flexible new radiation code, I, Choosing a configuration for a large-scale model, *Q. J. R. Meteorol. Soc.*, *122*, 689–719, 1996.
- Gaffen, D., M. Sargent, R. Habermann, and J. Lanzante, Sensitivity of tropospheric and stratospheric temperature trends to radiosonde data quality, *J. Clim.*, *13*, 1776–1796, 2000.
- Gordon, C., C. Cooper, C. A. Senior, H. Banks, J. M. Gregory, T. C. Johns, J. F. B. Mitchell, and R. A. Wood, The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments, *Clim. Dyn.*, *16*, 147–168, 2000.
- Haigh, J. D., The role of stratospheric ozone in modulating the solar radiative forcing of climate, *Nature*, *370*, 544–546, 1994.
- Haigh, J. D., Impact of solar variability on climate, *Science*, *272*, 981–984, 1996.
- Hansen, J., M. Sato, and R. Ruedy, Radiative forcing and climate response, *J. Geophys. Res.*, *102*, 6831–6864, 1997.
- Hasselmann, K., On the signal-to-noise problem in atmospheric response studies, in *Meteorology Over the Tropical Oceans*, edited by D. B. Shaw, pp. 251–259, Royal Meteorol. Soc., Reading, UK, 1979.
- Hasselmann, K., Optimal fingerprints for the detection of time-dependent climate change, *J. Clim.*, *6*, 1957–1971, 1993.
- Hasselmann, K., Multi-pattern fingerprint method for detection and attribution of climate change, *Clim. Dyn.*, *13*, 601–611, 1997.
- Hegerl, G., and M. Allen, Origins of model-data discrepancies in optimal fingerprinting, *J. Clim.*, in press, 2002.
- Hegerl, G. C., and G. R. North, Comparison of statistically optimal ap-



- proaches to detecting anthropogenic climate change, *J. Clim.*, *10*, 1125–1133, 1997.
- Hegerl, G. C., K. Hasselmann, U. Cubasch, J. F. B. Mitchell, E. Roeckner, R. Voss, and J. W. W. Wigley, Multi-fingerprint detection and attribution analysis of greenhouse gas, greenhouse gas-plus-aerosol and solar forced climate change, *Clim. Dyn.*, *13*, 613–634, 1997.
- Hegerl, G. C., P. A. Stott, M. R. Allen, J. F. B. Mitchell, S. F. Tett, and U. Cubasch, Optimal detection and attribution of climate change: Sensitivity of results to climate model differences, *Clim. Dyn.*, *16*, 737–754, 2000.
- Hegerl, G. C., P. D. Jones, and T. P. Barnett, Effect of observational sampling error on the detection of anthropogenic climate change, *J. Clim.*, *14*, 198–207, 2001.
- Hill, D. C., M. R. Allen, N. P. Gillett, S. F. B. Tett, P. A. Stott, G. S. Jones, W. Ingram, and J. F. B. Mitchell, Natural and anthropogenic causes of recent climate change, in *Detecting and Modelling Regional Climate Change*, edited by I. M. Brunet and B. D. Lopez, pp. 275–290, Springer-Verlag, New York, 2001.
- Hoyt, D. V., and K. H. Schatten, A discussion of plausible solar irradiance variations, 1700–1992, *J. Geophys. Res.*, *98*, 18,895–18,906, 1993.
- Johns, T. C., R. E. Carnell, J. F. Crossley, J. M. Gregory, J. F. B. Mitchell, C. A. Senior, S. F. B. Tett, and R. A. Wood, The second Hadley Centre coupled ocean-atmosphere GCM: Model description, spinup and validation, *Clim. Dyn.*, *13*, 103–134, 1997.
- Johns, T. C., et al., Anthropogenic climate change for 1860 to 2100 simulated with the HadCM3 model under updated emissions scenarios, *HCTN 22*, Hadley Centre, Met Office, Bracknell, UK, 2001. (Available at [http://www.metoffice.com/research/hadleycentre/pubs/HCTN/HCTN\\_22.pdf](http://www.metoffice.com/research/hadleycentre/pubs/HCTN/HCTN_22.pdf).)
- Jones, A., D. L. Roberts, and M. J. Woodage, The indirect effects of anthropogenic sulphate aerosol simulated using a climate model with an interactive sulphur cycle, *HCTN 14*, Hadley Centre, Met Office, Bracknell, UK, 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., chap. 6, pp. 285–358, Cambridge Univ. Press, New York, 1996.
- Lacis, A., J. Hansen, and M. Sato, Climate forcing by stratospheric aerosols, *Geophys. Res. Lett.*, *19*, 1607–1610, 1992.
- Lean, J., J. Beer, and R. Bradley, Reconstruction of solar irradiance since 1610: Implications for climate change, *Geophys. Res. Lett.*, *22*, 3195–3198, 1995a.
- Lean, J. L., O. R. White, and A. Skumanich, On the solar ultraviolet spectral irradiance during the Maunder Minimum, *Global Biogeochem. Cycles*, *9*, 171–182, 1995b.
- Mann, M. E., R. S. Bradley, and M. K. Hughes, Global-scale temperature patterns and climate forcing over the past six centuries, *Nature*, *392*, 779–787, 1998.
- Mardia, K. V., J. T. Kent, and J. M. Bibby, *Multivariate Analysis*, Academic, San Diego, Calif., 1979.
- Marsh, N., and H. Svensmark, Cosmic rays, clouds, and climate, in *Solar Variability and Climate*, vol. 94, *Space Sci. Rev.*, edited by E. Friis-Christensen et al., pp. 215–230, Kluwer, Norwell, Mass., 2000a.
- Marsh, N. D., and H. Svensmark, Low cloud properties influenced by cosmic rays, *Phys. Rev. Lett.*, *85*, 5004–5007, 2000b.
- Mitchell, J. F. B., T. C. Johns, J. M. Gregory, and S. F. B. Tett, Climate response to increasing levels of greenhouse gases and sulphate aerosols, *Nature*, *376*, 501–504, 1995.
- Nakićenović, et al., *Special Report on Emissions Scenarios*, 599 pp., Cambridge Univ. Press, New York, 2000.
- North, G. R., and K.-K. Kim, Detection of forced climate signals, II, Simulation results, *J. Clim.*, *8*, 409–417, 1995.
- North, G. R., and M. J. Stevens, Detecting climate signals in the surface temperature record, *J. Clim.*, *11*, 563–577, 1998.
- North, G. R., K.-K. Kim, S. S. P. Shen, and J. W. Hardin, Detection of forced climate signals, I, Filter theory, *J. Clim.*, *8*, 401–408, 1995.
- Orn, G., U. Hansson, and H. Rodhe, Historical worldwide emissions of anthropogenic sulfur: 1860–1985, *Tech. Rep. CM-91*, Stockholm Univ., Stockholm, 1996.
- 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.
- Parker, D., M. Gordon, D. Cullum, D. Sexton, C. Folland, and N. Rayner, A new global gridded radiosonde temperature database and recent temperature trends, *Geophys. Res. Lett.*, *24*, 1499–1502, 1997.
- Pope, V. D., M. L. Gallani, P. R. Rowntree, and R. A. Stratton, The impact of new physical parameterizations in the Hadley Centre climate model, HadAM3, *Clim. Dyn.*, *16*, 123–146, 2000.
- Randel, W. J., and F. Wu, A stratospheric ozone trends data set for global modeling studies, *Geophys. Res. Lett.*, *26*, 3089–3092, 1999.
- Rowell, D. P., Assessing potential seasonal predictability with an ensemble of multidecadal GCM simulations, *J. Clim.*, *11*, 109–120, 1998.
- Salby, M., and P. Callaghan, Connection between the solar cycle and the QBO: The missing link, *J. Clim.*, *13*, 2652–2662, 2000.
- Santer, B. D., et al., A search for human influences on the thermal structure of the atmosphere, *Nature*, *382*, 39–45, 1996a.
- Santer, B. D., T. M. L. Wigley, T. P. Barnett, and E. Anyamba, Detection of climate change and attribution of causes, in *Climate Change 1995: The Science of Climate Change*, edited by J. T. Houghton et al., chap. 8, pp. 407–444, Cambridge Univ., New York, 1996b.
- 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.
- Schimel, D., et al., Radiative forcing of climate change, in *Climate Change 1995: The Science of Climate Change*, edited by J. T. Houghton et al., chap. 2, pp. 65–131, Cambridge Univ. Press, New York, 1996.
- Shine, K., R. Derwent, D. Wuebbles, and J.-J. Morcrette, Radiative forcing of climate, in *Climate Change: The IPCC Scientific Assessment*, edited by J. Houghton, G. Jenkins, and J. Ephraums, chap. 5, pp. 45–68, Cambridge Univ. Press, New York, 1990.
- Shine, K. P., Y. Fouquart, V. Ramaswamy, S. Solomon, and J. Srinivasan, Radiative forcing, in *Climate Change 1994*, edited by J. T. Houghton et al., chap. 4, pp. 168–203, Cambridge Univ. Press, New York, 1995.
- Stevenson, D. S., C. E. Johnson, W. J. Collins, R. G. Derwent, K. P. Shine, and J. M. Edwards, Evolution of tropospheric ozone radiative forcing, *Geophys. Res. Lett.*, *25*, 3819–3822, 1998.
- Stott, P. A., and S. F. B. Tett, Scale-dependent detection of climate change, *J. Clim.*, *11*, 3282–3294, 1998.
- Stott, P. A., S. F. B. Tett, G. S. Jones, M. R. Allen, J. F. B. Mitchell, and G. J. Jenkins, External control of twentieth century temperature by natural and anthropogenic forcings, *Science*, *290*, 2133–2137, 2000.
- Stott, P. A., S. F. B. Tett, G. S. Jones, M. R. Allen, W. J. Ingram, and J. F. B. Mitchell, Attribution of twentieth century temperature change to natural and anthropogenic causes, *Clim. Dyn.*, *17*, 1–21, 2001.
- Svensmark, H., and E. Friis-Christensen, Variation of cosmic ray flux and global cloud coverage—A missing link in solar-climate relationships, *J. Atmos. Sol. Terr. Phys.*, *59*, 1225–1232, 1997.
- 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*, *247*, 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.
- Twomey, S. A., Pollution and the planetary albedo, *Atmos. Environ.*, *8*, 1251–1256, 1974.
- Williams, K. D., C. A. Senior, and J. F. B. Mitchell, Transient climate change in the Hadley Centre models: The role of physical processes, *J. Clim.*, *14*, 2659–2674, 2001.
- Wilson, D., and S. Ballard, A microphysically based precipitation scheme for the UK Meteorological Office Unified Model, *Q. J. R. Meteorol. Soc.*, *125*, 1607–1636, 1999.

M. R. Allen, Space Science and Technology Department, Rutherford Appleton Laboratory, Chilton OX11 0QX, UK. (m.allen1@physics.ox.ac.uk)

D. C. Hill, Clinical Trial Service Unit, University of Oxford, Harkness Building, Radcliffe Infirmary, Oxford OX2 6HE, UK. (david.hill@ctu.ox.ac.uk)

W. J. Ingram, T. C. Johns, C. E. Johnson, A. Jones, G. S. Jones, J. F. B. Mitchell, D. L. Roberts, D. M. H. Sexton, P. A. Stott, S. F. B. Tett, and M. J. Woodage, Hadley Centre for Climate Prediction and Research, Met Office, London Road, Bracknell, Berkshire RG12 2SY, UK. (william.ingram@metoffice.com; tim.johns@metoffice.com; colin.johnson@metoffice.com; andy.jones@metoffice.com; garth.s.jones@metoffice.com; john.f.mitchell@metoffice.com; david.roberts@metoffice.com; david.sexton@metoffice.com; peter.stott@metoffice.com; simon.tett@metoffice.com; margaret.woodage@metoffice.com)