ESTIMATION OF NATURAL AND ANTHROPOGENIC CONTRIBUTIONS TO TWENTIETH CENTURY TEMPERATURE CHANGE

Simon F. B. Tett, Gareth S. Jones, Peter A. Stott, David C. Hill[†], John F. B. Mitchell, Myles R. Allen[†‡],

William J. Ingram, Tim C. Johns, Colin E. Johnson,
Andy Jones, David L. Roberts, David M. H. Sexton and Margaret J. Woodage
Hadley Centre for Climate Prediction and Research,
The Met. Office, London Road, Bracknell, Berkshire RG12 2SY, UK.
[†]Space Science and Technology Dept., Rutherford Appleton Laboratory,
Chilton, OX11 0QX, UK
[‡]Dept. of Physics,Clarendon Laboratory, University of Oxford, Oxford,OX1 3PU, UK

ABSTRACT

This is a shortened version of a paper with the same title and authors in preparation for J. Geophys. Res.

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 parametrisation of the effect of aerosols on cloud albedo (first indirect effect) and a radiation scheme which 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; tropospheric anthropogenic forcings (tropospheric ozone, well-mixed greenhouse gases and the direct and first indirect effects of sulphate aerosol); 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 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 ± 0.24 K/century, offset by cooling from other anthropogenic forcings of 0.4 ± 0.26 K/century giving a total anthropogenic warming trend of 0.5 ± 0.15 K/century. Over the entire century natural forcings give a linear trend close to zero. Observed surface temperature changes are generally consistent with our simulations but the simulated tropospheric response, since the 1960s, is about 50% too large.

Our analysis suggests that the early 20th 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, plus 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.

1. INTRODUCTION

Several authors (e.g. Santer et al. (1996); Hegerl et al. (1997); North & Stevens (1998); Tett et al. (1999); Hegerl et al. (2000); Stott et al. (2000)) 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 & Kim 1995; Hasselmann 1997; Hegerl & North 1997; Allen & Tett 1999).

Tett et al. (1999) (T99 from hereon) and Stott et al. (2000) (S00 from hereon) computed responses from the Atmosphere/Ocean General Circulation Model (AOGCM) HadCM2(Johns et al. 1997) to solar, volcanic, greenhouse and the direct anthropogenic sulphate forcing. They compared the responses with observations of surface temperature using a spatio-temporal methodology and concluded that natural causes alone could not explain observed changes in surface temperature from 1946–1996. HadCM2 included an ocean model with a resolution of $2.5^{\circ} \times 3.75^{\circ}$ and

Proc. 1st Solar & Space Weather Euroconference, 'The Solar Cycle and Terrestrial Climate', Santa Cruz de Tenerife, Tenerife, Spain, 25-29 September 2000 (ESA SP-463, December 2000)

needed a flux adjustment[1] to keep the control simulation stable and its climate close to the current climate. It represented all greenhouse gases as equivalent CO_2, and the direct effect of sulphates as changes in surface albedo.

Barnett et al. (1999) compared simulations from several different models with observations and found that there were cases in which simulated linear trends in northern summer temperature were inconsistent with observations. Most of those models used a simple parametrisation of the effects of sulphate aerosols similar to that used in HadCM2. However they found that the amplitude of the "sulphate" component computed from a single simulation of ECHAM4 (a model with a representation of the indirect effect of aerosols and an interactive sulphur cycle) was, in one case, inconsistent with observations. If this result were confirmed by other models, which include physically based parametrisations of the direct and indirect effect of sulphates, then the hypothesis that sulphates alone have significantly offset greenhouse gas warming would be unlikely to be true.

The aim of this paper is to examine the contributions of natural and anthropogenic forcings to temperature change during the 20th century using 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 $1.25^\circ \times 1.25^\circ$. In addition to an increase in oceanic resolution it includes many improvements on HadCM2 which have removed the need for a flux adjustment. HadCM3 represents the radiative effects of CO_2, N_2O, CH_4, and some of the (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(Twomey 1974), which was not represented at all in HadCM2, is now modelled using a relatively simple, non-interactive technique.

The control simulation is stable for multi-century integrations and the temperature variability near the surface, though not in the free atmosphere, compares well with observations(Collins et al. 2000b). HadCM2 and HadCM3 show similar global-mean temperature responses to increases in greenhouse gases during the 20th and the 21st centuries but HadCM3 shows less tropical warming than HadCM2 due to changes in details of the physics parametrisations(Williams et al. 2000).

The rest of the paper is organised as follows. First we describe the simulations and observations. We then describe the simulated responses and compare them with observations. Next we describe the detection and attribution methodology. In section 5 we show the results of the analyses and in section 6 we conclude.

[1]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.

2. SIMULATIONS

The control simulation for HadCM3 (CONTROL) has constant, near pre-industrial[2], forcing and we use the first 1200 years of the simulation in our analysis. Four ensembles with different external forcings were carried out using HadCM3. Each ensemble consisted of four simulations. The ensembles are:

GHG The simulations were forced with historical changes in well-mixed greenhouse gases.

TROP-ANTHRO 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.

ANTHRO As TROP-ANTHRO except from 1974 stratospheric ozone decline was included.

NATURAL The simulations were forced with the solar irradiance timeseries of Lean et al. (1995) and a timeseries of stratospheric aerosol due to explosive volcanic eruptions (Sato et al. 1993). Both forcing timeseries have been extended to 1997.

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 in 1/Dec/1859 and the twelve anthropogenic simulations ended on 30/Nov/1999. The NATURAL simulations were integrated to 30/Nov/1997. Initial conditions for TROP-ANTHRO were taken from ANTHRO on 1/Dec/1974.

2.1. Observed datasets and data processing

We compare the results of the model simulations with an updated version of the surface temperature dataset of Parker et al. (1994) and with the HadRT2.1s radiosonde temperature dataset – an updated version of Parker et al. (1997). Radiosonde data from the Indian subcontinent (60°E – 90°E, 0 – 30°N) was removed because of apparent problems with its quality and the remaining data corrected for known changes in instruments by comparison with co-located MSU data(Parker et al. 1997).

Annual averages of both the surface and radiosonde datasets were computed from monthly-mean temperature anomalies. At each location we required there to be at least eight months of observations; otherwise we discarded the annual-mean value.

The annual-mean surface observations were decadally averaged, with periods ending in 1997. For each decade

[2]The concentrations (ppbv) used for the well-mixed greenhouse gases are: CO_2: 289600, CH_4: 792.1 N_2O: 285.1. The (H)(C)FCs all had zero concentrations.

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 100-year timescales using decadal data with the 100-year average removed. Locations in the observations at which less than five decades were present were omitted. This data was then filtered, using spherical harmonics, to remove scales below 5,000 km (T99, S00). Harmonics were further weighted by $1/\sqrt{2l+1}$ (l is the total spherical harmonic wavenumber) to give each spatial scale included equal weight(Stott & Tett 1998). Simulated data was decadally averaged, bilinearly interpolated in latitude and longitude to the observational grid. Simulated data was discarded where there were no observational data and then processed in the same way as the observations were.

When computing global-mean timeseries 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. As the observed data are anomalies relative to 1961–90 we computed the 1961–90 climate mean for each simulation and 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 timeseries.

Annual-mean simulated data from throughout the atmosphere was trilinearly (pressure, longitude and latitude) interpolated to the three-dimensional observed grid and discarded where there was 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 1971–90 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. 2000b).

Changes in surface temperature observed over the century show warming (Fig. 1(a)) 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 (Fig. 1(b)) in the stratosphere and 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 0.6K occurring at about 50°N and almost no warming at 30°N. Differences between the observations shown here and that of T96 (see their Fig. 2D) are due to the continued development of the radiosonde dataset and removal of data from the Indian sub-continent.

3. MODEL AND OBSERVED TEMPERATURE RESPONSES

Annual means of global-mean temperature from the ensemble averages (Fig. 2) show that the simulated responses are all inconsistent with the observations. From the 1920s until the 1950s GHG warms less than the observations. From the 1940s onwards it begins to warm and by the end of the 20th century 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 till 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 (not shown). We believe that this small response is due to the stratospheric ozone forcing being concentrated over Antarctica.

Natural forcings, in our simulations, 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.

The patterns of simulated response from the 20th century are shown in Fig. 3. 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 (Fig. 3(a)). ANTHRO and TROP-ANTHRO are in reasonable agreement with the observations (Fig. 1(a)), and both warm less than GHG especially in the mid-latitudes 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.

We now examine temperature changes throughout the atmosphere between the decade 1985–1995 and the twenty year period 1961–1980. All three anthropogenic ensembles have similar warming in the troposphere, greatest warming in the upper tropical troposphere and warm more in the northern hemisphere than the southern (Fig. 4). 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 the observations (Fig. 1(b)). Inclusion of stratospheric ozone decline in ANTHRO produces large stratospheric cooling (of up to 6K over Antarctica), especially in high latitudes, which brings this ensemble into better agreement with the observations (Fig. 1(c)). 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.

The boundary between cooling and warming is close to the tropopause in all ensembles except over Antarctica in ANTHRO (Fig. 4). 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 is 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.

Qualitative comparison of our ensembles with the observations suggests that ANTHRO is the most similar to the observations (compare Fig 4(c) with Fig 1(b)). As all the anthropogenic ensembles are quite similar in the troposphere it appears that 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

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 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). This assumes that the observations ($\mathbf{y}$) may be represented as a linear sum of simulated signals (X) and internal climate variability ($\mathbf{u}$):

$$\mathbf{y} = \mathsf{X}\beta + \mathbf{u} \tag{1}$$

where β_i is the scaling factor, or amplitude, that we apply to the i^{th} signal ($\mathbf{x_i}$) to obtain the best fit to the observations. In this paper the signals are ensemble averages from the simulations described earlier. 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 (β_i). Errors in the *patterns* of forcing and response are not taken into account by this procedure. The values of β which give the best fit (the best-estimate value $\tilde{\beta}$) to observations, using the standard linear regression approach are (AT99):

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

where C_N is the covariance matrix of natural variability ($\mathcal{E}(\mathbf{u}\mathbf{u}^T)$) estimated, in our case, from simulations of coupled-atmosphere ocean GCMs. We do not normally have enough data to accurately estimate the inverse covariance matrix (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.

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

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

where C_{N_2} is an estimate of $\mathcal{E}(\mathbf{u}\mathbf{u}^T)$ using data which is statistically independent of that used to estimate C_N.

We perform two related tests:

detection This 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 $\tilde{\mathsf{V}}(\tilde{\beta})$ and testing whether it includes zero. Rejection of this null and a positive value of $\tilde{\beta}_i$ implies detection.

amplitude-consistency This 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 $\tilde{\mathsf{V}}(\tilde{\beta})$ and testing whether it includes unity. In this test we inflate $\tilde{V}(\tilde{\beta})_{ij}$ by a factor of $\sqrt{(1+1/m_i)}\sqrt{(1+1/m_j)}$ to compensate for sampling noise in the signals, where m_i and m_j are the ensemble sizes. 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%.

The best estimate of the temperature trend (or any other linear diagnostic such as changes 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 $\tilde{V}(\tilde{\beta})_{ij}$, inflated to compensate for signal-noise, by the trends of the $_{\text{I}}{}^{\text{th}}$ and $_{\text{J}}{}^{\text{th}}$ ensembles.

Covariance matrices are estimated from intra-ensemble 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 eqn. (1). In all our analysis segments were overlapped by ten years. When computing covariance matrices from intra-ensemble variance we remove the ensemble average and scale each realisation of $\mathbf{u}$ (segment) by a factor of $\sqrt{(m-1)/m}$ where m is the number of ensemble members.

In Section 5 we analyse changes in near-surface temperature on 100-year timescales (century) and on 50-year timescales (50-year), and 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 thirty-five year period (Section 2.1).

For both the 50-year and the free-atmosphere analysis we use intra-ensemble variability from the GHG, ANTHRO and NATURAL ensembles to estimate C_N and data from CONTROL to estimate C_{N_2}. Any significant differences between C_N and C_{N_2} would reduce the power of the optimisation algorithm (i.e. increase uncertainty ranges) but would not introduce a bias in the estimated signal amplitudes.

For the century analysis we believe that nine realisations of century timescale variability from the intra-ensemble variability of HadCM3 is not enough to generate a sufficiently reliable estimate of C_N. Therefore we use control and intra-ensemble variability from five ensembles of HadCM2 (S00) to estimate C_N while C_{N_2} is estimated using HadCM3 CONTROL and intra-ensemble variability from the GHG, ANTHRO and NATURAL ensembles.

4.1. Consistency

We test that the best-estimate combination of signals is consistent with our linear statistical model (Eqn 1) by computing the residual sum of squares:

$$R^2 = \sum_{i=1}^{\kappa} \frac{(y_i - \sum_{j=1}^{n} X_{ij}\tilde{\beta}_j)^2}{C_{N_2 ii}}. \qquad (4)$$

where i is an index over the ranked eigenvectors of C_N, j is an index over signals and κ is the number of eigenvectors used to filter signals and observations (see subsection 4.3 for details).

In the case of noise-free signals R^2 is F-distributed (AT99). As an *ad hoc* correction for noise in the signals we scale R^2 by $1/(1+s)$, and assume that it is still F-distributed, 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 degrees of freedom for covariance matrices

In order to compute uncertainties and truncations we need an estimate of the degrees of freedom (dof) of the covariance matrices we compute. These matrices are computed from various different datasets and their dof is the sum of the dof of the individual datasets. For CONTROL the estimated dof, assuming maximally overlapped segments, is the number of non-overlapping segments multiplied by 1.5 (Allen & Smith (1996); S00) and rounded down to the nearest integer. For each ensemble the estimated dof is the number of non-overlapping 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 $\tilde{\mathsf{V}}(\tilde{\beta})$ is that of C_{N_2}.

The estimated degrees of freedom for the century analysis (see Table 1) may be over-optimistic as the individual HadCM2 ensemble members were all initialised from the same 1700-year control. Furthermore the last three simulations of each of the two solar ensembles were initialised by applying small random perturbations to the first solar simulation in each ensemble. Similarly the three HadCM3 ensembles were all initialised from the same HadCM3 control. 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 3%. The estimated dof of C_N is used to determine the maximum allowable truncation (see below) and so we explore the sensitivity of our results to truncation.

4.3. Truncation

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 at the 10% level all further analysis is carried out at this truncation (κ). All data is then filtered by projection onto the leading κ 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 at the 10% level 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

We used the same three tests as T99 and S00 to test for signal degeneracy or co-linearity (see pages 243–248 of Mardia et al. (1979)). We wish to err on the side of including too many signals as by not including a signal in an analysis we assume that its amplitude is zero. Therefore the largest value from the three tests determined the number of signals we considered.

If two signals are degenerate, the usual consequence is that neither is individually detectable, since a range of linear combinations are equally consistent with the data

including those which assign zero amplitude to one signal or the other. However, specific combinations of these signals may easily be detectable.

4.5. Transformations

We assume that the three anthropogenic signals (GHG, ANTHRO and TROP-ANTHRO) are linear combinations of the following physically-based signals:

G Response to well-mixed greenhouse gases alone.

O_T Response to stropospheric ozone changes.

O_S Response to stratospheric ozone decline.

O Response to both stratospheric and tropospheric ozone changes.

S Response to sulphates (indirect and direct)

namely,

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

The amplitudes and covariance matrices of these physically based signals are given by a linear transformation of the original amplitudes and of $\tilde{\mathsf{V}}(\tilde{\beta})$. 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:

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

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

Amplitude uncertainty ranges, and particularly the upper bound, estimated from signals with a low signal-to-noise ratio are likely to be incorrect(Allen & Stott 2000). We use the following summary statistic for the j^{th} signal to give us some guidance when this may be occurring:

$$(\text{SNR})^2 = \frac{m_j}{\kappa}\sum_{i=1}^{\kappa}\frac{X_{ij}^2}{C_{N_2 ii}}$$

where κ 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 as $F(\kappa,\nu_2)$ where ν_2 is the dof of C_{N_2}.

5. DETECTION AND ATTRIBUTION OF OBSERVED TEMPERATURE CHANGES

5.1. Changes in near-surface temperature on century timescales

We now examine changes in near-surface temperature from 1897–1997 using both spatial and temporal information. For most of the 20th century TROP-ANTHRO and ANTHRO are identical and therefore we use the latter in subsequent analyses. We transform the amplitudes of GHG and ANTHRO to obtain amplitudes of *G* (greenhouse gases) and *SO* (sulphates and ozone) as described in Section 4. Tests for degeneracy suggest that we can reliably estimate the amplitude of *G*, *SO* and NATURAL signals simultaneously. Thus all further analysis is done using this combination of signals.

The filtered observations (see Section 4) contain more than 96% of the observed variance (Table 1) and the residuals are consistent with those expected from CONTROL at all truncations. All three signals are detected demonstrating that all have had a significant impact on changes in near-surface temperature over the 20th 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.

Signal-to-noise ratio is large for the anthropogenic signals but small for NATURAL (Table 1) suggesting it is significantly noise-contaminated. Though our detection of NATURAL is probably robust, its estimated amplitude ranges, and in particular the upper range, are sensitive to this noise contamination(Allen & Stott 2000).

We reconstruct the global-mean temperature changes from the best-estimate signal amplitudes and simulated responses (Fig. 5). Well-mixed greenhouse gases and other anthropogenic effects (largely the indirect effect of sulphate aerosols) almost balance giving a total anthropogenic warming of approximately 0.1K from the beginning of the 20th century to the 1960s. Thereafter anthropogenic effects warm the planet by approximately 0.5K. From the 1950s onwards natural and anthropogenic non-greenhouse gas forcings each cause a cooling of about 0.1K. Together they offset about 0.2K of the estimated 0.6K warming due to greenhouse gases over the same period.

While Fig. 5 shows the best-estimate combination of signals, it is even more important to consider uncertainty ranges. These are most easily summarised in terms of

linear trends (Fig. 6 over selected periods (the entire century, 1897-1947 and 1947-1997 – see Section 4 for details.) Over the 20th century anthropogenic forcings cause a warming trend of 0.5 ± 0.15K/century. The trend due to greenhouse gases is 0.9 ± 0.24 K/century while the remaining anthropogenic factors cool at a rate of 0.4 ± 0.26 K/century. Over the century natural forcings contribute little to the observed trend.

During the early century greenhouse gases and natural forcings cause warming trends of about 0.2 to 0.3 K/century while other anthropogenic factors produce negligible cooling trends (Fig. 6). Over the last half of the century greenhouse gases warm the climate at a rate of 1.7 ± 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 about 0.3 ± 0.2 K/century. Thus, since 1947 changes in aerosol concentrations (anthropogenic and natural) have offset at least a third of the greenhouse gas warming.

5.2. Free atmosphere changes

We now examine the difference between the 10-year zonal-mean from 1986–1995 and the 20-year zonal-mean for 1961–1980 as in AT99.

Earlier we showed that the changes in the free atmosphere simulated by TROP-ANTHRO and GHG are 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 subsection 4.5 for details.

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 three to five times too large (Fig. 7(a)). At truncation seven the filtered observations contain 48% of the observed mass-weighted variance (Table 1) compared to 71% at truncation 36 (the truncation we believe the largest we can reasonably consider given the estimated dof of C_N — Table 1).

The SNR for the two anthropogenic signals is reasonably high (Table 1), while the SNR for the natural signal is less than one. We also find that the amplitude of at most two signals can reliably be simultaneously estimated. There are three reasons why we neglect natural effects in further analysis of the free-atmosphere changes. First, the simulated response to natural forcings (Fig. 4(d)) is small. Second, we did not detect NATURAL in any combination including it. Thirdly no linear amplification of it alone is consistent, using the F-test of subsection 4.1, with the observations. Thus we consider the GSO_T and O_S signals.

Failure to detect NATURAL does not rule out the possibility of a significant natural influence on climate. The simulated signal is weak and noise noise contaminated and so our failure to detect it does not strongly rule out the possibility of some process which preferentially amplifies the response to solar or volcanic forcing. 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 & Callaghan (2000); Hill et al. (2000).

The GSO_T and O_S case has residual variance consistent with CONTROL for all truncations less than or equal to seven (Fig. 7(a)). Detection of GSO_T, but not of O_S, occurs at those truncations (Fig. 7(b)). While the amplitude of O_S is consistent with unity the same is not true of GSO_T which has a best-estimate value of 0.65. This suggests that the simulated tropospheric response is about 50% stronger than the observed response.

Above truncation seven the residual variance is approximately three to five times larger than that of CONTROL (Fig. 7(a)) 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 (Fig. 8(a)). However, at this truncation, the filtered observations do not show the observed stratospheric cooling (Fig. 1(b)) as seen more clearly in the difference between the raw and the filtered observations (Fig. 8(b)). The raw observations are cooler in the stratosphere and approximately 0.1K warmer throughout large regions of the troposphere than 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 seven the best-estimate warming from GSO_T is similar to the filtered observations (Fig. 8(a)) in the troposphere.

6. SUMMARY AND CONCLUSIONS

We have presented results from a set of simulations of HadCM3. It has a physically based interactive sulphur cycle, a simple parametrisation of the first indirect effect of sulphate aerosols(Twomey 1974) and a better 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 multi-century integrations

We forced the model with "historical" changes in greenhouse gas concentrations, sulphate emissions, tropospheric and stratospheric ozone, solar irradiance changes and changes in volcanic stratospheric aerosol in four ensembles each of four simulations.

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 20th century. The best-fit combination of simulations was consistent with observations.

We found that the early 20th 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.

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 ± 0.24 K/century while other anthropogenic forcings cool at a rate of 0.4 ± 0.26 K/century giving a total anthropgenic warming of 0.5 ± 0.15 K/century.

We detected the effect of other anthropogenic forcings on the radiosonde record of temperature change in the free atmosphere from 1961-95 but with a simulated tropospheric response about 50% too large. We found no evidence 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.

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 five. Collins et al. (2000a) compared the variability of simulated summer near-surface temperatures from CONTROL with a proxy temperature dataset from circa 1400 to 1950. These results suggest that the internal variance of HadCM3 is two to three 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 five we still detected the effect of greenhouse gases though not other factors.

Before 1979 there is little direct measurement of the changes in solar irradiance and thus considerable uncertainty in its timeseries. For example we could have used the timeseries of Hoyt & Schatten (1993) rather than Lean et al. (1995). There is also some uncertainty in the forcing from explosive volcanic eruptions. Lacis et al. (1992) quote a forcing from volcanoes of 30 W/m^2 (without stratospheric adjustment) per unit aerosol optical depth. We find a forcing of 20 W/m^2 per unit aerosol optical depth once we include stratospheric adjustment. In the century analysis we found no evidence that the model's response to natural forcings was incorrect but found several 50-year periods when it was. As we only carried out simulations with total natural forcing we were not able to explore differential error in the solar and volcanic forcings.

European surface observations indicate that the model has about half the anthropogenic sulphate aerosol concentrations observed. Non-sulphate 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. 1995) 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.

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 1937–87 and 47–97 periods.

We have not considered the effects of other forcings such as changes in land-surface properties and mineral dust which could have effected 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 towards 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 to 50 years.

ACKNOWLEDGMENTS

Financial support to carry out the simulations and fund SFBT, PAS, GSJ, DCH, AJ, CEJ, DLR, DMHS and MJW was provided by U.K. Dept. of Environment Transport and the Regions contract PECD 7/12/37. JFBM, WJI and TCJ were all supported by the U.K. Public Met. Service Research contract. MRA was supported by a Research Fellowship from the U.K. 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.

REFERENCES

Albrecht B.A., 1989, Science, 245, 1227

Allen M.R., Smith L.A., 1996, J. Climate, 9, 3373

Allen M.R., Stott P.A., 2000, Cli. Dyn, in preparation

Allen M.R., Tett S.F.B., 1999, Cli. Dyn., 15, 419

Barnett T.P., Hasselmann K., Chelliah M., et al., 1999, BAMS, 2631–2659

Collins M., Osborn T.J., Tett S.F.B., Briffa K.R., Schweingruber F.H., 2000a, J. Clim., (Submitted)

Collins M., Tett S.F.B., Cooper C., 2000b, Clim. Dyn., (In press)

Gaffen D., Sargent M., Habermann. R., Lanzante J., 2000, J. Clim., 13, 1776

Gordon C., Cooper C., Senior C.A., et al., 2000, Climate Dyn., 16, 147

Hasselmann K., 1979, In: Shaw D.B. (ed.) Meteorology over the tropical oceans, 251–259, Royal Meteorological Society

Hasselmann K., 1993, J. Clim., 6, 1957

Hasselmann K., 1997, Cli. Dyn., 13, 601

Hegerl G.C., North G.R., 1997, J. Clim., 10, 1125

Hegerl G.C., Hasselmann K., Cubasch U., et al., 1997, Cli. Dyn., 13, 613

Hegerl G.C., Stott P.A., Allen M.R., et al., 2000, Cli. Dyn., (in press)

Hill D., Gillett N., Allen M., Stott P., Tett S., 2000, J. Geophys. Res., (In Prep.)

Hoyt D.V., Schatten K.H., 1993, J. Geophys. Res., 98, 18895

Johns T.C., Carnell R.E., Crossley J.F., et al., 1997, Cli. Dyn., 13, 103

Kattenberg A., Giorgi F., Grassl H., et al., 1996, In: Houghton J.T., Meira Filho L.G., Callander B.A., et al. (eds.) Climate Change 1995. The Science of Climate Change, chap. 6, 285–358, Cambridge University Press, Cambridge

Lacis A., Hansen J., Sato M., 1992, GRL, 19, 1607

Lean J., Beer J., Bradley R., 1995, Geophys Res Lett, 22, 3195

Mardia K.V., Kent J.T., Bibby J.M., 1979, Multivariate Analysis, Academic Press

North G.R., Kim K.K., 1995, J. Clim., 8, 409

North G.R., Stevens M.J., 1998, J. Clim., 11, 563

North G.R., Kim K.K., Shen S.S.P., Hardin J.W., 1995, J. Clim., 8, 401

Parker D., Gordon M., Cullum D., et al., 1997, Geo. Res. Lett., 24, 1499

Parker D.E., Jones P.D., Folland C.K., Bevan A., 1994, J. Geophys. Res, 99, 14373

Pope V.D., Gallani M.L., Rowntree P.R., Stratton R.A., 2000, Climate Dyn., 16, 123

Salby M., Callaghan P., 2000, 13, 2652

Santer B.D., Taylor K.E., Wigley T.M., et al., 1996, Nature, 382, 39

Sato M., Hansen J.E., McCormick M.P., Pollack J.B., 1993, J. Geophys. Res., 98, 22987

Schimel D., Alves D., Enting I., et al., 1995, In: Houghton J.T., Filho L.G.M., Callander B.A., et al. (eds.) Climate Change 1995. The Science of Climate Change, chap. 2, 65–131, Cambridge University Press, Cambridge

Shine K.P., Fouquart Y., Ramaswamy V., Solomon S., Srinivasan J., 1995, Climate Change 1994, chap. 4 Radiative Forcing, 168–203, CUP

Stott P.A., Tett S.F.B., 1998, J. Clim., 11, 3282

Stott P.A., Tett S.F.B., Jones G.S., et al., 2000, Cli. Dyn., (to appear)

Tett S.F.B., Stott P.A., Allen M.R., Ingram W.J., Mitchell J.F.B., 1999, Nature, 399, 569

Twomey S.A., 1974, Atmos. Environ., 8, 1251

Williams K.D., Senior C.A., Mitchell J.F.B., 2000, J. Climate, submitted

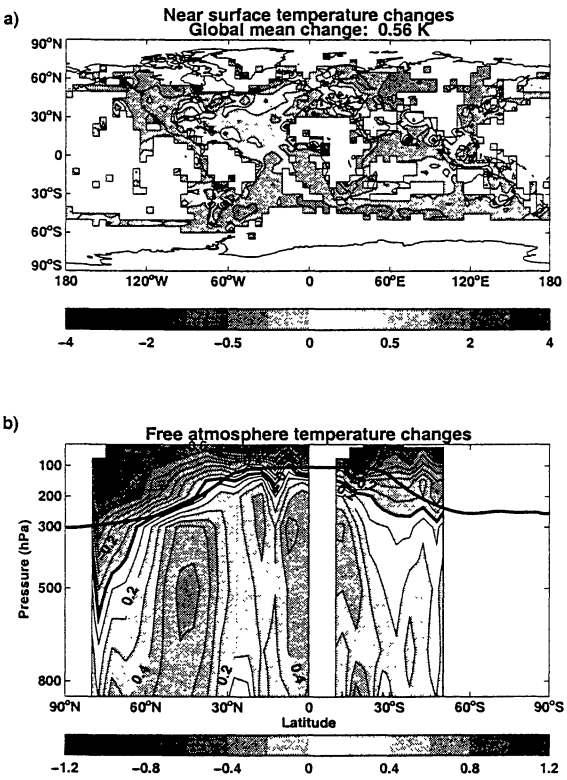

Figure 1. Observed temperature changes
a): Observed changes in near-surface temperature (1977–97 minus 1881–1920). A contour interval of 1K is used from -4K to 4K with additional contours at ±0.5K and ±0.25K.
b): Observed changes in zonal-mean temperature (1985–95 minus 1961–80). A contour interval of 0.1K is used with every second contour labelled from −1.2K to 1.2K. The black line denotes the zonal-mean position of the tropopause from the NCEP/NCAR reanalysis for the period 1985–95 using data provided by the NOAA-CIRES Climate Diagnostics Center, Boulder, Colorado, from `http://www.cdc.noaa.gov/`.

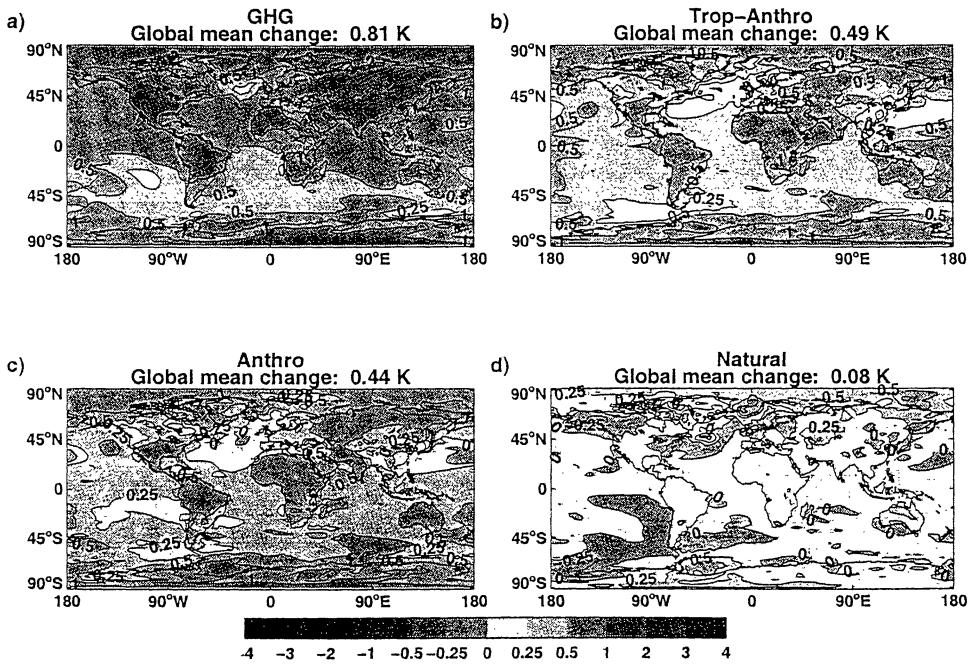

Figure 3. Simulated 20th century temperature changes.
Temperature difference (K) between the 20-year average 1977–1997 and the 40-year average 1881–1920 for the four ensembles, GHG (a), TROP-ANTHRO (b), ANTHRO (c), and NATURAL (d). Note these plots show the *raw* model data (i.e. without the observed mask). All other details are as Fig.1(a).

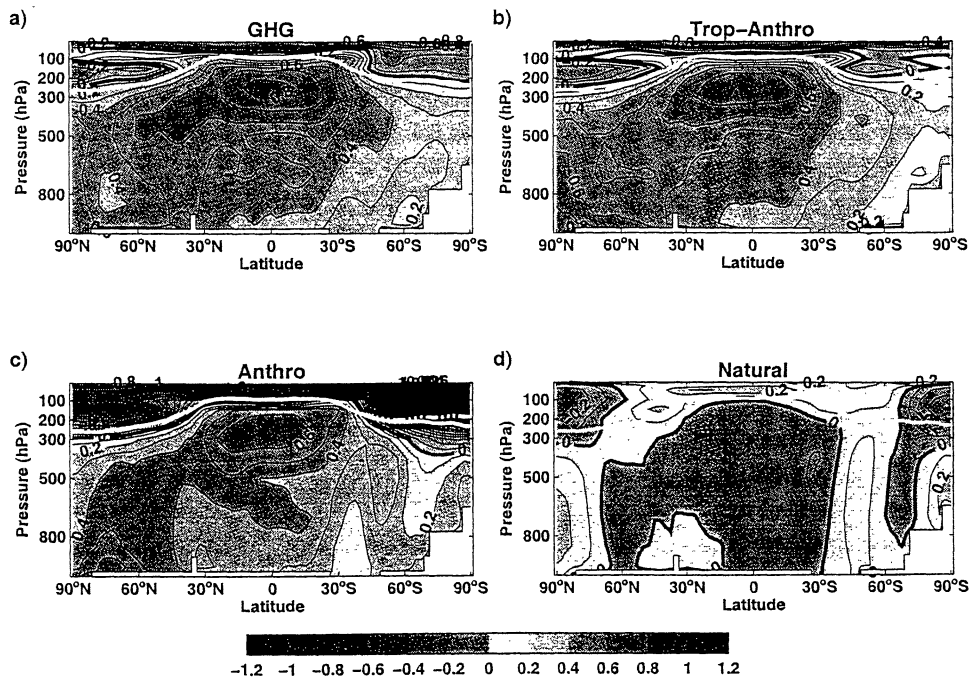

Figure 4. Simulated zonal-mean temperature differences
Differences (K) between 1985–1995 and 1961–1980 for the four ensembles: GHG (a), TROP-ANTHRO (b) ANTHRO (c) and NATURAL (d). White lines show the position of the mean tropopause in CONTROL while the dashed white line in (c) shows the mean position of the tropopause in a atmosphere only simulation with 1990 stratospheric ozone. The maximum difference between the two lines is approximately 50 hPa. All other details are as Fig. 1 (b).

Case	Period	Trunc.	% Var.	GHG	T-A	ANTHRO	NATURAL	ν_1	ν_2
Surface	1897–97	40	96.4	5.85	3.77	3.35	1.17*	40	27
(century)	1897–97	20	91.3	7.45	4.50	3.93	1.43	40	27
Free Atmos.	1961–95	7†	48.0	–	6.11	5.90	0.97*	36	42

Table 1. Signal properties

Shown for each analysis are the truncation used (third column), and the fraction of the observed variance (after processing) after filtering in the truncated eigenvector space (fourth column). 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. † denotes cases in which the truncation used is less than the largest possible.

The centre columns show the signal-to-noise ratio (SNR—see Section 4.6 for details) of the simulated signals. T-A (TROP-ANTHRO) is identical to ANTHRO before 1975. SNR values shown with a * are 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.

Shown in the right-hand columns are the estimated dof of C_N (ν_1) and C_{N_2} (ν_2).

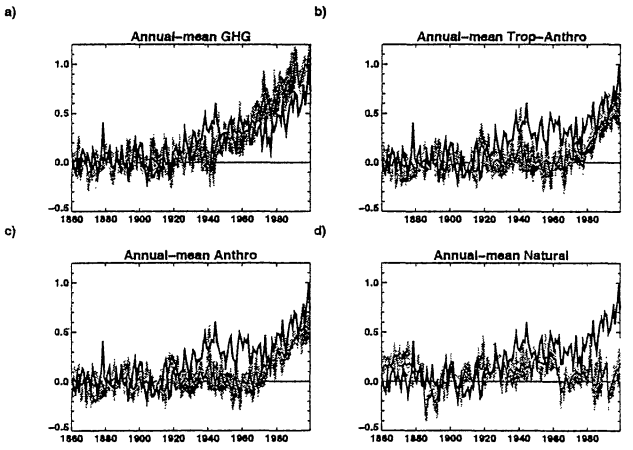

Figure 2. Global-mean near-surface temperature changes

Near-surface changes in global-mean temperature, relative to the 1881–1920 mean for the observations (thick black line) and the ensemble-mean of the GHG(a), TROP-ANTHRO(b), ANTHRO(c) and NATURAL(d) simulations (thin black line). The maximum and minimum range from the individual simulations is shown in gray.

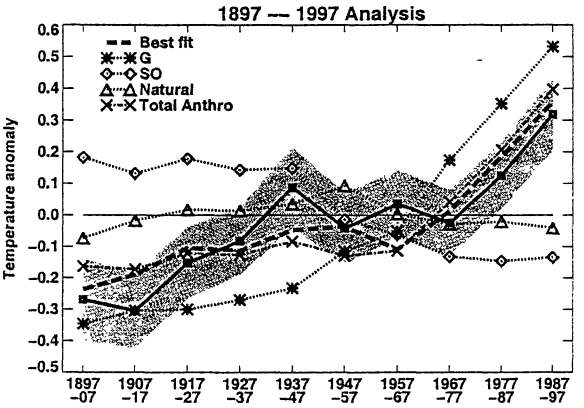

Figure 5. 100-year best-estimate reconstruction of near-surface temperature changes

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

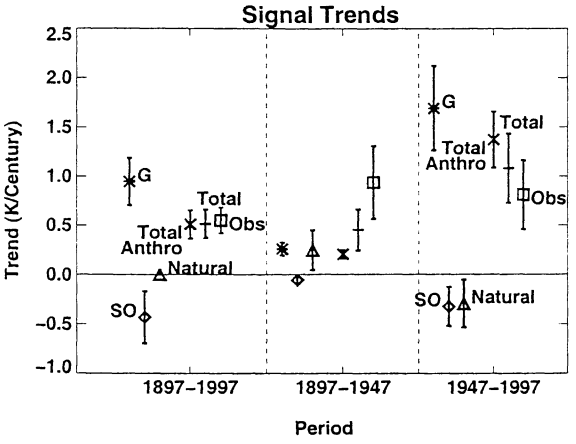

Figure 6. Linear trends from century analysis

Best-estimate linear trend and uncertainty ranges (K/century) for *G* (asterisk), *SO* (iamond), NATURAL (triangle), total anthropogenic trend (x), total trend (+) and observed trends (square). Symbols show best-estimate trend whilst error bars show the 5–95% uncertainty range inflated to allow for four member ensembles.

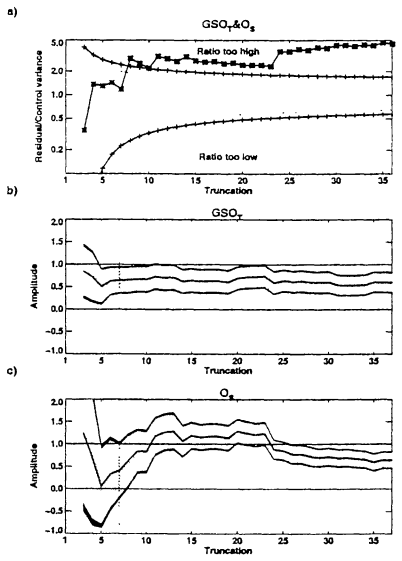

Figure 7. Sensitivity to truncation for free atmosphere analysis.
The ratio of the residual to the CONTROL variance (solid line with asterisks), using a logarithmic scale, is plotted in (a). Note that CONTROL variance has been inflated (see Section 4 for details). The vertical dotted line shows truncation seven – the largest truncation for which the residual and CONTROL variance are consistent.
Shown as a function of truncation are the best-estimate amplitudes (solid line), 5–95% "detection" uncertainties (light-gray shading), 5–95% "amplitude-consistency" uncertainties (thin black shading) for GSO_T (b) and O_S (c).

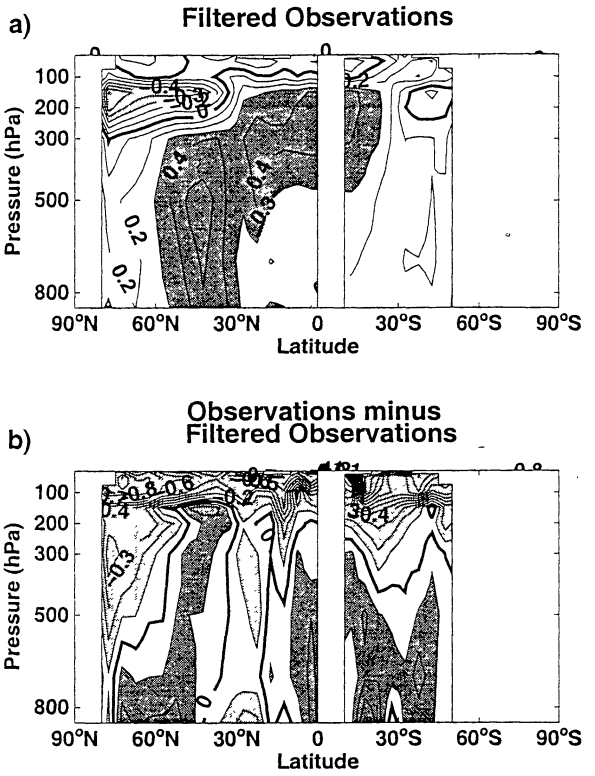

Figure 8. Filtered observations

a: Observed changes in zonal mean temperature filtered by projection onto the leading seven eigenvectors of C_N. A contour interval of 0.1 is used with dark (light) shading for values above (below) 0.3K (-0.3K) and the zero contour drawn bold.
b: Raw observations minus (a) (i.e. what the filtering removes). A contour interval of 0.1K is used with dark (light) shading for values above (below) 0.1K (-0.1K). The zero contour is drawn bold.