UCRL-JRNL-209353

LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

Detecting and Attributing External Influences on the Climate System: A Review of Recent Advances

T. Barnett, F. Zwiers, G. Hegerl, M. Allen, T. Crowley, N. Gillett, K. Hasselmann, P. Jones, B. Santer, R. Schnur, P. Stott, K. Taylor, S. Tett

February 2, 2005

Journal of Climate

Disclaimer

Detecting and Attributing External Influences on the Climate System: A Review of Recent Advances

By The International ad hoc Detection and Attribution Group[1]

Resubmitted to J Climate

14th September 2004

[1]Citation to this article should be to the International ad hoc Detection and Attribution Group (IDAG). The members of the IDAG who contributed to this article are Tim Barnett (group coordinator up to 2002, Scripps Inst. Ocean.), Francis Zwiers (Canadian Centre for Climate Modelling and Analysis, Meteorological Service of Canada), Gabriele Hegerl (Duke Univ.), Myles Allen (Univ. of Oxford), Tom Crowley (Duke Univ., group coordinator 2002 onwards), Nathan Gillett (Univ. Victoria), Klaus Hasselmann (Max Planck Inst. for Meteor.), Phil Jones (Univ. of East Anglia), Ben Santer (PCMDI), Reiner Schnur (Max Planck Inst. for Meteor.), Peter Stott (Hadley Centre, UK Met Office), Karl Taylor (PCMDI), and Simon Tett (Hadley Centre, UK Met Office).

Corresponding author: Francis Zwiers (e-mail: francis.zwiers@ec.gc.ca)

Abstract

We review recent research that assesses evidence for the detection of anthropogenic and natural external influences on the climate. Externally driven climate change has been detected by a number of investigators in independent data covering many parts of the climate system, including surface temperature on global and large regional scales, ocean-heat content, atmospheric circulation, and variables of the free atmosphere, such as atmospheric temperature and tropopause height. The influence of external forcing is also clearly discernible in reconstructions of hemispheric scale temperature of the last millennium. These observed climate changes are very unlikely to be due only to natural internal climate variability, and they are consistent with the responses to anthropogenic and natural external forcing of the climate system that are simulated with climate models. The evidence indicates that natural drivers such as solar variability and volcanic activity are at most partially responsible for the large-scale temperature changes observed over the past century, and that a large fraction of the warming over the last 50 years can be attributed to greenhouse gas increases. Thus the recent research supports and strengthens the IPCC Third Assessment Report conclusion that "most of the global warming over the past 50 years is likely due to the anthropogenic increase in greenhouse gases".

1. Introduction

The "International Ad Hoc Detection group" (IDAG) is a group of specialists on climate change detection, who have been collaborating on assessing and reducing uncertainties in the detection of climate change since 1995. Early results from the group were contributed to the IPCC Second Assessment Report (SAR; IPCC 1996). Additional results were reported by Barnett et al. (1999) and contributed to the IPCC Third Assessment Report (TAR; IPCC 2001). The weight of evidence that humans have influenced the course of climate during the past century has accumulated rapidly since the inception of the IDAG. While little evidence was reported on a detectable anthropogenic influence on climate in IPCC (1990), a "discernible" human influence was reported in the SAR, and the TAR concluded that "most of the observed warming over the last 50 years is likely to have been due to the increase in greenhouse gas concentrations". The evidence has continued to accumulate since the TAR. This paper reviews some of that evidence, and refers to earlier work only where necessary to provide context.

Climate change detection assumes that climate (meaning the statistical characteristics of our atmospheric, oceanic and cryospheric environment) is implicitly predictable in the sense that if a (known) change in external forcing occurs, the climate will respond by displaying a predictable change in its statistical characteristics. This should hold even if the climate displays "regime-like" behavior, as is characteristic of many chaotic systems (Palmer 1999a), because regime occupancy characteristics are part of the full description of the behaviour of the climate system. If changes occur in these occupancy characteristics, then we expect that they will be reflected in the statistics (means, variances, auto- and cross-covariances, and higher order moments) that characterize the climate. A key statistical characteristic of the climate system is its mean state; an aspect of climate that we assume can be changed in predictable ways by both natural and anthropogenic external influences. For example, climate is expected to respond to aerosols ejected into the stratosphere by strong volcanic eruptions, to variations in solar irradiance, to changes in greenhouse gases and to changes in the composition of the atmosphere associated with human activity, particularly the burning of fossil fuel. Similarly, we assume that if external influences cause changes in the climate's variability, then the characteristics of those changes (e.g., a change in regime occupation frequency) will also occur in predictable ways.

Evidence from coupled global climate models (CGCMs; oceanic and atmospheric general circulation models that are coupled together with land surface and cryospheric components), together with our rapidly increasing understanding of the role of external forcing in paleoclimates, suggests that this assumption is well founded. Thus one of the main goals of detection and attribution research during the past several years has been to compare observed changes in climate, primarily during the past century, against CGCM simulations that have been forced with estimates of historical changes in anthropogenic and natural external forcing.

The most easily obtainable evidence of externally forced change has come from global-scale analyses of the combined instrumental surface air temperature and sea-surface temperature records (e.g., Jones et al. 1999; Jones and Moberg 2003). This record, which extends into the 19th century, is well suited for climate change research because it is of high quality and has broad spatial coverage. It has been extensively scrutinized (e.g., Folland et al. 2001b), and is expected to exhibit the response to external forcing with high signal-to-noise ratio. Therefore, climate change detection research initially relied mainly on the surface temperature record.

While temperature variables continue to be investigated in order to better understand and reduce uncertainty, investigation is now proceeding with several other climate variables. Some recent studies have also begun to assess whether the climate response to external forcing is detectable on regional scales. In addition, some investigators are now evaluating the prospects of detecting externally forced change in the frequency and intensity of climatic extremes. Both of these developments are important because, ultimately, policy makers will be most strongly influenced by evidence of impacts in regions that are of direct interest to them, and by evidence that anthropogenic greenhouse gas emissions are having an influence on the occurrence of high impact climate events such as heat waves and flooding.

The plan for the remainder of this paper is as follows. Section 2 briefly reviews the detection techniques that have been used in recent research. Our consideration of new scientific developments begins in Section 3 with a description of the considerable progress that has been made in improving our understanding of the climate of the last millennium and the external factors that have influenced its variability. Sections 4, 5 and 6 then review advances that are based on the instrumental surface temperature record (Section 4), free atmosphere temperature and circulation records (Section 5), and oceanic records (Section 6). Section 7 deals briefly with rainfall and climate extremes, and Section 8 discusses some recent progress with the use of Bayesian methods. We complete the paper with a summary of our main findings in Section 9.

2. Methodological considerations

Any discussion on the methodology that is used for detection and attribution should begin with an understanding of these terms. The definitions we use are those given by Mitchell et al. (2001) in the TAR (IPCC 2001). Quoting from that report, "*Detection* is the process of demonstrating that an observed change is significantly different (in a statistical sense) than can be explained by natural internal variability" where natural internal variability is the chaotic variation of the climate system that occurs in the absence of anomalous external forcing. Detection does not immediately imply attribution of the cause of the detected change. As noted in the SAR (IPCC 1996) and the TAR unequivocal attribution would require controlled experimentation with our climate system. That, of course, is not possible, and thus from a practical perspective, attribution of anthropogenic climate change is understood to mean (a) detection as defined above, (b) demonstration that the detected change is consistent with a combination of external forcing including anthropogenic changes in the composition of the atmosphere and natural internal variability, and (c) that it is “not consistent with alternative, physically-plausible explanations of recent climate change that exclude important elements of the given combination of forcings” (IPCC 2001).

In this section we very briefly review the statistical methods that have been used in recent detection and attribution work. Two statistical approaches have been used in recent studies. Standard ‘frequentist’ methods (methods based on the relative frequency concept of probability) continue to predominate, but there is increasing interest in the use of Bayesian methods of statistical inference. One reason is that information from multiple lines of evidence can be combined in the Bayesian framework. We will briefly review the optimal fingerprinting technique in the following subsection. This will be followed by a short discussion on the differences between the standard and Bayesian approaches to statistical inference that are relevant to detection and attribution.

a. Optimal fingerprinting

Optimal fingerprinting is generalized multivariate regression that has been adapted for the detection of climate change and the attribution of change to externally-forced climate change signals (Hasselmann 1979, 1997; Allen and Tett 1999). The multiple regression model that is used has the form $\mathbf{y} = \mathbf{X}\mathbf{a} + \mathbf{u}$ where vector $\mathbf{y}$ is a filtered version of the observed record, matrix $\mathbf{X}$ contains the estimated response (signal) patterns to the external forcings that are under investigation, $\mathbf{a}$ is a vector of scaling factors that adjusts the amplitudes of those patterns, and $\mathbf{u}$ is a realization of internal climate variability. Vector $\mathbf{u}$ is assumed to be a realization of a Gaussian random vector with a covariance matrix $\mathbf{C}$. Vector $\mathbf{a}$ is estimated with $\mathbf{a} = (X^T C^{-1} X)^{-1} X^T C^{-1} y = (\tilde{X}^T \tilde{X}) \tilde{X}^T \tilde{y}$ where matrix $\tilde{X}$ represents the signals patterns after normalization by the climate's internal variability, and vector $\tilde{y}$ represents the observations after normalization. The normalizations transform the signals and observations so as to maximize the signal-to-noise ratio (see Mitchell et al 2001, and references cited therein).

The matrix **X** typically contains signals that are estimated with either a CGCM, an atmospheric general circulation model (AGCM; see Sexton et al. 2001, 2003) or a simplified climate model such as an energy balance model (EBM). Because CGCMs simulate natural internal variability as well as the response to specified anomalous external forcing, the GCM simulated climate signals are typically estimated by averaging across an ensemble of simulations (for a discussion of optimal ensemble size and composition, see Sexton et al., 2003). By allowing us to scale the signal patterns to best match the pattern of change that is contained in the observations, the vector of scaling factors $\mathbf{a}$ accounts for the possibility of error in the amplitude of the anomalous external forcing, and for the possibility that the amplitude of the climate model response to the forcing may not be correct.

Fitting the multiple regression model requires an estimate of the climate's natural internal variability (i.e., the covariance matrix $\mathbf{C}$). The instrumental record is not long enough to provide a reliable estimate and may also be contaminated by the effects of external forcing. Thus long control simulations with CGCMs (i.e., without anomalous external forcing) are typically used for this purpose. It is understood that CGCMs may not simulate natural internal climate variability accurately, particularly on small spatial scales, and thus a residual consistency test (Allen and

Tett 1999) is typically used to assess the model simulated variability on the scales that are retained in the analysis. It is also recognized that the uncertainty of the estimate of the vector of scaling factors $\mathbf{a}$ should be assessed with a second, statistically independent estimate of the covariance matrix $\mathbf{C}$. This second covariance estimate is typically obtained from an additional, independent control simulation.

Signal estimates obtained with CGCMs contain remnants of the climate's natural internal variability even though they are obtained by averaging across an ensemble of forced climate change simulations. The presence of this noise in the signal may bias ordinary least squares estimates of $\mathbf{a}$ downward, particularly if only a small ensemble is available to estimate signals that have small signal-to-noise ratios (as is the case in the 20th century). Thus several recent studies that use CGCM derived signals have estimated $\mathbf{a}$ with the total least squares algorithm (Allen and Stott 2003).

There is considerable variation in the details of the implementation of the optimal fingerprinting approach, and in the way data are processed prior to its application. However, recent research has shown that different approaches to detection and attribution yield consistent results. Methodological aspects that have been investigated include the use of stepwise versus multiple regression (Hegerl and Allen 2002), various data-treatment methods (Gillett et al. 2002a) and the use of signals and noise estimates constructed from multiple models (Gillett et al. 2002b). This demonstration of consistency and robustness of results increases our confidence in the detection of anthropogenic climate change.

b. Methods of Inference

Detection and attribution questions are assessed through a combination of deductive reasoning (to determine whether there is evidence that other mechanisms of change not included in the climate model could plausibly explain the observed change) and by evaluating specific hypotheses on the scaling factors $\mathbf{a}$. The hypotheses continue to be assessed most often using standard methods (Hasselmann 1979, 1997; Hegerl et al. 1997; Allen and Tett 1999; Allen et al. 2004). However, some researchers are now also beginning to use Bayesian methods (Hasselmann 1998; Leroy 1998; Berliner et al. 2000; Schnur and Hasselmann 2004; Lee et al. 2004).

i) Standard approach

In the standard approach, detection of a postulated climate change signal occurs when its amplitude in observations is shown to be significantly different from zero. This is handled by testing the null hypothesis $H_D : \mathbf{a} = \mathbf{0}$ where $\mathbf{0}$ is a vector of zeros. The second attribution requirement (consistency with a combination of external forcings and natural internal variability) is assessed with the assistance of the *attribution consistency test* (Hasselmann 1997; see also Allen and Tett 1999) which evaluates the null hypothesis $H_A : \mathbf{a} = \mathbf{1}$ where $\mathbf{1}$ denotes a vector of units. Consistency between the observed and climate model simulated response to forcing (i.e., a finding that there is insufficient evidence to reject H_A) lends support to an attribution assessment, but does not on its own provide strong evidence in support of attribution (Berliner et al. 2000; Lee et al. 2004). A complete attribution assessment would take into account not just evidence from this test, but would also account for competing mechanisms of climate change as completely as possible, as discussed in Mitchell et al. (2001).

ii) Bayesian approach

Interest in a Bayesian approach is motivated by several factors. These include the ability to integrate information from multiple lines of evidence, and the ability to incorporate independent prior information into the analysis. Two distinct approaches to Bayesian detection and attribution have been taken to date. These are exemplified by Hasselmann (1998) and Schnur and Hasselmann (2004) on the one hand, and Berliner et al. (2000) and Lee et al. (2004) on the other. In both cases inferences are based on a posterior distribution that blends evidence from the observations with independent prior information that is represented by a prior distribution. This ability to incorporate prior information, which may include information on the uncertainty of external forcing estimates, climate models, and their responses to forcing, is a strength of the Bayesian approach even though some of the prior information may be subjective. This is because all information that enters into the analysis is declared explicitly. Another strength is that Bayesian inferences are probabilistic (i.e., based on the posterior likelihoods of detection and attribution), which means that they can better feed into decision making processes that balance risks and benefits. Also, the Bayesian approach provides a more satisfactory inference on attribution by assessing the likelihood of attribution consistency.

Schnur and Hasselmann (2004) approach the problem by developing a filtering technique appropriate for the Bayesian method that optimizes the impact of the data on the prior in a manner similar to the way in which optimal fingerprints maximize the ratio of the anthropogenic signal to natural variability noise in the conventional approach. The optimal filter in the Bayesian approach depends on the properties of both the natural climate variability and model errors. In contrast, Berliner et al. (2000) and Lee et al. (2004) use an approach that does not optimize the impact of the data on the prior distribution. Instead, they use Bayesian methods only to make inferences about the estimate of $\mathbf{a}$ that is obtained from a conventional optimal fingerprinting approach.

3. Analysis of paleo-climate reconstructions

Both instrumental measurements and proxy data can be used to investigate climate change and climate variability of Earth's recent past. However, prior to the mid-19th century, only European instrumental data are adequate to extend temperature time series back to about 1750. Before then it is necessary to use proxy records to estimate temperatures. Although the proxy evidence is less reliable than instrumental data (e.g., Jones et al. 2001; Esper et al. 2002), estimates for earlier centuries (particularly the last millennium) are vital as they enable the last 140 years to be placed in a broader context. These records also provide estimates of the range of variability on decadal to century time scales that can occur naturally due to external forcing from solar output changes and explosive volcanism, and due to internal variability of the climate system.

a. The last millennium

Compilations of proxy records developed during the last few years clearly show that the earth has warmed rapidly in the last 100 years or so (see the reviews by Jones et al. 2001, and Jones and Mann 2004). There is surprising agreement between the temperature time series developed by different authors (Figure 1a) even though the different reconstructions generally use different proxies and represent somewhat different variables. Data in some reconstructions cover only the mid-to-high latitudes. Some reconstructions are representative only of the growing season while others reflect temperature variations in all seasons. Some are representative only of land areas, while others represent both land and ocean areas. Most reconstructions are only available to the 1970s and 1980s, beyond which many proxies are not available.

The reconstructions show that average temperatures during the last two decades of the 20th century, which were 0.2 K warmer than the 1961 to 1990 average, were probably the warmest of the last millennium. The 1980s and 1990s are estimated to have been about 0.2 K warmer than the warmest decadal periods in the 11th and 12th centuries. The first half of the millennium was generally cooler than the 20th century mean, but milder than the 1500 to 1900 period. The coolest century was the 17th followed by the 19th, separated by a milder 18th century. The warming during the 20th century, which is approximately 0.6 K (IPCC 2001), is considerably larger than the variation of warming/cooling estimates of approximately ±0.2 K within each century of the rest of the millennium that are obtained from paleo reconstructions. Thus, proxy-derived series suggest that 20th century warming is unique in the last millennium for both its mean value and its rapidity of change. Both the early and late decades of the 20th century are clearly unusual in a millennial context.

Attribution results generally indicate that the early 20th century warming can be explained as a combination of a signal from natural forcing and an emerging greenhouse gas signal of similar magnitude, possibly in combination with natural climate variability. However, the contribution from individual natural forcings and internal variability remain somewhat controversial. Some studies indicate that solar forcing dominates (e.g. Stott et al. 2003), while others find a significant greenhouse gas signal (Tett et al., 2002; Hegerl et al., 2003) and a warming due to a cessation of volcanism (Hegerl et al. 2003) or a large contribution from internal climate variability (Tett et al. 2002, Delworth and Knutson 2000).

The proxy-based view of temperature change over the last 500 years was challenged by reconstructions of past surface temperatures derived from boreholes (Huang et al. 2000). Boreholes represent the temperature of the ground below the land/atmosphere interface, which integrates the temperature variations of the overlying atmosphere that diffuse into depth. Boreholes therefore predominately provide information on long timescales. Borehole temperature readings suggest that Northern Hemisphere surface temperatures may have warmed by over 1 K since 1500, compared to the conventional proxy estimate of about 0.5 ± 0.2 K (Figure 1a). Part of the discrepancy is due to the seasonal differences in the response of different proxy series. Tree ring data are more weighted towards growing season temperatures, while boreholes tend to represent mean annual data. Questions have also arisen about possible effects of snowcover on the borehole reconstructions (see e.g. Mann and Schmidt, 2003). Area weighting of the 453 Northern Hemisphere borehole series that were available, rather than simple averaging, reduces the warming since 1500 by 0.2 K. Recently studies (see Mann et al. 2003, Pollack and Smerdon 2004, and Rutherford and Mann 2004; Huang 2004) suggest that if the borehole records are recalibrated with instrumental data over the 20^{th} century, they can be shown to support the conventional proxy view of the millennium. Nevertheless, further work should go into fully reconciling these two proxies. Like traditional proxy data, many borehole records were collected as early as the 1970s and 1980s, so will not fully represent recent late-20th century warming. A clear resolution to the issue is most likely to come from an extensive updating of both traditional and borehole proxy data.

Despite the differences, the borehole/conventional proxy issues have served paleoclimatology well by forcing a reassessment of exactly what part of the year a proxy represents. It also demonstrates that it is important to process proxy data in a manner that takes their specific characteristics and coverage into account, and to include estimates of uncertainty.

b. Paleoclimate forcing and modeling studies

Substantial progress has been made over the last few years in refining paleo-proxy forcing time series and consequently, in interpreting paleoclimate reconstructions with simple energy balance climate models (EBMs) that are driven with these forcing time series (c.f., IPCC 2001). These results have been confirmed using improved forcing and paleo reconstruction time series (Crowley et al. 2004; Crowley 2004) using a more detailed EBM with geographic resolution in the horizontal domain and a seasonal cycle. Results confirm previous conclusions that about 50% of the decadal pre-anthropogenic Northern Hemisphere temperature variance can be attributed to a direct response to solar and volcanic variability (Crowley et al. 2004). A small additional amount of variance is explained by the small decrease in CO_2 concentration in the 17th and 18th century, which is thought to reflect a combination of the ocean response to Little Ice Age cooling and possibly decreased land respiration (Joos et al. 1999). Almost 60% of the decadal temperature variance in the full proxy record (1005-1960, and up to 77% from 1400 on; see Figure 2) can be explained by external forcing (solar, volcanic, greenhouse gas, and sulphate aerosol in the 20th century). Additional experiments employing land-use changes suggest that some of the late 19th century model-data discrepancies reported in Crowley (2000) can be eliminated by specifying known land-use changes (Bertrand et al. 2002; Bauer et al. 2003). Presently, a number of CGCM simulations of the last millennium are becoming available (e.g., Tett et al. pers. comm.; Zorita et al. 2003) in addition to simulations of selected periods of the past (e.g. Rind et al. 2004). These simulations, in conjunction with the increasing availability of proxy reconstructions, should improve our understanding of the origins of past climate change (see also Jones and Mann, 2004). They can also be used to test reconstruction methods and estimate uncertainties in the shape and variance of reconstructed temperature signals (see, for example, Zorita et al., 2003).

Multiple regression can be used to diagnose the influence of external forcing on climate variability in the last millennium as represented by a range of paleo reconstructions (Hegerl et al. 2003). The response to volcanism and greenhouse gases can be clearly detected in most records (Briffa et al. 2001; Crowley et al. 2004; Mann et al. 1999; Esper et al. 2002) and can be distinguished from each other and solar forcing with only small differences in results for different records. The response to solar forcing is detectable only in some periods and some records, although the EBM simulation of solar forcing is not inconsistent with the records. It is possible that errors in the forcing history, particularly early in the record, may have obscured the solar signal. Low-frequency solar forcing is quite uncertain (Lean et al. 2002). The anthropogenic signal can be detected and distinguished from other forcings in all paleo reconstructions towards the end of the 20th century, and is detectable in many reconstructions by the middle of the 20th century (Hegerl et al. 2003).

Simulations of the last few hundred years can also be used to estimate changes in global ocean heat storage (Figure 3). Using a comparison between model simulation and the surface proxy data set for a best-fit sensitivity and a plausible estimate of ocean diffusivity, substantial changes in ocean-heat storage are postulated during the Little Ice Age and a warming that commenced in the mid-19th century (Crowley et al. 2004). The first phase of the ocean-heat storage increase represents, in part, a relaxation after an intense period of volcanism, and in part, anthropogenic greenhouse gas concentration increases due to deforestation and early industrialization. The more recent increases in ocean heat content, which agree very well with the Levitus et al. (2001) reconstruction, are driven mostly by the anthropogenic forcing.

4. Surface Temperature

Averaged globally, surface temperature is estimated to have increased by 0.6 K between the late 19th century and the end of the 20th century (Folland et al 2001a, b) with a 95% confidence bound, taking into account all known systematic and sampling errors in the basic land and marine data, that is estimated to be ±0.2 K. In the following subsections we will briefly discuss the instrumental surface temperature record, some recent global and regional detection and attribution studies, and results from some studies that have used the observational record to constrain estimates of key climate parameters and future temperature change.

a. Data

Land and marine temperature data archives evolve continually, and thus the land surface temperature database has recently been enhanced by Jones and Moberg (2003). Concurrently, work is underway in many countries to improve data quality and extend data availability (particularly at the daily timescale), including a very substantial effort in the United States to provide digital access to their 18th and 19th century records. Such data rehabilitation efforts are also required in the developing world, particularly in Central and South America, Africa and southern Asia, to fill large gaps in the record. Improvements in availability of marine temperature data are also expected in the next few years, particularly for the 19th century and the two world war periods (Diaz et al. 2002).

While it is very important to continue to fill current data voids to better define and monitor regional climate change, it is unlikely that additional data will significantly alter our estimates of changes in the global mean temperature. However, additional data will further reduce the uncertainty of these estimates. Improvements in other aspects of the instrumental record (such as daily series of pressure, precipitation, cloudiness, sunshine and water vapor) are also anticipated and will help to produce a more complete picture of climate change from the mid-19th century onwards.

b. Global scale detection and attribution results

The IPCC TAR concluded that “... most of the warming observed over the last 50 years is likely to have been due to the increase in greenhouse gas concentrations” (IPCC 2001). This assessment was made with a very broad range of evidence, including evidence from a number of optimal detection studies using several different coupled climate models. Subsequent research has further increased the pool of evidence that supports the IPCC conclusion. For example, an optimal fingerprinting study that uses climate change signals estimated from an array of climate models produced results that are broadly consistent with the TAR (Allen et al. 2004).

Despite the qualitative similarity in detection results, estimates of the magnitude of the contributions of individual anthropogenic forcing agents to the observed warming remain sensitive to which model was used to estimate the climate change signal and the natural internal climate variability. This sensitivity is partly due to differences in forcing and response to sulphate aerosols (Hegerl and Allen 2002), and to forcing mechanisms that were not included in all models used in the TAR (such as the indirect effect of sulphate aerosols or the effect of the changing ozone distribution). Responses attributed to natural external climate forcing are also model dependent. Thus it is important to explicitly include ‘model uncertainty’ in detection and attribution assessments. Different choices in the implementation of the optimal detection formalism, such as the use of annual or summer season data, the choice of base period for calculating anomalies, and the use of a time-evolving signal pattern as opposed to fixed spatial trend patterns cause only minor differences in results that can be fully accounted for by these choices (Gillett et al. 2002a).

Gillett et al. (2002b) recently published a first attempt to overcome model uncertainty by using a multi-model approach to detection. Detection results from five models (HadCM2, HadCM3, CGCM1, CGCM2 and ECHAM3) were synthesized using the mean response patterns as fingerprints in a detection of greenhouse gas and sulfate aerosol influence, including an estimate of model uncertainty. Results indicate that the inter-model differences do not greatly increase detection and attribution uncertainties as applied to temperature data, and that averaging fingerprints actually improves detection results. This statement does not preclude the possibility that systematic errors common to all models may still be important, and inter-model differences may be much more important in the detection and attribution of change in other variables such as precipitation (Allen and Ingram 2002; Hegerl et al. 2004) or regional temperature.

c. Regional results

The impact of global warming on society is largely determined by spatial scales far smaller than those considered in detection and attribution studies to date. Thus detection and attribution research is beginning to focus on sub-global scales. Two approaches are being used in this work – one to assess the extent to which global studies can provide information on sub-global scales, and another to assess the influence of external forcing on the climate in specific regions.

Our approach to the former is to consider a global data set from which we have systematically removed global mean information. The result is shown in Figure 4a-c (after Allen et al. 2004). The detection diagram with full global mean information included is shown in Figure 4a (corresponding to Figure 12.12 in the TAR). The combined greenhouse gas and sulphate aerosol ("GS") signal is detected in observations regardless of the model that is used to estimate the signal, and the estimated scaling factors on the GS signals are generally consistent with unity. Estimates of the contribution to observed warming from GS forcing are very similar regardless of the model used to estimate the GS signal. The estimated magnitude and detectability of aerosol and other anthropogenic and natural signals in multi-signal analyses is less consistent between fingerprints from different models. However, a greenhouse gas signal is detected consistently if other signals are explicitly considered, providing strong evidence for the attribution of a part of the observed warming to greenhouse gas increases.

Similar results are obtained when the linear trend in the global mean is removed from observations (Figure 4b), although uncertainties on the scaling factors, or equivalently, the estimated contribution to 20th century warming, is increased. Figure 4c shows the result that is obtained when all global mean information is removed from the GS signal estimates. We see that the uncertainty of the scaling factors is further increased, and that detection of the space-time evolution of the estimated GS pattern of global change occurs less consistently with the different GS signal estimates. However, detection still occurs at the 5% (one-sided) significance level, or lower, in four out of nine cases, and with only slightly weaker significance level, in 7 out of nine cases. Note that the increase in uncertainty is expected when global mean information, which has a high signal-to-noise ratio, is disregarded. Figures 4b and 4c provide evidence that the detection of anthropogenic climate change is also driven by the pattern of the observed warming in space and time, not only by consistent global mean temperature trends between models and observations. This suggests that greenhouse warming should also be detectable on subglobal scales.

A related approach applies indices that reflect features of the anticipated response to anthropogenic forcing. Such indices might include the global-mean surface temperature, the land-ocean temperature contrast, the magnitude of the annual cycle in surface temperature over land, the Northern Hemisphere meridional temperature gradient and the hemispheric temperature contrast (Braganza et al. 2004, see also Karoly and Braganza 2001, and Karoly et al. 2003). Braganza et al. (2004) find changes in these temperature indices during 1950-and 1999 that are statistically significant, and similar to those simulated in anthropogenic climate change simulations. An attribution analysis suggests that the anthropogenic forcings account for almost the entire temperature change in the analysis period, while the early 20th century change is explained by equal contributions from anthropogenic and natural forcing, and a contribution from internal climate variability.

Another approach for assessing the regional influence of external forcing is to apply detection and attribution formalisms to observations in specific regions (e.g., Zwiers and Zhang 2003, and Stott 2003). Zwiers and Zhang (2003) assess the detectability of the GS signal as estimated by the Canadian Centre for Climate Modelling and Analysis (CCCma) CGCMs in a series of nested regions, beginning globally and descending to separate continental domains for North America and Eurasia. They find that observations in both domains taken during the latter half of the 20th century contain evidence that their climates have been influenced by anthropogenic emissions (Figure 5). This finding is robust to the exclusion of North Atlantic Oscillation (NAO/AO) related variability (Thompson and Wallace 1998), which may itself be related to anthropogenic forcing (see below; also see Gillett et al. 2000, 2003b, Fyfe et al. 1999, Shindell et al. 1999).

Stott (2003) assesses the detectability of the response to natural and anthropogenic forcing as simulated by HadCM3 in six continental scale regions, each composed of a small number of subregions. For each continental scale region, greenhouse warming can be detected and separated from the effect of natural forcing and sulfate aerosol and ozone forcing. The ability of HadCM3 to reproduce observed change in these regions, which is illustrated in Figure 6, is compelling evidence of the influence humans have probably had on regional climates. Stott reports detection of an anthropogenic signal in all continental regions considered, although HadCM3 appears to have overestimated sulphate aerosol cooling in some regions (notably Asia).

North American continental scale temperature change has also been analyzed by Karoly et al. (2003) in a similar way as in the global study of temperature indices by Braganza et al (2004) mentioned above. The indices reflect several large-scale aspects of North American temperature change, namely the regional mean, the mean land-ocean temperature contrast, an estimate of the mean meridional temperature gradient, an estimate of the mean annual cycle, and an estimate of the regional mean diurnal temperature range. The variability of the indices as simulated by five CGCMs was compared with estimates from observations, and observed changes in the 20th century were compared with those simulated in response to historical natural and anthropogenic forcing. Confounded variability caused by changes in instrumental coverage over time was controlled by applying a fixed observational data mask to both observations and model. The model-simulated decadal variability in the indices was similar to that estimated from observations, with the possible exception of meridional temperature gradient variability, which tends to be greater in models than in observations. Simulated trends in the indices were found to be consistent with observed trends under historical GS forcing (Figure 7), while there appears to be only a small likelihood of agreement with trends driven by natural forcing only for both the entire 20th century and the second half of the 20th century.

d. Observational constraints on future temperature change

By validating climate models against observations, detection and attribution studies such as those described above provide observational constraints that can be used to reduce uncertainty in model-based predictions. This approach has been used to provide objective forecasts of global temperature over the coming decades (Allen et al. 2000; Stott and Kettleborough 2002; Allen and Stainforth 2002) that are less sensitive to the choice of model and insensitive to details of the external forcing scenarios used to drive the models.

Detection formalisms can also be used to obtain observational constraints on key climate model parameters (such as the climate sensitivity to radiative forcing and ocean diffusivity) by determining parameter settings that yield simulations that are consistent with observations. Such studies have generally been performed with EBMs or a model of intermediate complexity because of cost considerations, although some studies are now underway with more complex climate models (e.g., see http://www.climateprediction.net/index.php). Results using 20th century instrumental data (Forest et al. 2000, 2002; Andronova and Schlesinger 2000; Gregory et al. 2002a, Knutti et al. 2003) are in broad agreement with each other and generally yield a skewed distribution for climate sensitivity with peak values around 2-3 K for CO_2 doubling and a wide tail to the right. The 10-90% uncertainty range on climate sensitivity is typically around 1.8-6.5 K (Forest et al. 2002). Since aerosol forcing is a major uncertainty in simulating 20th century climate change, all studies either make assumptions about total aerosol forcing (e.g. Forest et al. 2000), allow for a range of total aerosol forcing (Forest et al. 2002) or apply an inverse calculation of the sulfate forcing given the scaling for sulfate aerosol and greenhouse gas forcing (Gregory et al. 2002a; Knutti et al. 2003;. Anderson et al 2003). If reconstructions of Northern Hemispheric mean temperature from paleo data are used, it may be possible to further reduce the upper bound of this uncertainty range (G. C. Hegerl et al 2004, pers. com.). This is consistent with Hansen et al (1984) who looked at forcing and temperature at the last glacial maximum and concluded that the climate sensitivity associated with processes on the decadal to century scale probably does not exceed 5 K.

5. Free Atmosphere

a. Temperature

One of the major remaining puzzles in our understanding of the causes of late 20th century climate change relates to the apparent “differential warming” of the surface and troposphere. Surface thermometer measurements indicate that the Earth’s near-surface temperatures have warmed at a rate of 0.15 to 0.20°C/decade since 1979. The reality of this recent surface warming has been confirmed by numerous investigations (see, e.g., Jones et al., 1999; National Research Council, 2000). In contrast, radiosondes and the satellite-based Microwave Sounding Unit (MSU) record developed at the University of Alabama at Huntsville (UAH) show little or no tropospheric warming over the past 25 years (Parker et al. 1997 and Christy et al. 1998). This contrasts with model simulations of the response to anthropogenic forcing, which generally show substantial tropospheric warming over this period (e.g. Santer et al. 2001, and Tett et al. 2003).

Radiosonde measurements provide a longer-term perspective on surface and tropospheric warming rates. They show that the tropical troposphere warmed relative to the surface over 1960-1978, and thereafter cooled relative to the surface (Gaffen et al. 2000). Analyses of radiosonde data by Angell (2000) and others have concluded that there is no discrepancy between the overall warming rates at the surface and in the lower troposphere for the period 1958-1998. As noted above, this consistency between surface and tropospheric warming rates appears to break down over the shorter, MSU era.

A number of recent studies have sought to understand the possible causes of differential surface and tropospheric warming rates in models and observations. Santer et al. (2000) showed that spatial differences in coverage between satellite data, which are globally complete, and surface observations, which are spatially incomplete, could explain up to one-third of the difference between surface and lower tropospheric temperature (2LT) trends over 1979-1999. External forcing also helps to reconcile some of the differential warming, since both volcanic eruptions and stratospheric ozone depletion may have cooled the troposphere by more than the surface over the last several decades. This can be demonstrated in model simulations (IPCC 2001; Santer et al. 2000) and from observational studies (Santer et al. 2001; Free and Angell 2002). There are, however, substantial uncertainties in quantifying the differential cooling caused by these forcings. These arise primarily from uncertainties in our estimates of historical forcings, and from errors in the model responses to these forcings.

Several recent investigations have also assessed the differential effects of natural modes of variability (such as ENSO and the AO/NAO) on observed surface and tropospheric temperatures. These differential effects arise from differences in the amplitudes and spatial expression of these modes at the surface and in the troposphere, and make only minor contributions to the overall differences in observed surface and tropospheric warming rates (Santer et al. 2001; Hegerl and Wallace 2002).

Accounting for all these effects (spatial coverage, external forcing, and natural variability) cannot fully explain the apparent differential warming of the surface and troposphere in observations, or why models do not replicate this differential warming (Santer et al. 2001).

It is possible that observational error may explain these remaining discrepancies. A recently completed reprocessing of the MSU channel 2 temperature data (Mears et al. 2003) yields an estimated tropospheric temperature trend over 1979-1998 that is roughly 0.1 K per decade warmer than the trend in the Christy et al. (2000) version of the channel 2 data (Figure 8). This discrepancy is primarily related to differences in the estimated calibration coefficient for the

MSU instrument on the NOAA-9 satellite. It also arises from differences in how both groups account for the effects of satellite orbital drift on the sampling of the diurnal temperature cycle.

If the Mears et al. (2003) analyses of the satellite data is more reliable, then there is no serious inconsistency between modeled and observed tropospheric temperature trends (Figure 9; Santer et al., 2003a). There is some evidence to support this interpretation. For example, significant tropospheric warming occurs also in another recent reanalysis of the MSU channel 2 data (Vinnikov and Grody 2003), which uses a different strategy from that in Mears et al. (2003) and Christy et al. (2000) to account for drift in sampling the diurnal cycle. Also, a recent study by Fu et al. (2004) applied a regression-based technique to adjust MSU channel 2 data for the contribution it receives from the cooling stratosphere. This statistical adjustment reveals pronounced tropospheric warming, even in the Christy et al. (2000) MSU channel 2 product. Further study of the robustness of this result is underway. A recent study suggests that radiosonde products (whose trends are quite similar to UAH trends) may also have underestimated the tropospheric warming since 1979 (Lanzante et al. 2003). Additionally, synthetic MSU temperatures computed from the recently-completed ERA-40 reanalysis project also show tropospheric warming (Santer et al., 2004), and are in close agreement with the Mears et al. (2003) version of MSU channel 2. There are, however, valid scientific concerns regarding some of this supporting evidence, and further research is needed to validate the vertical coherence in model simulations.

Alternatively, if the Christy et al. (2000) analysis is closer to the “true” tropospheric temperature change over the satellite era, then we do not understand the factors that influence observed lapse rate variability on multi-decadal timescales, and climate models cannot reproduce the “observed” differential warming. This highlights the importance of reducing uncertainties in satellite- and radiosonde-based estimates of recent tropospheric temperature changes. If we cannot achieve this, we admit the possibility of very different outcomes in comparisons between modeled and observed atmospheric temperature trends, ranging from “close correspondence” to “fundamental inconsistency”.

b. Tropopause Height

The tropopause marks the transition between the turbulently-mixed troposphere and the more stably-stratified stratosphere. Measurements from radiosondes indicate that the height of the tropopause has increased in recent decades (Ramaswamy et al. 2001; Seidel et al. 2001). Reanalysis products show similar changes (Randel et al. 2000; Santer et al. 2003b). Superimposed on these decadal-scale increases in tropopause height are short term decreases associated with major volcanic eruptions. Long term change in the location of the tropopause may be a useful fingerprint of human effects on climate. The height of the tropopause reflects an integrated response to temperature changes in both the troposphere and stratosphere, leading to signal-to-noise (S/N) characteristics than are rather different from those of atmospheric temperatures in discrete layers (Santer et al. 2003c).

Model results suggest that the observed decadal-scale increase in tropopause height is largely driven by the GHG-induced warming of the troposphere and the ozone-induced cooling of the stratosphere, and that natural variability alone (internal, solar, and volcanic) cannot explain this increase. Model fingerprints of anthropogenically-forced tropopause height change are readily identifiable in reanalysis data (Santer et al. 2003c). Positive detection results are robust to a number of different processing options, such as the vertical resolution of the temperature data used to define the lapse-rate tropopause, inclusion or removal of global means, choice of reanalysis dataset, etc (Santer et al., 2004). Significant uncertainties remain, however, in the relative contributions of tropospheric warming and stratospheric cooling to recent tropopause height change.

c. Circulation

Anomalous external forcing on the climate system may produce dynamical (e.g., Palmer 1999a) as well as thermal responses in the atmosphere. One such dynamical change which has received considerable attention is the recent trend toward the positive phase of the Arctic Oscillation (AO; Thompson and Wallace 1998) and its Atlantic sector version, the North Atlantic Oscillation (NAO), to linger in its positive state (see Gillett et al. 2003a). This has been attributed by some to anthropogenic greenhouse gas increases and/or ozone forcing (e.g., Shindell et al. 2001). Although all anthropogenically forced models consistently produce some increase in a pattern-based NAO index (varying from very small for some models to similar to observed for others; Osborne 2002), regional details are not consistent between models (Gillett et al. 2003a and Osborn 2002). Natural variability of the NAO/AO cannot be fully eliminated as longer episodes showing similar changes have occurred in the early 20th century and may have occurred in earlier centuries (Luterbacher et al. 2002). The change in the pattern based AO/NAO index (obtained using the first empirical orthogonal function of surface pressure) is more unusual with respect to the observed earlier variability than that in the NAO index based on longer records of the difference between surface pressure readings at the Azores (or Gibraltar) and Iceland. Also, the change in the hemispheric scale AO (also referred to as the Northern Annular Mode) is more unusual relative to the earlier record than in the Atlantic sector based NAO (Ostermeier and Wallace 2003). Note that recent winters have seen a decline in the NAO back to 1951-1980 levels.

While attribution of the observed AO/NAO trend to external influences is uncertain (Osborn

2002; Gillett et al. 2003a), evidence of a human contribution to changes in the global surface pressure distribution observed during 1948-1998 has been obtained using signals from a number of climate models (Gillett et al. 2003b). While the signals are detected in several independently constructed surface pressure datasets (HadSLP – Basnett and Parker 1997; NCEP/NCAR reanalysis – Kalnay et al. 1996; Trenberth and Paolino 1980), the strength of the response to anthropogenic forcing that is simulated by the models is a factor of 3-5 weaker than the apparent response in the observations. This suggests that the effects of future circulation changes may be underestimated by climate models and points to a need for research to reconcile the discrepancies between models and observations.

Circulation changes that are largely associated with a trend towards the positive phase of the Southern Annular Mode with anomalously low pressure over Antarctica have also been noted in the Southern Hemisphere (IPCC 2001; Thompson et al. 2000; Thompson and Solomon 2002). Modeling evidence suggests that these changes are largely consistent with forcing due to ozone depletion (Sexton, 2001; Gillett and Thompson 2003), though greenhouse gas increases may have also played a role (Fyfe et al. 1999, Kushner et al. 2001).

Recently, attention has also been given to Southern Hemisphere (SH) synoptic variability (Simmonds and Keay 2000; Fyfe 2003) as represented in the NCEP/NCAR reanalysis. These studies have noted a shift in the distribution of cyclones over the 1960-1999 period, with fewer cyclones in the sub-Antarctic Southern Ocean, and an increase over the Antarctic Ocean. While it is not clear whether this apparent change is the result of increases in SH data availability, Simmonds and Keay (2000) use circumstantial evidence and studies conducted with independent data to argue that the decline equator ward of 60^{o}S is probably real. Fyfe (2003) shows that these changes are associated with a shift toward the positive phase of the Southern Annular mode and that similar changes occur in an ensemble of anthropogenically forced CGCM simulations beginning early in the 20th century. The simulated change emerges beyond the range of the model's natural variability by mid century, and simulated changes between 1960 and 1999 are similar to those seen in the NCEP/NCAR reanalysis.

6. Oceans

The oceans, because of their large heat storage capacity, are the thermal flywheel of the climate system. They store most of the heat contained in the system and, along with the atmosphere, redistribute it to maintain the climate system as we know it today. As global warming increases, it is critical to know how oceanic heat storage and sea level will be affected.

a. Heat Content

Recent work by Levitus et al. (2001) has made it possible to document changes in oceanic heat content from 1955 onwards. These changes have been evaluated on an ocean-by-ocean basis by Barnett et al. (2001), comparing model results with observations only where the latter existed, thereby avoiding many sampling problems (cf. Gregory et al, 2004). They found that the heat content in the upper 3000m of each of the world's oceans has increased steadily since 1955 (and apparently had been increasing well before that – Section 3b; Figure 3). Anthropogenically forced simulations obtained from a CGCM showed changes in heat content similar to those calculated from observations. A detection and attribution analysis, using methods similar to those described in Section 2, demonstrates consistency between the observed and simulated changes in ocean heat content and indicates that the observed change is probably not the result of internal climate system variability alone. Similar results were reported by Reichert et al. (2002) using ocean heat content signals from a different CGCM. It should be noted that the forcings from these two simulations were somewhat different from each other.

Other potential sources of the observed ocean warming have since been investigated. Changes in solar forcing can potentially explain only about 2% of the observed increase in ocean heat content (Crowley et al. 2004). Geothermal heat escaping to the oceans from the great rifts may explain perhaps 15% of the observed change (W. Munk and J. Orcutt 2003, pers. com.) and thus sea floor heating is probably not a major factor. In contrast, estimates of changes in ocean heat content caused by anthropogenic warming provide a much closer fit to the observations, and thus provide a more plausible explanation of the underlying cause.

The ocean heat content analyses cited above are based on basin integrated values for the different ocean basins, in part because observational coverage, particularly at lower levels in the ocean, remains thin. Future efforts must seek to evaluate the heat input into the oceans, and their response, on a full three dimensional basis, else they risk ignoring what should be a strong, spatially dependent signal in the warming (e.g. Gregory 2000; Gregory et al 2004; Sun and Hansen 2003; and Sokolov et al 2003). In one such a detailed study of ocean climate change, Banks and Bindoff (2003) find that observed watermass changes in the Indo-Pacific appear to be consistent with those simulated by a coupled climate model (HadCM3), and that the correlation between simulation and observations is significant relative to the variability of the control simulation. But the comparison is still difficult due to the inhomogeneous space-time coverage of the observations and uncertainties introduced through the methods of filling in the observational dataset outside the well observed part of the ocean (Gregory et al 2004) and the fact that many ocean models underestimate the levels of natural variability of phenomena such as the Pacific Decadal Oscillation and ENSO. In the end, better regional ocean information will be necessary for regional response and detection studies and biological impacts analyses.

b. Sea Level and Sea Ice

Sea level has been measured at a large number of locations around the world since before 1900. Several studies conducted over the last 20 years have estimated a linear rise in sea level of approximately 10-20 cm over the last century (IPCC 2001) after taking factors such as continental rebound into account. This value has a substantial uncertainty attached to it. A continued rise associated with anthropogenically-induced ocean warming (steric effects) and melting of land ice could have potentially large impacts on low lying coastal regions, including the Gulf Coast of the United States, and small islands such as the Maldives. But how much of the observed rise to date can be attributed to anthropogenic causes?

It has recently become possible to attribute portions of the observed increase in sea level to various physical mechanisms (c.f. Munk 2002). The Levitus et al. (2000, 2001) global ocean temperature data implies a sea level rise of 3 cm during the latter half of the 20th century. Thermal expansion may have caused an additional 3 cm rise in the first half of the century (Crowley et al. 2004). The IPCC TAR estimates total sea level rise between 1910 and 1990 from eustatic processes, including the effects of thermal expansion and melting land ice, of 7 cm, but with a very wide uncertainty range. This raises a serious problem as noted by Munk (2002), because about 50% of sea level rise, again with very large uncertainty, remains unexplained: the observed sea level rise starts too early, is too linear and is too large to be due to anthropogenic effects alone. Munk notes that the differential might be explained by melting of the polar caps and corresponding changes in the earth's moment of inertia, but observed changes in the Earth's rotation characteristics only partially support such conjecture.Other possible explanations include uncertainties in current estimates of sea level rise and ocean heat storage, especially in the deep ocean where there are few observations (Cabanes et al. 2001; Munk 2002). This discrepancy emphasizes that more research is needed to understand previous as well as future predicted sea level rise.

Some attention has also recently been paid to observed changes in Arctic sea-ice extent. A recent detection exercise showed that these changes are outside the range of natural internal variability and are consistent with changes in model forced with both anthropogenic and natural external forcing agents. At the same time, the observed changes are not consistent with model results when only changes in natural external forcing are prescribed (Gregory et al. 2002b). This suggests that Arctic sea ice extent is also beginning to show the signature of anthropogenic climate change.

7. Rainfall and climate extremes

It is recognized that the impact of climate change will probably be felt most strongly through changes in precipitation and short-term climate extremes such as heavy rainfall and/or flooding, extreme temperature and heat waves. It is therefore important to determine whether these aspects of climate are changing and whether the changes can be attributed to human activity. We therefore very briefly review a few of the developments related to these questions in this section.

a. Observed and simulated changes

Evidence for observed changes in short duration extremes generally depends on the region considered and the analysis method (IPCC 2001). So far, only a few global analyses have been performed, mainly due to the poor availability of quality controlled and homogenized daily station data. However, as noted in Section 2, the availability of historical daily data is improving. Also, indices for temperature and precipitation extremes that are calculated from station data are becoming available, including some indices from regions where daily station data are not released (Frich et al. 2002; Klein-Tank and Können, 2003).

A first analysis of a subset of indices suggests detectable signals in indices of temperature extremes, but limited agreement between modeled and observed changes in rainfall and drought extremes (Kiktev et al. 2003). However, significant increases in observed extreme precipitation have been reported over some parts of the world, for example over the United States (Karl and Knight 1998), where the increase is similar to changes expected under greenhouse warming (for example, Semenov and Bengtsson 2002; Groisman et al. 2004). However, a quantitative comparison between area-based extreme events simulated in models and point-observations by weather stations, as required in detection efforts, remains difficult because of the different scales involved (Osborn and Hulme 1997). In the absence of detection results for extreme events, particularly for non-temperature related events, studies based on model data alone can be used to develop suitable approaches for early detection.

Simulated changes in globally averaged annual mean and extreme precipitation appear to be quite consistent between models and follow physical principles (Allen and Ingram 2002). However, the spatial pattern of projected precipitation change is very different between models. The latter makes changes in annual mean precipitation difficult to detect in output from a given model when using a precipitation change signal from another model (e.g., Hegerl et al. 2004). This, coupled with the currently low signal-to-noise ratio for anthropogenically forced change in model output, suggest that detection in observations will remain a challenge. Interestingly, Allen and Ingram (2002) demonstrate that global mean precipitation variations simulated by a model that includes both natural and anthropogenic external forcing tend to follow observed variations in global mean precipitation, although Lambert et al (2004) indicate that most of this agreement is probably due to the precipitation response to natural forcing. Gillett et al (2004) confirm this: they detect volcanic influence in global precipitation, but no anthropogenic influence, using integrations of the Parallel Climate Model. Robock and Liu (1994) previously showed a significant reduction in global mean precipitation in a model in response to a volcanic eruption. This result has a clear physical explanation: a short-wave forcing anomaly, due for example to solar or volcanic activity, that gives the same temperature response as a longwave forcing anomaly due to a change in CO_2 concentrations has a much greater impact on precipitation. The reason is that the direct impact of a CO_2 increase on the tropospheric energy budget is to reduce precipitation (e.g., Yang et al 2003). Subsequent surface and tropospheric warming more than compensates for this reduction in almost all models, but still leaves a net precipitation change induced by CO_2 smaller than the change induced by a similar-magnitude (in terms of temperature impact) short-wave forcing. This has important implications for detection and for placing constraints on projections of future precipitation amounts (Section 4d), since different forcings affect precipitation in different ways.

Model results suggest that future changes in precipitation extremes will probably be greater than changes in mean precipitation (see, for example, Meehl et al. 2000; Kharin and Zwiers 2000, 2004; Semenov and Bengtsson 2002). They also indicate that simulated changes in temperature extremes cannot be explained just by a shift in the temperature distribution (Hegerl et al. 2004). Therefore, changes in extremes can not be inferred from changes that are detected in the more widely available monthly mean data. A model-model detection study, where fingerprints from one model were used to detect precipitation change in simulations from another model, suggests that changes in moderately extreme precipitation (i.e., the magnitude of events that occur a few times per year) may be more robustly detectable using signals from different models than changes in annual total rainfall (Figure 10; Hegerl et al. 2004). This is mainly because precipitation extremes increase over a large fraction of the globe, making detection results less sensitive to the spatial pattern of change.

b. Attribution of changes in climate-related risk

An emerging challenge in detection and attribution research is that many important impacts of climate change will probably manifest themselves through a change in the frequency or likelihood of occurrence of events that, taken individually, could be explained as natural (e.g. Palmer 1999b). Whenever an extreme climate event occurs, such as the floods in the central Europe in the Autumn of 2002 or the Mississippi-floods of 1993 in the USA, the question arises as to whether this event has been "caused" by climate change. If the event in question might have occurred naturally, the only answer that can be given to this question, as stated, is "no", although it might be possible to add that this kind of event is expected to become more likely in the future as a result of climate change. There is, however, a clear demand for a more quantitative assessment, not only for assessing how such risks may change in the future (Palmer and Raissanen 2002) but also to provide a more complete understanding of the reasons they are occurring now.

The concept of 'attributable risk' is well established in the epidemiological literature. If P_1 is the probability of an event (such as a flood) occurring now, and P_0 is the probability of it occurring, all other things being equal, if greenhouse gas concentrations had not increased over the past century, then the fraction of the current risk that is attributable to past greenhouse gas emissions is simply $1 - P_0 / P_1$ (Allen 2003). This concept applies straightforwardly to events that occur frequently both with and without the risk factor in question (i.e. in both present-day climate, under the influence of past greenhouse gas emissions, and the climate that would have been obtained at present, all other things being equal, in the absence of those emissions), since both P_1 and P_0 can be interpreted simply in terms of frequency of occurrence. Analyses of attributable risk provide the basis for such statements as “half the deaths due to X are attributable to environmental risk factor Y” and are subject to well-documented hazards of interpretation (Greenland & Robins 1988) which need to be borne in mind as they are extended to the climate problem. The “frequency of occurrence” interpretation becomes more problematic when we are dealing with the most extreme events that, by definition, occur very infrequently in both present-day and pre-industrial climate where the probability of events is much more difficult to assess. Changes in the probability and recurrence time of extreme rainfall, temperature and storminess events are expected under climate change conditions (e.g., Kharin and Zwiers 2000, 2004), suggesting that this kind of quantitative risk analysis will become more important in the future.

8. Bayesian studies

As indicated in Section 2b, there is increasing interest in using Bayesian methods in detection and attribution work, and several studies using these techniques are now available. This subsection reviews three of these studies. Two studies (Schnur and Hasselmann 2004; Lee et al. 2004) used geographically distributed surface temperature data and geographical patterns of anthropogenically forced climate change simulated by CGCMs. A third study (Smith et al. 2003) used time series of Northern and Southern hemisphere mean surface temperature and hemispheric mean anthropogenic signals from an EBM.

Schnur and Hasselmann (2004) report on a study that uses an optimal Bayesian filtering technique. They applied the technique to recent 31-year trends of near-surface temperature, precipitation, and summer and winter diurnal temperature range, and considered three competing hypotheses: namely that the climate changes observed late in the 20^{th} century can be explained by natural internal variability alone, by natural internal variability and greenhouse gas forcing

(G), or by natural internal variability and the combined effect of greenhouse gas and sulfate aerosol forcing (GS). These three possibilities were assumed to be equally likely apriori. Schnur and Hasselmann constructed a filter that maximizes the impact of the observations on the prior likelihood of detection. As with the standard optimal fingerprinting approach, a filter is required because there is insufficient data to perform the analysis without reducing the dimensionality of the problem. However, because a different optimization criterion is used, the Bayesian filter does not require as large a dimension reduction as the standard approach (Figure 11a). Consequently, the the G and GS signals are more sharply defined.

Figure 11b shows the posterior probabilities that were obtained in the reduced (optimal) detection space for each variable and hypothesis. Also shown is the net posterior that is obtained when all four variables are used simultaneously. For temperature, the odds of hypotheses G and GS against the natural-variability hypothesis are seen to be very large, i.e. the G and GS signals are clearly detected in the observations. However, GS is found to be only twice as likely as G, which does not represent decisive evidence for one hypothesis over the other (Kass and Raftery 1995). For diurnal temperature range (DTR) and precipitation, detection is not achieved. In the combined analysis, G and GS are still both detected, with the odds of GS over G being only slightly larger than one. Compared to conventional analyses, the inclusion of the model error structure in the Bayesian analysis leads to a downgrading of the information on the impact of sulfate forcing. This occurs because the model uncertainty in the response to aerosols is much larger than that for greenhouse gases.

Lee et al. (2004) report on a Bayesian study using a version of the technique described by Berliner et al. (2000). Their approach is to perform a conventional optimal fingerprinting analysis using multiple regression as discussed in Section 2a, and then to use Bayesian techniques to evaluate detection and attribution hypotheses about the scaling factors $\mathbf{a}$. This study considers only the GS signal, but evaluates the evidence for its presence in observations for several 5-decade windows, beginning with 1900-1949, and ending with 1950-1999. The GS signal and the necessary estimates of natural internal variability were obtained from two versions of the CCCma CGCM. Dimension reduction in this study was performed in the conventional manner.

The resulting posterior distributions were analysed to assess whether there was evidence to support detection and attribution. Evidence supporting detection was assessed by comparing prior and posterior probabilities that $\mathbf{a} \in D$, where $D = [0.1, \infty)$. This criterion, in effect, requires high posterior likelihood that the observed response to GS forcing is positive, but requires that the model response is not larger than 10-times the apparent observed response. Evidence supporting attribution was similarly assessed by comparing prior and posterior probabilities that $\mathbf{a} \in A$, where $A = (0.8, 1.2)$. This is a stringent requirement that, in effect, requires a high posterior likelihood that the amplitude of the model response to historical GS forcing is within 20% of the apparent observed response. Similar to Schnur and Hasselmann (2004), the Lee et al. evaluation of the evidence was made by means of Bayes factors (Kass and Raftery, 1995), which are ratios that compare the posterior odds of detection (or attribution) to the corresponding prior odds.

The results of this analysis provide "Very Strong" evidence (defined as a Bayes factor greater than 150) in support of detection during the early and latter halves of the 20th century regardless of the choice of prior distribution. On the other hand, evidence for attribution, as stringently defined above, is weak. "Positive evidence" (Bayes factor in the range 3 to 20) for attribution was obtained when using noncommittal priors and a less stringent attribution criterion that requires the model response to the GS forcing to be within 50% of the apparent observed response. Lee et al. estimated that “Strong evidence” for attribution may emerge within the next two decades as the anthropogenic signal strengthens.

The third study using Bayesian methods cited above (Smith et al. 2003) considers the joint behaviour of the Northern and Southern hemispheric mean surface temperature time series during the 20th century. A Bayesian approach was used in this case to faciliate intercomparison between a number of competing time series models of the temporal covariability between the hemispheres. Some of these models included estimates of responses to historical variations in external forcing as covariates. The main finding was that a bivariate time series model that includes the responses to historical greenhouse gas, aerosol, and solar forcing factors is clearly better than similar time series model that exclude one or more of these forcing factors.

All three of these studies demonstrate the utility of the Bayesian approach, both as a tool for selecting between competing statistical models (Smith et al. 2003) or as a means for taking prior information and sources of uncertainty into account (Schnur and Hasselmann 2004; Lee et al. 2004). Bayesian techniques have often been criticized because they incorporate prior information that may be subjective in nature. However, many authors have shown that the prior plays a relatively minor role in determining the form of the posterior distribution. Bayesian approaches are better able to quantify the evidence for attribution and offer the flexibility to easily incorporate evidence from multiple lines of evidence together with model and observational uncertainty into a single comprehensive analysis. The interest in these methods can be expected to continue to increase.

9. Summary and Conclusions

Recent detection studies based on changes in surface temperature are consistent with the findings reported in the IPCC TAR. Additional uncertainties, such as the effect of differences in model fingerprints and simulations have meanwhile been addressed. It has been demonstrated that averaging fingerprints from multiple models increases our confidence in detecting anthropogenic climate change and that different implementations of the optimal detection method show consistent results. New detection studies show that anthropogenic climate change is detectable in the surface temperature records of individual continents and that it can be distinguished from climate change due to natural forcing.

Important progress has also been made in our understanding of Northern Hemisphere mean temperature over the last millennium. Paleo reconstructions of hemispheric mean temperature show that the 20th century warming is unique in the last millennium both for its size and rapidity. Recent evidence suggests that the borehole –record of climate change may be consistent with that of conventional proxy records. However, reconstruction techniques are still being evaluated. New detection studies based on proxy reconstructions yield generally consistent results with studies based on the instrumental period. The combined response to volcanism and greenhouse gases can be detected in most reconstructions and the individual responses to greenhouse gas, volcanic and solar forcing can be distinguished from each other with only small differences in results for different records. Simulations of the last millennium with coupled climate models are becoming available, which should help in further improving our understanding of the climate of the last millennium.

Progress has also been made towards understanding the temperature evolution in the free atmosphere. Recently, an independently processed version of the Microwave Sounding Unit

(MSU), channel 2 recorded temperature of the mid-troposphere has become available. This new product shows global warming at a rate that is consistent with model-simulations of the warming in the mid-troposphere and the differential warming between surface and mid-troposphere. While it is still unclear which of the MSU2 temperature datasets is more reliable, this work suggests that the differential warming between surface and lower troposphere is within present observational uncertainty. Further work is needed to assess and reduce uncertainties in estimates of observed tropospheric warming. Furthermore, observed variations in tropopause height have been shown to be consistent with model simulated changes in tropopause height resulting from anthropogenic and natural forcing.

The anthropogenic climate change signal has also been detected in ocean heat content, which has increased in all ocean basins. The model simulated trend in ocean heat content during the latter half of the 20th century is similar to that estimated from observations and does not seem explainable via natural forcings. The simultaneous warming in all basins could not be explained by solar variability or geothermal forcing. However, questions remain about the spatial similarity of the observed and modeled signals within each ocean basin, as well as sampling problems associated with the distribution of the observed ocean temperature data. An anthropogenic signal has also been detected in global sea level pressure data, although in this case, the model simulated response to anthropogenic forcing is significantly weaker than the changes that have been observed.

We conclude that there is now further evidence supporting the findings of the IPCC TAR. Our understanding of the importance of external forcing in determining Earth's climate and its variability have been increased through the continued development and analysis of millenial scale proxy climate records. The detection of climate change from anthropogenic forcing has been extended to continental scales and to other variables including sea level pressure and tropopause height. Several methodological advances have been made, including the use of multi-model methods and the increasing use of Bayesian methods to more comprehensively assess the available detection and attribution evidence. Important steps have been made towards a better understanding of the temperature evolution at the earth's surface compared to the free atmosphere in the satellite era. Progress is being made towards understanding ocean climate change and changes in climatic extremes, although we are still far from having a full understanding of both of these important issues.

Many open questions remain, for example, the role of forcings not yet fully included in CGCM simulations, such as land use change or forcing by black carbon and non-sulfate aerosols. Also, due to poor signal-to-noise ratios and model uncertainty, anthropogenic rainfall changes can not presently be detected even on a global scale, although a volcanic signal is detectable in global mean land rainfall. It will remain to be investigated whether projections of future changes in the hydrological cycle can, in part, be constrained by the apparent rainfall response to natural forcing, and the possibility that change in intense precipitation may be more detectable.

Acknowledgements

The International ad hoc Detection and Attribution Group (IDAG) is a self-organized group of climate experts interested in the problem of the detection and attribution of externally forcing climate signals in the observations. This report, like our others, expresses our collective views and not those of any international or national agency. While most of the Group obtains funding from their normal sources, we wish to acknowledge the Climate Change Detection and Attribution Project, a jointly funded effort by NOAA's Office of Global Programs and the Department of Energy's Office of Science, for providing support for meeting and specialized studies of the Group members. PAS was funded by the UK department for Environment, Food and Rural Affairs under contract PECD 7/12/37. GH was also funded by NSF grant ATM-0296007. We also thank Karen King of DOE's Pacific Northwest National Laboratory for handing the Group's administrative matters and Jesse Kenyon for help with the manuscript.

This work in part was performed under the auspices of the U.S. Department of Energy by University of California, Lawrence Livermore National Laboratory under contract No. W-7405-Eng-48.

References

Allen, M. R., 2003: Liability for climate change. *Nature*, **421**, 891-892.

Allen, M.R. and S.F.B. Tett, 1999: Checking for model inconsistency in optimal fingerprinting. *Climate Dyn.*, **15**, 419-434.

Allen, M. R. and W. J. Ingram, 2002: Constraints on future changes in climate and the hydrologic cycle. *Nature*, **419**, 224-232.

Allen, M. R. and D. A. Stainforth, 2002: Towards objective probabilistic climate forecasting, *Nature,* **419**, 228.

Allen, M. R. and P.A. Stott, 2003: Estimating signal amplitudes in optimal fingerprinting, Part I: Theory. *Climate Dyn.*, **21**, 477-491.

Allen, M. R., P. A. Stott, J. F. B. Mitchell, R. Schnur and T. Delworth, 2000: Quantifying the uncertainty in forecasts of anthropogenic climate change. *Nature*, **407**, 617-620.

Allen, M.R. and Coauthors, 2004: Quantifying anthropogenic influence on recent near-surface temperature. *Surveys in Geophysics*, in press.

Anderson, T.L., R.J. Charlson, S.E. Schwartz, R. Knutti, O. Boucher, H. Rodhe and J. Heintzenberg, 2003: Climate forcing by aerosols - a hazy picture. *Science*, **16**, 1103-1104.

Andronova, N. G. and M. E. Schlesinger, 2000: Causes of global temperature changes during the 19^{th} and 20^{th} centuries. *Geophys. Res. Lett.*, **27 (14)**, 2137-2140.

Angell, J. K., 2000: Difference in radiosonde temperature trend for the period 1979-1988 of

MSU data and the period 1959-1998 twice as long. *Geophys. Res. Lett*., **27**, 2177-2180.

Banks, H. T. and N. L. Bindoff, 2003: Comparison of observed temperature and salinity changes in the Indo-Pacific with results from the coupled climate model HadCM: Processes and mechanisms. *J. Climate*, **16**, 156-166.

Barnett, T. P., and Coauthors, 1999: Detection and attribution of recent climate change: A status report. *Bull. Amer. Meteor. Soc*., **80**, 2631-2660.

Barnett, T.P., D.W. Pierce and R. Schnur, 2001: Detection of anthropogenic climate change in the world's oceans. *Science*, **292**, 270-274.

Basnett, T.A. and D.E. Parker, 1997: Development of the global mean sea-level pressure data set GMSLP2. Clim. Res. Tech. Note 79, Hadley Centre, Met Office, Bracknell, UK.

Bauer E., M. Claussen and V. Brovkin, 2003: Assessing climate forcings of the Earth system for the past millennium. *Geophys. Res. Let.*, **30,** 1276, doi: 10.1029/GL016639.

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

Berliner, L.M, R.A. Levine, and D.J. Shea, 2000: Bayesian climate change assessment. *J. Climate*, **13,** 3805-3820.

Bertrand, C., M. F. Loutre, M. Crucifix and A. Berger, 2002: Climate of the last millennium: A sensitivity study. *Tellus*, **54A**, 221-244.

Braganza, K., D. J. Karoly, A. C. Hirst, P. Stott, R. J. Stouffer, and S. F. B. Tett, 2004: Simple indices of global climate variability and change: Part II – Attribution of climate change during the 20th century. *Climate Dyn*., **22**, 823-838, doi:10.1007/S00382-004-0413.

Briffa, K.R. and T.J Osborn, 2002: Blowing hot and cold. *Science,* **295**, 2227-2228.

Briffa, K.R., T.J. Osborn, F.H. Schweingruber, I.C. Harris, P.D. Jones, S.G. Shiyatov, and E. A. Vaganov, 2001: Low-frequency temperature variations from a northern tree-ring density network, *J. Geophys. Res*., **106**, 2929-2941.

Cabanes, C., A. Cazenave, and C. Le Provost, 2001: Sea level rise during past 40 years determined from satellite and in situ observations. *Science,* **294**, 840-842.

Christy, J.R., R.W. Spencer and E. Lobl, 1998: Analysis of the merging procedure for the MSU daily temperature time series. *J. Climate*, **5**, 2016-2041.

Christy, J.R., R.W. Spencer, and W.D. Braswell, 2000: MSU tropospheric temperatures: Dataset construction and radiosonde comparison. *J. Atmos. Oceanic Technol.*, **17**, 1153-1170.

Christy, J.R., D.E. Parker, S.J. Brown, I. Macadam, M. Stendel and W.B. Norris, 2001: Differential trends in tropical sea surface and atmospheric temperatures. *Geophys. Res. Lett*., **28**, 183-186.

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

Crowley, T.J., 2004: A 2000 year composite estimate of global volcanism from ice cores. *J. Geophys. Res*., in review.

Crowley and Lowery, 2000: How warm was the medieval warm period? A comment on “Man-made versus natural climate change.” *Ambio,* **29**, 51-54.

Crowley, T. J., S. K. Baum, K.-Y. Kim, G. C. Hegerl and W. T. Hyde, 2004: Modeling ocean heat content changes during the last millennium. *Geophys. Res. Lett*., in press.

Delworth, T.L. and T.R. Knutson, 2000: Simulation of early 20th century global warming. *Science*, **287**, 2246-2250.

Diaz, H.F., C.K. Folland, T. Manabe, D.E. Parker, R.W. Reynolds and S.D. Woodruff, 2002: Workshop on advances in the use of historical marine climate data . *WMO Bulletin*, **51**, 377-380.

Esper, J., E. Cook, and F.H. Schweingruber, 2002: Low-frequency signals in long tree-ring chronologies for reconstructing past temperature variability**.** *Science*, **295,** 2250-2253.

Folland, C.K., and Coauthors, 2001a: Global temperature change and its uncertainties since 1861. *Geophys. Res. Lett.*, **28**, 2621-2624.

Folland, C.K., and Coauthors, 2001b: Observed climate variability and change. *Climate Change 2001. The Scientific Basis. The Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change.* J. T. Houghton et al. Eds., Intergovernmental Panel on Climate Change (IPCC), 880pp.

Forest, C.E., M.R. Allen, P.H. Stone, and A.P. Sokolov, 2000: Constraining uncertainties in climate models using climate change detection techniques. *Geophys. Res. Lett*., **27**, 569-572.

Forest, C.E., P.H. Stone, A.P. Sokolov, M.R. Allen, and M. D. Webster, 2002: Quantifying uncertainties in climate system properties with the use of recent observations. *Science*, **295**, 113-117.

Free, M. and J.K. Angell, 2002: Effect of volcanoes on the vertical temperature profile in radiosonde data. *J. Geophys. Res*., doi:10.1029/2001JD001128

Frich, P., L.V. Alexander, P. Della-Marta, B. Gleason, M. Haylock, A.M.G. Klein-Tank and T. Peterson, 2002: Observed coherent changes in climatic extremes during the second half of the twentieth century. *Climate Research*, **19**, 193-212.

Fu, Q., C.M. Johanson, S.G. Warren, and D.J. Seidel, 2004: Contribution of stratospheric cooling to satellite-inferred tropospheric temperature trends. *Nature*, **429**, 55-58.

Fyfe, J.C., 2003: High latitude Southern Hemisphere cyclones: Harbingers of climate change? *J. Climate*, **16**, 2802-2805.

Fyfe, J.C., G.J. Boer, and G.M. Flato, 1999: The Arctic and Antarctic Oscillations and their projected changes under global warming. *Geophys. Res. Lett*., **26**, 1601-1604.

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

Gillett, N.P. and D. W. J. Thompson, 2003: Simulation of recent Southern Hemisphere climate change. *Science*, **302**, 273-275.

Gillett, N. P., G.C. Hegerl, M.R. Allen, and P.A. Stott, 2000: Implications of changes in the Northern Hemispheric circulation for the detection of anthropogenic climate change. *Geophys. Res. Lett.*, **27**, 993-996.

Gillett, N.P., G.C. Hegerl, M.R. Allen, P.A. Stott, and R. Schnur, 2002a: Reconciling two approaches to the detection of anthropogenic influence on climate. *J. Climate*, **15**, 326-329.

Gillett, N.P., F.W. Zwiers, A.J. Weaver, G.C. Hegerl, M.R. Allen, and P.A. Stott, 2002b: Detecting anthropogenic influence with a multi-model ensemble. *Geophys. Res. Lett.,* **29**, 1970, doi:10.1029/2002GL015836.

Gillett, N.P., H. Graf, and T. Osborn. 2003a: Climate Change and the NAO, *The North Atlantic Oscillation*. Hurrell, Kushnir, Visbeck and Ottersen, Eds., AGU, Washington, D.C. 193-210.

Gillett, N.P., F.W. Zwiers, A.J. Weaver, and P.A. Stott, 2003b: Detection of human influence on sea level pressure. *Nature*, **422**, 292-294.

Gillett, N.P., A.J. Weaver, F.W. Zwiers, and M.F. Wehner, 2004: Detection of volcanic influence on global precipitation. *Geophys. Res. Lett.*, **31**, L12217, doi:10.1029/2004GL020044.

Greenland, S. and J. M. Robins, 1988: Conceptual problems in the definition and interpretation of attributable fractions. *Amer. J. Epidemiol*., **128(6)**, 1185-1197.

Gregory, J.M., 2000: Vertical heat transports in the ocean and their effect on time-dependent climate change. *Climate Dyn.*, **16**, 501-515.

Gregory, J.M., R. J. Stouffer, S. C. B. Raper, P. A. Stott, and N. A. Rayner, 2002a: An observationally based estimate of the climate sensitivity. *J. Climate*, **15**, 3117-3121.

Gregory, J.M., P. A. Stott, D. J. Cresswell, N. Rayner, C. Gordon, and D. M. H. Sexton, 2002b: Recent and future changes in Arctic sea ice simulated by the HadCM3 AOGCM. *Geophys. Res. Lett*., **29**, doi:10.1029/2001GL014575.

Gregory, J.M., H.T. Banks, P.A. Stott, J.A. Lowe, and M.D. Palmer, 2004: Simulated and observed decadal variability in ocean heat content, *Geophys. Res. Lett.*, **31**, L15312, doi:10.1029/2004GL020258.

Groisman, P., R. Knight, D. R. Easterling, T. R. Karl and G. Hegerl, 2004: Trends in precipitation intensity in the climate record. *J. Climate*, accepted.

Hansen, J., A. Lacis, D. Rind, G. Russell, P. Stone, I. Fung, M. Ruedy, and J. Lerner, 1984: Climate sensitivity: analysis of feedback mechanisms. *Climate Processes and Climate Sensitivity*, Geophysical Monograph 29, Amer. Geophys. Union, Washington, D.C., J.E. Hansen and T. Takahashi, Eds., 130-163.

Hasselmann, K., 1979: On the signal-to-noise problem in atmospheric response studies. *Meteorology of Tropical Oceans*, D.B. Shaw, Ed., Royal Meteorological Society, 251-259.

Hasselmann, K., 1997: Multi-pattern fingerprint method for detection and attribution of climate change. *Climate Dyn*., **13**, 601-612.

Hasselmann, K., 1998: Conventional and Bayesian approach to climate-change detection and attribution. *Quart. J. Roy. Meteor. Soc*., **124**, 2541-2565.

Hegerl, G.C. and M.R. Allen, 2002: Origins of model-data discrepancies in optimal fingerprinting. *J. Climate*, **15**, 1348-1356.

Hegerl, G.C. and J.M. Wallace, 2002: Influence of patterns of climate variability on the difference between satellite and surface temperature trends. *J. Climate*, **15**, 2412-2428.

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

Hegerl, G.C., T. J. Crowley, S. K. Baum, K-Y. Kim, and W. T. Hyde, 2003: Detection of volcanic, solar and greenhouse gas signals in paleo-reconstructions of Northern Hemispheric temperature. *Geophys. Res. Lett.*, **30(5)**, 1242, doi:10.1029/2002GL016635.

Hegerl, G. C., F.W. Zwiers, V.V. Kharin, and P.A. Stott, 2004: Detectability of anthropogenic changes in temperature and precipitation extremes. *J. Climate,* in press.

Huang, S., 2004: Merging information from different resources for new insights into climate chane in the past and future. *Geophys. Res. Lett.*, **31**, L13205, doi:10.1029/2004GL019781.

Huang, S., H.N. Pollack, and P.-Y. Shen, 2000: Temperature trends over the past five centuries reconstructed from borehole temperatures. *Nature*, **403**, 756-758.

IPCC, 1990: *Climate Change: The IPCC Scientific Assessment.* J.T. Houghton, G.J. Jenkins and J.J. Ephraums, Eds., Cambridge University Press, 365pp.

IPCC, 1996: *Climate Change 1995: The Science of Climate Change.* J.T. Houghton, L.G. Meira Filho, B.A. Callander, N. Harris, A. Kattenberg and K. Maksell, Eds., Cambridge University Press, 572pp.

IPCC, 2001: *Climate Change 2001: The Scientific Basis, Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change.* J.T. Houghton, Y. Ding, D.J. Griggs, M. Noguer, P.J. van der Linden, X. Dai, K. Maskell and C.A. Johnson, Eds., Cambridge University Press, 881pp.

Jones, P.D. and A. Moberg, 2003: Hemispheric and large-scale surface air temperature variations: An extensive review and an update to 2001. *J. Climate,* **16**, 206-213.

Jones, P.D. and M.E. Mann, 2004: Climate over past millennia. *Rev. Geophys.*, **42**, doi:10.1029/2003RG000143.

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

Jones, P.D., T.J. Osborn, and K.R. Briffa, 2001: The evolution of climate over the last millennium. *Science*, **292**, 662-667.

Joos, F., R. Meyer, M. Bruno, and M. Leuenberger, 1999: The variability in the carbon sinks as reconstructed for the last 1000 years. *Geophy. Res. Lett.*, **26**, 1437-1441.

Kalnay, E. and Coauthors, 1996: The NCEP/NCAR 40-year Reanalysis Project. *Bull. Amer. Meteor. Soc*., **77**, 437-471.

Karl, T.R. and R.W. Knight, 1998: Secular trends of precipitation amount, frequency, and intensity in the USA. *Bull. Amer. Meteor. Soc*., **79**, 231-241.

Karoly D.J. and K. Braganza, 2001: Identifying global climate change using simple indices. *Geophys. Res. Lett.,* **28**, 2205-2208.

Karoly, D.J., K. Braganza, P.A. Stott, J.M. Arblaster, G.A. Meehl, A.J. Broccoli and D.W. Dixon, 2003: Detection of a human influence on North American climate. *Science*, **302**, 1200-1203.

Kass, R.E. and A.E. Raftery, 1995: Bayes Factors. *J. Amer. Statist. Assoc.,* **90**, 773-795.

Kharin, V.V. and F.W. Zwiers, 2000: Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere-ocean GCM. *J. Climate*, **13**, 3760-3788.

Kharin, V.V. and F.W. Zwiers, 2004: Estimating extremes in transient climate change simulations. *J. Climate*, in review.

Kiktev, D., D. Sexton, L. Alexander, and C. Folland , 2003: Comparison of modelled and observed trends in indices of daily climate extremes. *J. Climate*, **16**, 3560-3571.

Klein-Tank, A. M. G., and G. P. Können, 2003: Trends in indices of daily temperature and precipitation extremes in Europe, 1946–99. *J. Climate*, **16**, 3665–3680.

Knutti, R., T. F. Stocker, F. Joos, and G.-K. Plattner, 2003: Probabilistic climate change projections using neural networks. *Climate Dyn*., **21**, 257-272, doi:10.1007/s00382-003-0345-1.

Kushner, P. J., I. M. Held, and T. L. Delworth, 2001: Southern Hemisphere atmospheric circulation response to global warming. *J. Climate*, **14**, 2238-2249.

Lambert, F. H., P. A. Stott, M. R. Allen, and M. A. Palmer, 2004: Detection and attribution of changes in 20th century terrestrial precipitation. *Geophys. Res. Lett*., in review.

Lanzante, J.R., S.A. Klein, and D.J. Seidel, 2003: Temporal homogenization of monthly radiosonde temperature data. Part II: Trends, sensitivity, and MSU comparison. *J. Climate*, **16**, 241-262.

Lean, J.L.,Y-M Wang, and N.R. Sheeley, 2002: The effect of increasing solar activity on the sun's total and open magnetic flux during multiple cycles: Implications for solar forcing on climate. *Geophys. Res. Lett.*, **29**, 2224, doi:10.1029/2002GL015880.

Lee, T.C.K., F.W. Zwiers, X. Zhang, G.C. Hegerl, and M. Tsao, 2004: A Bayesian approach to climate change detection and attribution. *J. Climate*, in review.

Leroy, S.S., 1998: Detecting climate signals: Some Bayesian aspects. *J. Climate*, **11**, 640-651.

Levitus, S., J. Antonov, T.P. Boyer, and C. Stephens, 2000: Warming of the World Ocean. *Science*, **287**, 2225-2229.

Levitus, S., J. Antonov, J. Wang, T.L. Delworth, K.W. Dixon, and A.J. Broccoli, 2001: Anthropogenic warming of the Earth's climate system. *Science*, **292**, 267-270.

Luterbacher, J. and Coauthors, 2002: Extending the N. Atlantic Oscillation reconstructions back to 1500. *Atmos. Sci. Lett.,* **2**, 114-124, doi:10.1006/asle.2001.0044.

Mann, M.E., and G.A. Schmidt, 2003: Ground vs. surface air temperature trends: implications for borehole surface temperature reconstructions, *Geophysical Research Letters,* **30,** 1607, doi: 10.1029/2003GL017170.

Mann, M.E., R.S. Bradley, and M.K. Hughes, 1999: Northern Hemisphere temperatures during the past millennium: Inferences, uncertainties and limitations. *Geophys. Res. Lett.*, **26**, 759-762.

Mann, M.E., S. Rutherford, R.S. Bradley, M.K. Hughes, and F.T. Keimig, 2003: Optimal surface temperature reconstructions using terrestrial borehole data. *J. Geophys. Res.*, **108**, 4203, doi:10.1029/2002JD002532.

Mears, C.A., M.C. Schabel, and F.J. Wentz, 2003: A reanalysis of the MSU Channel 2 tropospheric temperature record. *J. Climate*, **16**, 3650–3664.

Meehl, G.A., F. Zwiers, J. Evans, T. Knutson, L. Mearns and P. Whetton, 2000: Trends in extreme weather and climate events: Issues related to modeling extremes in projections of future climate change. *Bull. Amer. Meteor. Soc*., **81**, 427-436.

Mitchell J. F. B, D.J. Karoly, G.C. Hegerl, F.W. Zwiers, and J. Marengo, 2001: Detection of climate change and attribution of causes. *Climate Change 2001. The Scientific Basis. The Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change.* J. T. Houghton et al. Eds., Intergovernmental Panel on Climate Change (IPCC), 880pp.

Munk, W., 2002: Twentieth century sea level: An enigma. *Proc. Nat. Ac. Sciences*, **99(10**), 6550-6555.

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

Osborn, T., 2002. The winter N Atlantic Oscillation: Roles of internal variability and greenhouse forcing. *Exchanges: CLIVAR selected research papers*, No. 25, 54-58.

Osborn, T. J. and M. Hulme, 1997: Development of a relationship between station and grid-box rainday frequencies for climate model evaluation. *J. Climate*, **10**, 1885–1908.

Ostermeier, G. and J. M. Wallace, 2003: Trends in the North Atlantic Oscillation-Northern Hemisphere Annular Mode during the Twentieth Century. *J. Climate*, **16**, 336-341.

Palmer, T.N., 1999a: A nonlinear dynamical perspective on climate prediction. *J. Climate*, **12,** 575-591.

Palmer, T. N., 1999b: Predicting uncertainty in forecasts of weather and climate, *Rep. Prog. Physics*, **63**, 71-116.

Palmer, T.N. and J. Raisanen, 2002: Quantifying the risk of extreme seasonal precipitation events in a changing climate. *Nature*, **415**, 512-514.

Parker, D.E., M. Gordon, D.P.N. Cullum, D.M.H. Sexton, C.K. Folland and N. Rayner, 1997: A new global gridded radiosonde temperature data base and recent temperature trends. *Geophys. Res. Lett*., **24**, 1499-1502.

Pollack, H.N. and J.E. Smerdon, 2004: Borehole climate reconstructions: Spatial structure and hemispheric averages. *J. Geophys. Res*. **109**, D11106, doi:10.1029/2003JD004163.

Ramaswamy, V., and Coauthors, 2001: Stratospheric temperature trends: Observations and model simulations. *Rev. Geophys*., **39**, 71-122.

Randel, W.J., F. Wu, and D.J. Gaffen, 2000: Interannual variability of the tropical tropopause derived from radiosonde data and NCEP reanalyses. *J. Geophys. Res*., **105**, 15509-15523.

Reichert, K.B., R. Schnur, and L. Bengtsson, 2002: Global ocean warming tied to anthropogenic forcing. *Geophys. Res. Lett*., **29 (11)**, 1525, doi:10.1029/2001GL013954.

Rind, D., P. Lonergan, J. Lean, D. Shindell, J. Perlwitz, J. Lerner, and C. McLinden, 2004: On the relative importance of solar and anthropogenic forcing of climate change from the Maunder Minimum to the present. *J. Climate*, prov. accepted.

Robock, A. and Y. Liu, 1994: The volcanic signal in Goddard Institute for Space Studies three-dimensional model simulations. *J. Climate*, **7**, 44-55.

Rutherford, S. and Mann, M.E., 2004: Correction to "Optimal surface temperature reconstructions using terrestrial borehole data", *J. Geophys. Res.*, **109**, D11107, doi:10.1029/2003JD004290.

Santer, B.D. and Coauthors, 2000: Interpreting differential temperature trends at the surface and in the lower troposphere. *Science*, **287**, 1227-1231.

Santer, B.D. and Coauthors, 2001: Accounting for the effects of volcanoes and ENSO in comparisons of modeled and observed temperature trends. *J. Geophys. Res*., **106**, 28,033-28,059.

Santer, B.D. and Coauthors, 2003a: Influence of satellite data uncertainties on the detection of externally-forced climate change. *Science*, **300**, 1280-1284.

Santer, B.D. and Coauthors, 2003b: Behavior of tropopause height and atmospheric temperature in models, reanalyses, and observations: Decadal changes. *J. Geophys. Res*., **108(D1)**, 4002, doi:10.1029/2002JD002258.

Santer, B.D. and Coauthors, 2003c: Contributions of anthropogenic and natural forcing to recent tropopause height changes. *Science*, **301**, 479-483.

Santer, B.D., et al., 2004: Identification of anthropogenic climate change using a second-generation reanalysis. *J. Geophys. Res. (Atmospheres)*, in press.

Schnur, R. and K. Hasselmann, 2004: Optimal filtering for Bayesian detection of climate change. *Climate Dyn.*, in press.

Seidel, D.J., R.J. Ross, J.K. Angell, and G.C. Reid, 2001: Climatological characteristics of the tropical tropopause as revealed by radiosondes. *J. Geophys. Res.*, **106**, 7857-7878.

Semenov V. A. and L. Bengtsson, 2002: Secular trends in daily precipitation characteristics: Greenhouse gas simulation with a coupled AOGCM. *Climate Dyn.,* **19**, 123-140.

Sexton, D. M. H., 2001: The effect of stratospheric ozone depletion on the phase of the Antarctic Oscillation. *Geophys. Res. Lett*., **28(9)**, 3697-3700.

Sexton, D. M. H., D. P. Rowell, C. K. Folland and D. J. Karoly, 2001: Detection of anthropogenic climate change using and atmospheric GCM. *Climate Dynamics*, **7**, 669-685.

Sexton, D. M. H., H. Grubb, K. P. Shine and C. K. Folland, 2003: Design and analysis of climate model experiments for the efficient estimation of anthropogenic signals. *J. Climate,* **16,** 1320-1336.

Shindell, D.T., R.L. Miller, G.A. Schmidt, and L. Pandolfo 1999. Simulation of recent northern winter climate trends by greenhouse-gas forcing. *Nature* **399**, 452-455.

Shindell, D.T., G.A. Schmidt, R.L. Miller and D. Rind, 2001: Northern Hemispheric climate response to greenhouse gas, ozone, solar and volcanic forcing. *J. Geophys. Res*., **106**, 7193-7210.

Simmonds, I. and K. Keay, 2000: Mean southern hemisphere extratropical cyclone behavior in the 40-Year NCEP-NCAR reanalysis. *J. Climate*, **13**, 873–885.

Smith, R.L., T.M.L. Wigley, and B.D. Santer, 2003: A bivariate time series approach to anthropogenic trend detection in hemispheric mean temperatures. *J. Climate*, **16**, 1228-1240.

Sokolov, A., C.E. Forest, and P.H. Stone, 2003: Comparing oceanic heat uptake in AOGCM transient climate change experiments. *J. Climate*, **16**, 1573-1582.

Stott, P.A., 2003: Attribution of regional-scale temperature changes to anthropogenic and natural causes. *Geophys. Res. Lett.*, **30**, doi: 10.1029/2003GL017324.

Stott, P. A. and J. A. Kettleborough, 2002: Origins and estimates of uncertainty in predictions of 21st century temperature rise. *Nature*, **416**, 723-726

Stott, P.A., G.S. Jones, and J.F.B. Mitchell, 2003: Do models underestimate the solar contribution to recent climate change? *J. Climate*, **16**, 4079–4093.

Sun, S. and J. Hansen, 2003: Climate simulations for 1951-2050 with a coupled atmosphere-ocean model. *J. Climate*, **16**, 2807-2826.

Tett, S.F.B. and Coauthors, 2002: Estimation of natural and anthropogenic contributions to 20th century temperature change. *J. Geophys. Res*., **107**, doi:10.1029/2000JD000028.

Thompson, D.W.J. and S. Solomon, 2002: Interpretation of recent Southern Hemisphere climate change. *Science*, **296**, 895-899.

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

Thompson, D.W.J., J. M. Wallace, and G.C. Hegerl, 2000: Annual modes in the extratropical circulation Part II: Trends. *J. Climate*, **13**, 1018-1036.

Trenberth, K.E. and D.A. Paolino, 1980: The Northern Hemisphere sea-level pressure data set: Trends, errors and discontinuities. *Mon. Weath. Rev.*, **108**, 855–872.

Van der Plassche, O., K. von der Borg, and A. F. M. de Jong, 1998: Sea level – climate correlation during the past 1400 years. *Geology*, **26**, 319-322.

Vinnikov K. Y. and N. C. Grody, 2003: Global warming trend of mean tropospheric temperature observed by satellites. *Science*, **302**, 269-272.

Wentz, F., M. Schabel, C. Mears, and D. Seidel, 2001: Lower tropospheric air temperature derived from a blended analysis of MSU, SSM/T1, SSM/I, NCEP/NCAR reanalysis, and Reynolds SST. Remote Sensing Systems Technical Report 062801, Santa Rosa, CA, 46 pp.

Yang, F., A. Kumar, M. Schlesinger, and W. Wang, 2003: Intensity of hydrological cycles in warmer climates. *J. Climate*, **16**, 2419-2423.

Zorita, E., F. González-Rouco, and S. Legutke, 2003: Testing the Mann et al. (1998) approach to paleoclimate reconstructions in the context of a 1000-yr control simulation with the ECHO-G coupled climate model. *J. Climate*, **16**, 1378–1390.

Zwiers, F. W. and X. Zhang, 2003: Toward regional scale climate change detection. *J. Climate*, **16,** 793-797.

Figure Captions

Figure 1: Several compilations of Northern Hemisphere temperatures for the last millennium, all smoothed with a 50-year Gaussian filter. For details of the series see Briffa and Osborn (2002). Calendar-year averages are shown in a) and ‘summer’ (April-September) averages are shown in b). The two grey lines in a) show the borehole records (dashed as given by Huang et al. 2000, and solid after areally weighting). Error estimates for some of the compilations are discussed in Jones and Mann (2004).

Figure 2: Contribution of external forcing to an updated record of Northern Hemispheric mean temperature north of 30 N (Crowley and Lowery 2000). The top section shows a comparison between the paleoreconstruction (black), the instrumental record (green) and a best combination of solar, volcanic and anthropogenic forcing from an Energy Balance Model simulation (red). The bottom panel shows the contribution from each forcing estimated by a multiple regression (the thick curve indicates the best guess and the thin curves indicate the 5 to 95% uncertainty range). An asterisk “*” denotes a response that is detected at the 5% significance level. Internal climate variability is estimated from the residual paleo variability. After Hegerl et al. (2003).

Figure 3: Simulated ocean heat content change over the last millennium from a simulation that matched the surface temperature record (c.f. Figure 2) and a similar simulation with slightly higher sensivity (3.0 K) and no solar forcing. Both simulations were started at 1AD to allow the simulation to adjust to the radiative forcing perturbations. The 20th century response to natural forcing is shown for reference. Also shown are the ocean heat content values from Levitus et al. (2000) and an estimate of relative sea level change from marsh deposits near Clinton, CT (van

der Plassche et al. 1998). Dashed lines refer to periods of no deposition or erosion (see Crowley et al. 2004).

Figure 4: Examples of detection and attribution (D&A) analyses under different assumptions regarding data retention. Consider mainly results to the left of the first vertical dashed line. The bars represent 5-95% confidence intervals. The upper panels refer to the scaling factor needed to make the predicted and observed patterns match, while the lower panels give the incremental warming associated with various forcings. a) D&A using all available observed data, b) D&A results after removing linear trend in global mean from the observed data and c) results from D&A using data from which all global mean information has been removed. Results on the right hand side of the panel indicate that a separation between different external climnate forcings (G greenhouse gas, N natural, S Sulphate, So Solar, V Volcanic) is increasingly difficult if global scale information is disregarded. Please see Allen et al. (2004) for a complete description and discussion of this illustration.

Figure 5: Results of a continental scale analysis of the detectability of a model simulated response to greenhouse gas and sulfate aerosol (GS) forcing in North American decadal mean temperatures (upper panel) and Eurasian decadal mean temperatures (lower panel) for 1950-1999. Vertical bars indicate the uncertainty of the scaling (expressed as 95% confidence intervals) on the GS signal that is required to make the best fit with observations. The GS signal is obtained by combining two ensembles of transient climate change simulations performed with two versions of the Canadian Centre for Climate Modelling and Analysis global climate model. Results are shown as a function of the number of regional EOFs of decadal mean temperature

that are retained in the analysis. Two confidence intervals are shown for each truncation, reflecting two independent estimates of the internal variability obtained from control runs performed with the CCCma models. Confidence intervals that exclude zero provide evidence for the detection of the GS signal at the 5% level of significance, and confidence intervals that include unity provide evidence in support of an attribution claim. Results are generally not very sensitive to truncation. After Zwiers and Zhang (2003, Figure 2).

Figure 6: Comparison of HadCM3 all forcings runs with the observed decadal mean temperature changes in 1900-2000 for 16 sub-continental scale regions. The four dashed green curves represent the individual ensemble members, the solid green curve represents the ensemble mean, and observations are given in black. The model captures many features of the observed temperature changes, such as the steady warming in Southern Africa (SAF), Southern South America (SSA), and South Asia (SAS). Also apparent is the variability in Northern Europe (NRU) with an early century warming and a late century warming, and accelerating warming in Sahara region (SAH). Central America (CAM) also has early and late century warming in both model and observations. North America is not particularly well captured with no early century warming in Western N. America (WNA) in the model. South Australia (SAU) is also not well captured. In mid Asia (MAS) all model simulations suggest a cooling in the middle of the century while the observations show a general warming. Generally the model captures many features of both the variability and the trends in the different regions. Other regions shown are Eastern Southern N. America (ESNA), Eastern Northern N. America and Greenland (GRL), Northern Asia (NAS), Amazon (AMZ), Southern S. America (SSA), Central Africa (CAF),

Southern Australia (SAU), Mediterranean (MDB). The geographical domains of the regions are defined in Stott (2003, Table 1).

Figure 7: Trends in anthropogenically forced model simulations and in observations over 1950-1999. The error bars on the model trends are 5-95% confidence intervals for the ensemble-mean trends, estimated by resampling the long control simulations form the respective models and allowing for the number of members in each ensemble. The error bars about zero at the location of the observed trends are uncertainties of the trend estimates due to nature internal climate variability, as simulated by the models. They are 5-95% confidence intervals for a single realization, estimated from control simulations from the ECHAM4, HadCM2 and PCM models, which are the only one with DTR data available. From Karoly et al. (2004).

Figure 8: Changes in global-mean monthly-mean mid- to upper tropospheric temperatures (MSU channel 2) over 1979 to 2000. The Christy et al. (2000) version of the MSU channel 2 data shows virtually no overall warming (+0.014°C/decade). Independent reprocessing of the raw channel 2 radiances by Wentz et al. (2001) yields a warming of +0.142°C/decade over 1979-2000. Anomalies in both data sets are defined relative to their respective climatological monthly means over 1979-1997.

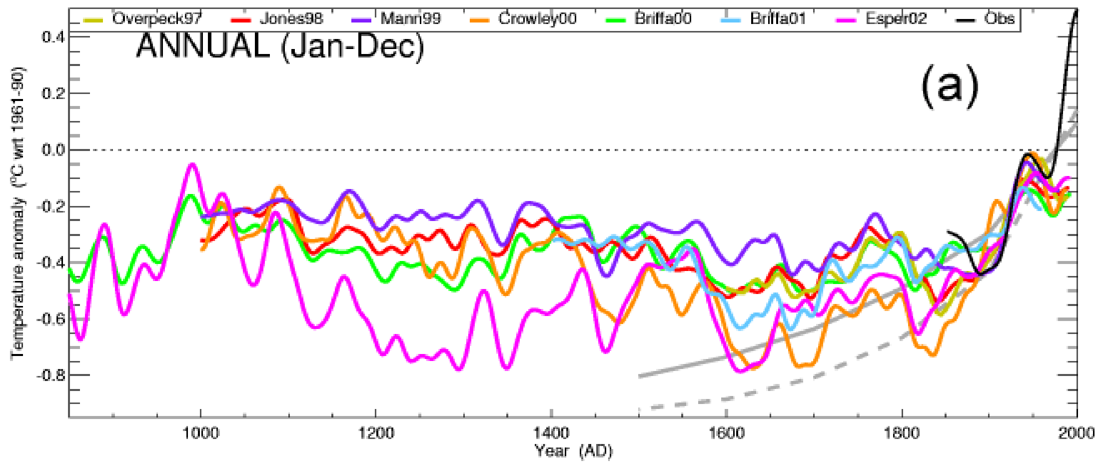
Figure 9: Statistical significance of differences between simulated and observed channel 2 temperature trends. Observed trends are from two different sources (Christy et al. 2000, and Wentz et al. 2001). Model equivalent channel 2 trends are from the “GSOP” experiment performed by Bengtsson *et al*. (1999) with the ECHAM4/OPYC coupled model. The comparison

involves “residual” trends after removal of estimated ENSO and volcano effects from model and observed data (Santer et al. 2001). The range of residual trends arises from uncertainties in τ (the assumed decay time for a volcanic signal) and from the choice of index used for removal of ENSO influences. All trends were computed with global-mean monthly-mean data spanning the 228-month period January 1979 through December 1997, the period of the GSOP integration. Model channel 2 trends are not significantly different from the Wentz *et al*. (2001) results, but are generally inconsistent with Christy et al. (2000).

Figure 10: Change in annual mean precipitation (left panels) and the wettest day per year (right panels) in CGCM1 (top) and HadCM3 (middle) at the time of CO2 doubling. Changes are expressed as a percentage of the present day climatological value. The scale ranges from -30% to +30%. Changes are only plotted where they are significant at the 10% level according to a Mann-Whitney test. The bottom two panels show the average of climate change patterns from both models where the large-scale (smoothed) changes are consistent between the models (inconsistency at the 10% level determined by a Mann-Whitney test using the ensemble member simulations). (See Hegerl et al. 2004).

Figure 11: Optimal Bayesian detection and attribution analysis for annual means of surface temperature (TAS), precipitation (PREC), and summer (JJA) and winter (DJF) means of diurnal temperature range (DTR). (a) The number of EOFs retained from the ECHAM3/LSG control run (spanning the initial phase space), and the number of patterns retained after applying the optimal filtering in the Bayesian and conventional detection and attribution analysis, respectively.(b) The prior and posterior probabilities for each of the climate change hypotheses (climate change can

be explained by natural internal variability (“Natural”), greenhouse gases alone (“GHG”), greenhouse gases plus sulfate aerosols (“GHG+S”)). Uniform priors were assumed giving each hypothesis a probability of 1/3. The net posteriors refer to the posterior probabilities if evidence from all four variables is considered simultaneously. The two small values for “Natural” are actually almost zero and have been inflated for display purposes only. See Schnur and Hasselmann (2004), for details.



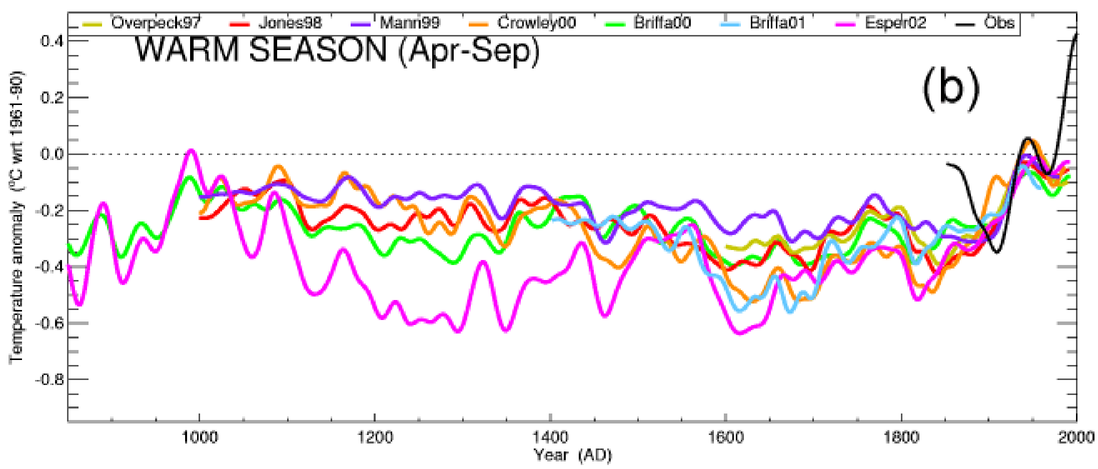


Figure 1: Several compilations of Northern Hemisphere temperatures for the last millennium, all smoothed with a 50-year Gaussian filter. For details of the series see Briffa and Osborn (2002). Calendar-year averages are shown in a) and 'summer' (April-September) averages are shown in b). The two grey lines in a) show the borehole records (dashed as given by Huang et al. 2000, and solid after areally weighting). Error estimates for some of the compilations are discussed in Jones and Mann (2004).

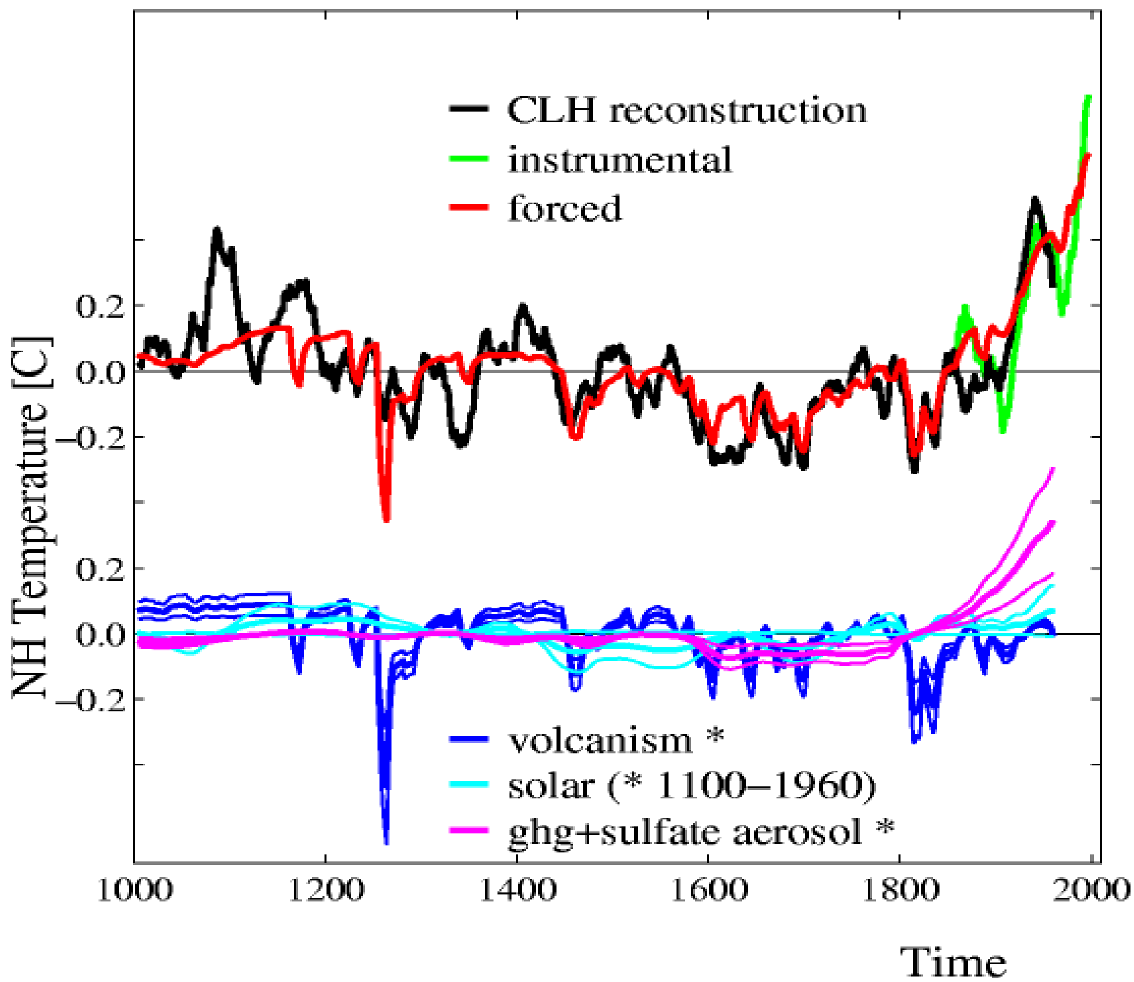


Figure 2: Contribution of external forcing to an updated record of Northern Hemispheric mean temperature north of 30 N (Crowley and Lowery 2000). The top section shows a comparison between the paleoreconstruction (black), the instrumental record (green) and a best combination of solar, volcanic and anthropogenic forcing from an Energy Balance Model simulation (red). The bottom panel shows the contribution from each forcing estimated by a multiple regression (the thick curve indicates the best guess and the thin curves indicate the 5 to 95% uncertainty range). An asterisk "*" denotes a response that is detected at the 5% significance level. Internal climate variability is estimated from the residual paleo variability. After Hegerl et al. (2003).

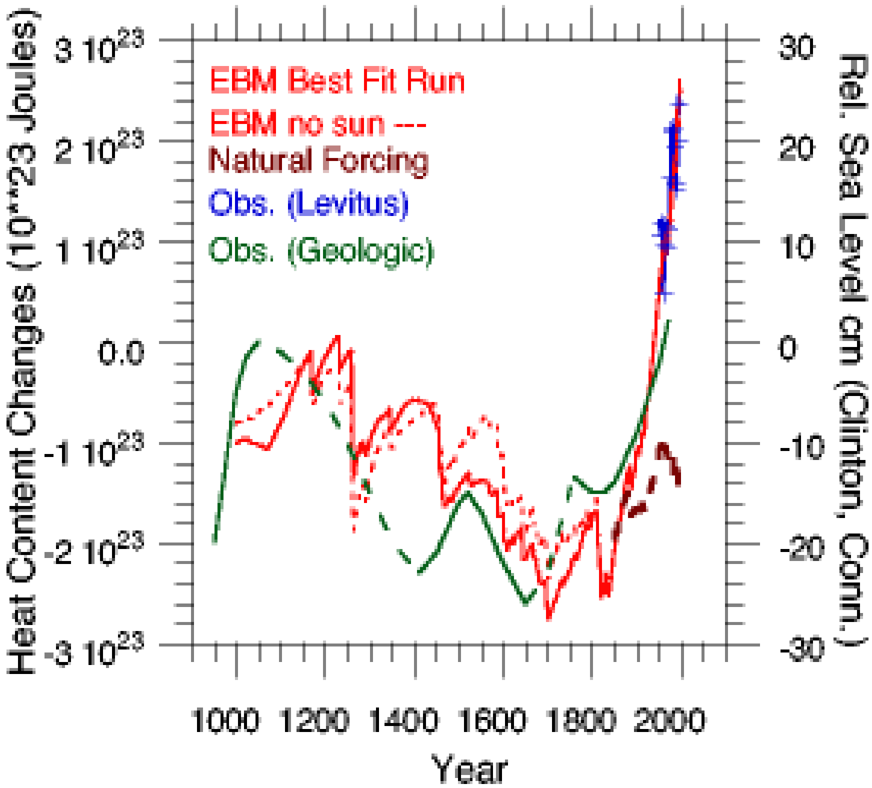


Figure 3: Simulated ocean heat content change over the last millennium from a simulation that matched the surface temperature record (c.f. Figure 2) and a similar simulation with slightly higher sensivity (3.0 K) and no solar forcing. Both simulations were started at 1AD to allow the simulation to adjust to the radiative forcing perturbations. The 20th century response to natural forcing is shown for reference. Also shown are the ocean heat content values from Levitus et al. (2000) and an estimate of relative sea level change from marsh deposits near Clinton, CT (van der Plassche et al. 1998). Dashed lines refer to periods of no deposition or erosion (see Crowley et al. 2004).

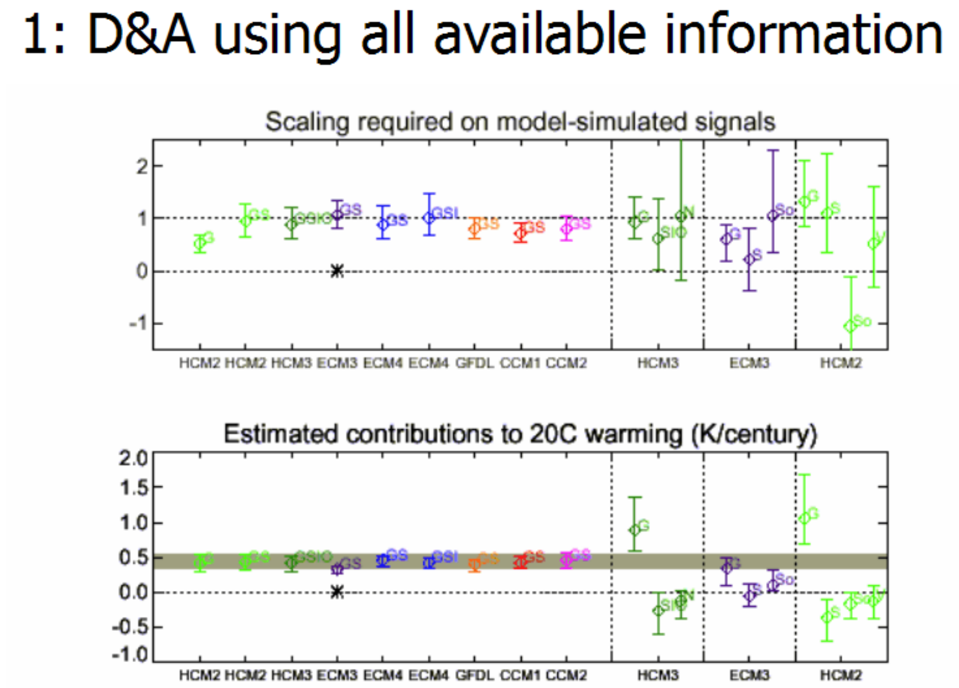


Figure 4a

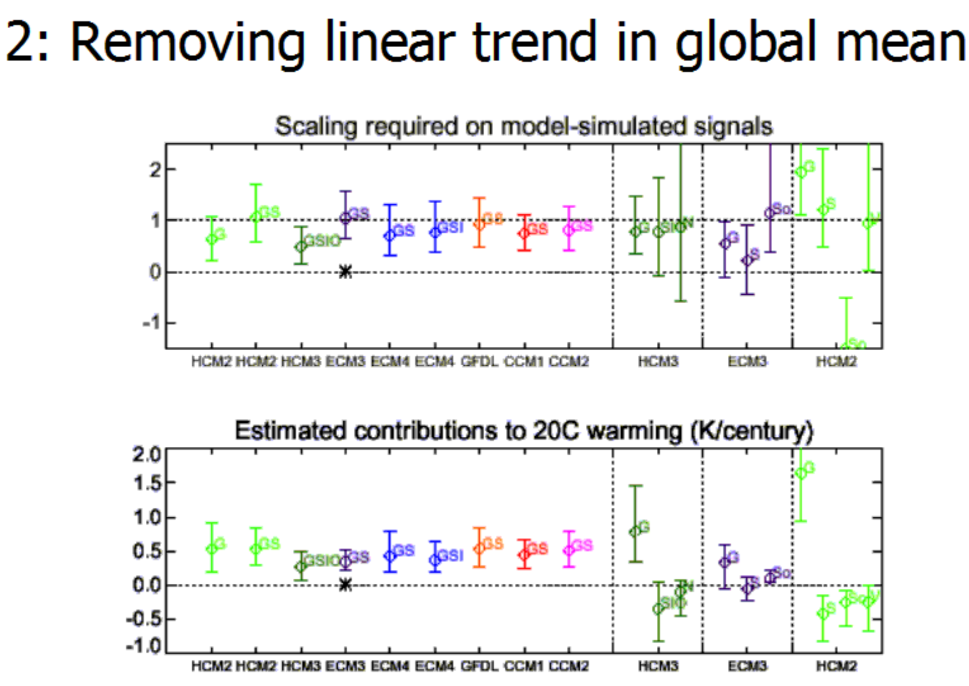


Figure 4b

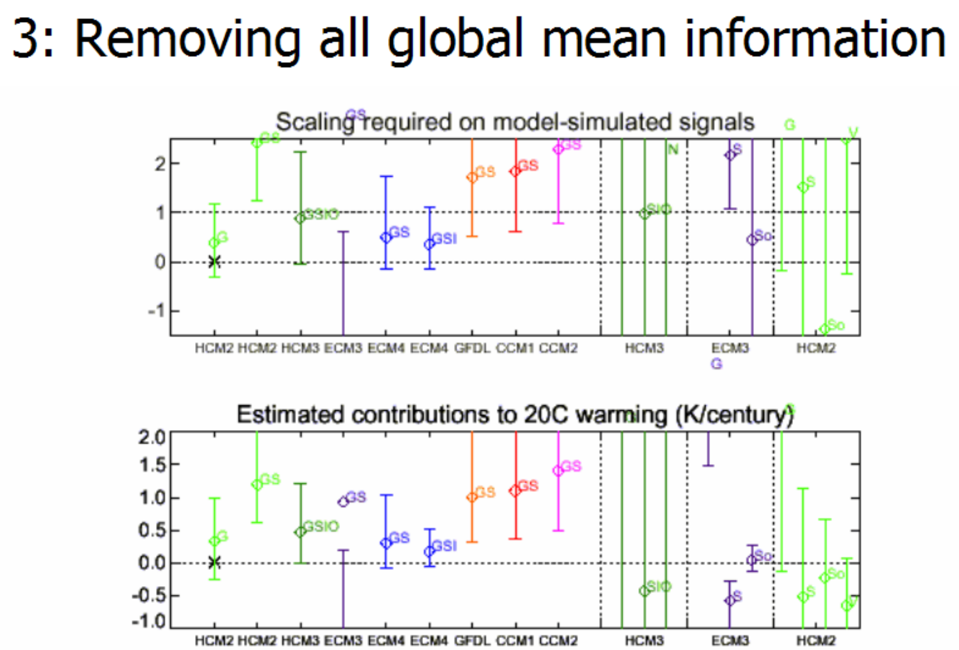


Figure 4c

Figure 4: Examples of detection and attribution (D&A) analyses under different assumptions regarding data retention. Consider mainly results to the left of the first vertical dashed line. The bars represent 5-95% confidence intervals. The upper panels refer to the scaling factor needed to make the predicted and observed patterns match, while the lower panels give the incremental warming associated with various forcings. a) D&A using all available observed data, b) D&A results after removing linear trend in global mean from the observed data and c) results from D&A using data from which all global mean information has been removed. Results on the right hand side of the panel indicate that a separation between different external climnate forcings (G greenhouse gas, N natural, S Sulphate, So Solar, V Volcanic) is increasingly difficult if global scale information is disregarded. Please see Allen et al. (2004) for a complete description and discussion of this illustration.

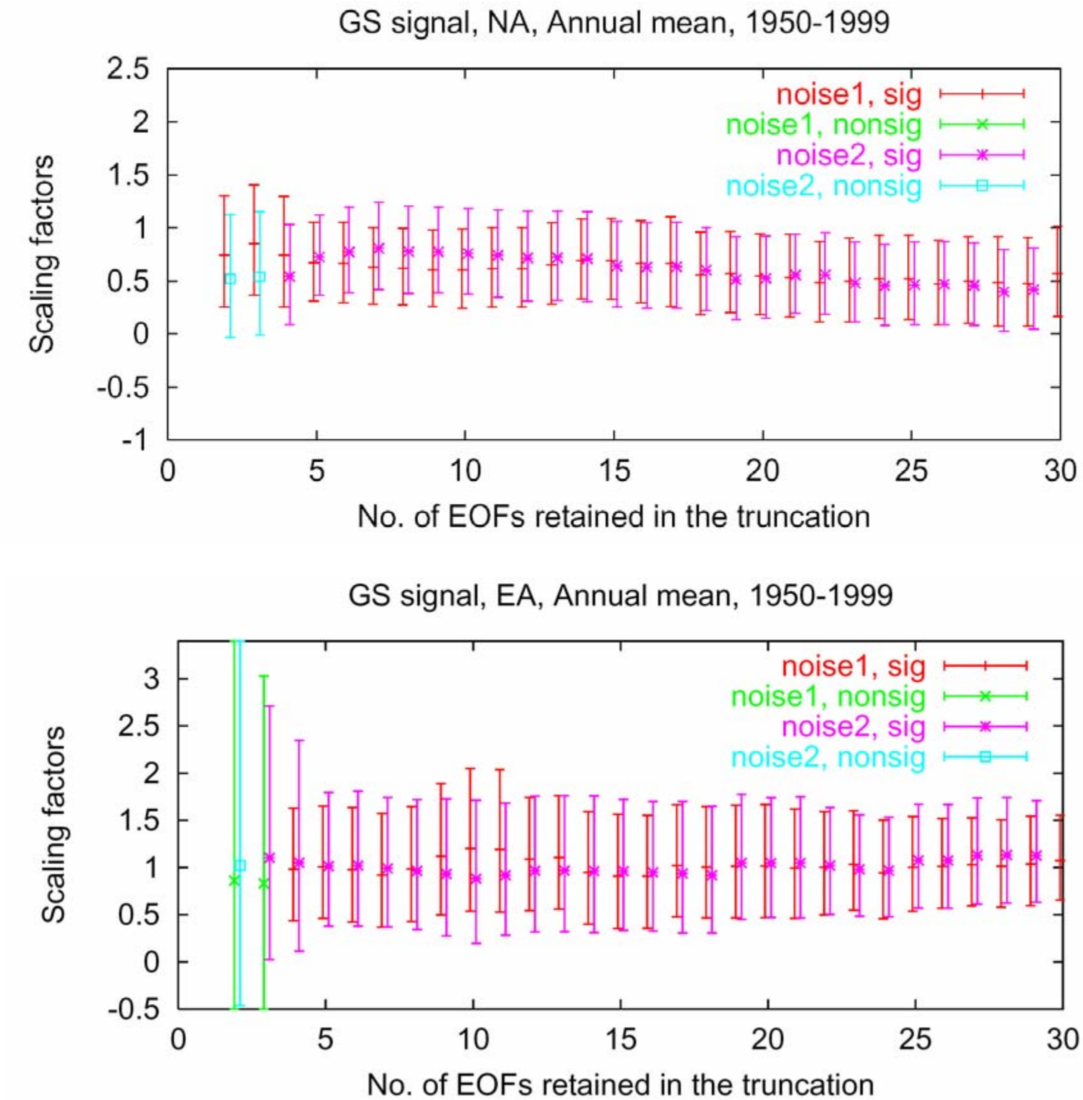


Figure 5: Results of a continental scale analysis of the detectability of a model simulated response to greenhouse gas and sulfate aerosol (GS) forcing in North American decadal mean temperatures (upper panel) and Eurasian decadal mean temperatures (lower panel) for 1950-1999. Vertical bars indicate the uncertainty of the scaling (expressed as 95% confidence intervals) on the GS signal that is required to make the best fit with observations. The GS signal is obtained by combining two ensembles of transient climate change simulations performed with two versions of the Canadian Centre for Climate Modelling and Analysis global climate model. Results are shown as a function of the number of regional EOFs of decadal mean temperature that are retained in the analysis. Two confidence intervals are shown for each truncation, reflecting two independent estimates of the internal variability obtained from control runs performed with the CCCma models. Confidence intervals that exclude zero provide evidence for the detection of the GS signal at the 5% level of significance, and confidence intervals that include unity provide evidence in support of an attribution claim. Results are generally not very sensitive to truncation. After Zwiers and Zhang (2003, Figure 2).

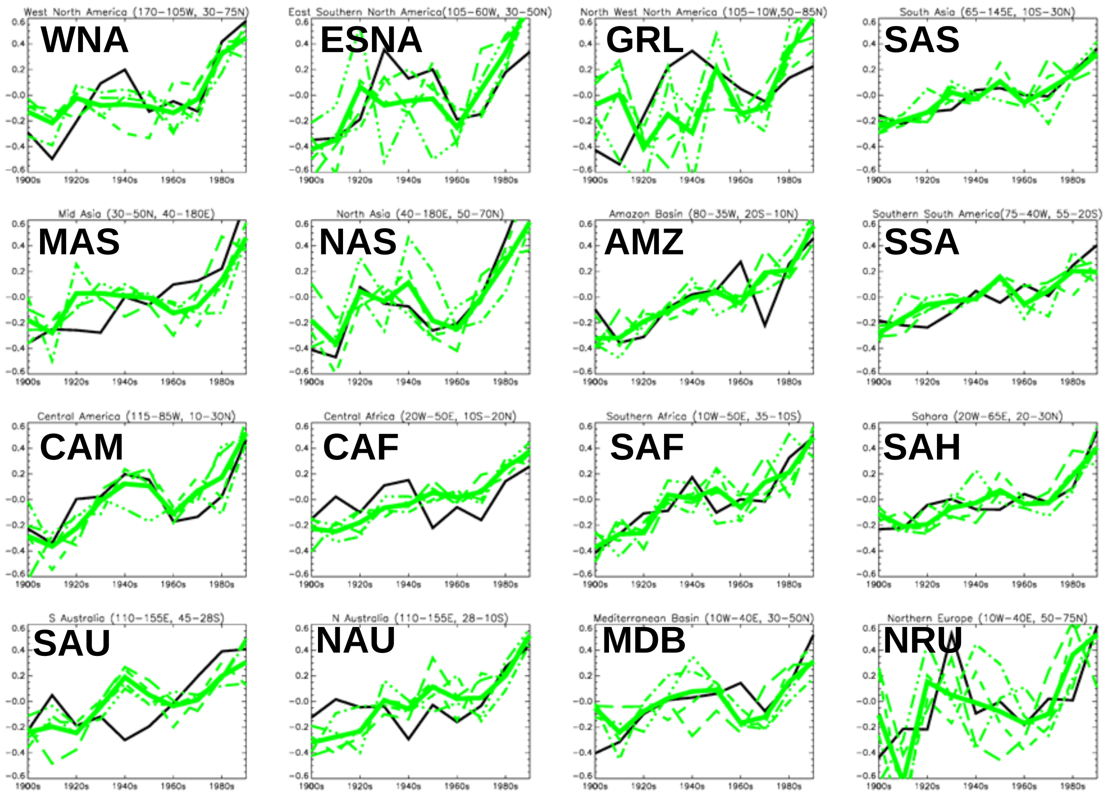


Figure 6: Comparison of HadCM3 all forcings runs with the observed decadal mean temperature changes in 1900-2000 for 16 sub-continental scale regions. The four dashed green curves represent the individual ensemble members, the solid green curve represents the ensemble mean, and observations are given in black. The model captures many features of the observed temperature changes, such as the steady warming in Southern Africa (SAF), Southern South America (SSA), and South Asia (SAS). Also apparent is the variability in Northern Europe (NRU) with an early century warming and a late century warming, and accelerating warming in Sahara region (SAH). Central America (CAM) also has early and late century warming in both model and observations. North America is not particularly well captured with no early century warming in Western N. America (WNA) in the model. South Australia (SAU) is also not well captured. In mid Asia (MAS) all model simulations suggest a cooling in the middle of the century while the observations show a general warming. Generally the model captures many features of both the variability and the trends in the different regions. Other regions shown are Eastern Southern N. America (ESNA), Eastern Northern N. America and Greenland (GRL), Northern Asia (NAS), Amazon (AMZ), Southern S. America (SSA), Central Africa (CAF), Southern Australia (SAU), Mediterranean (MDB). The geographical domains of the regions are defined in Stott (2003, Table 1).

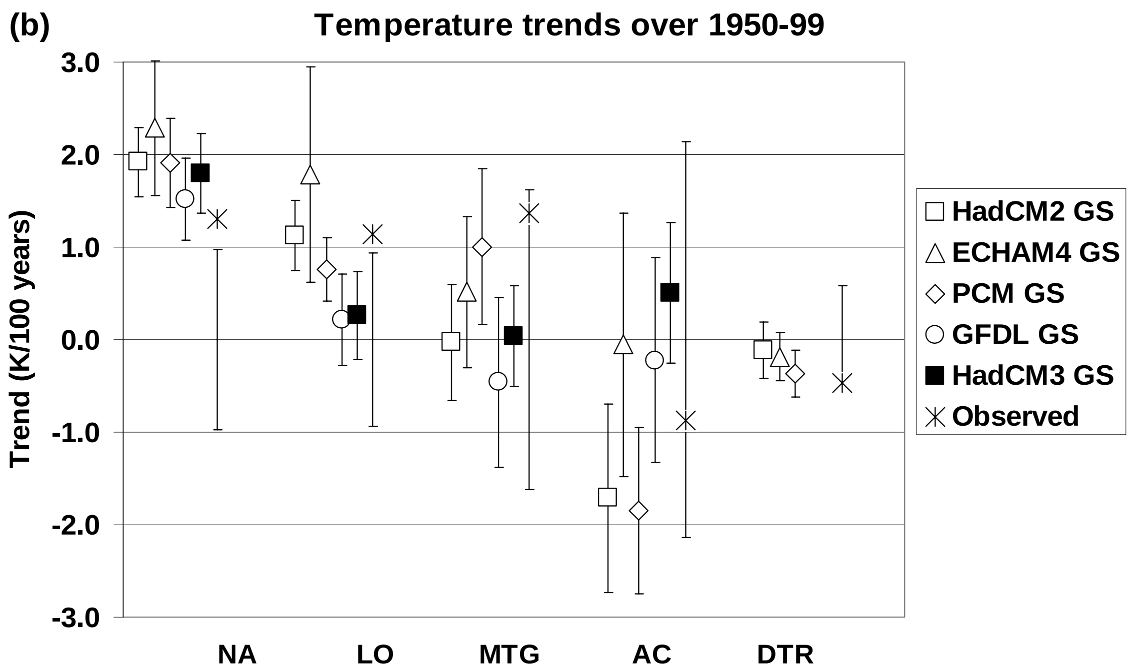


Figure 7: Trends in anthropogenically forced model simulations and in observations over 1950-1999. The error bars on the model trends are 5-95% confidence intervals for the ensemble-mean trends, estimated by resampling the long control simulations form the respective models and allowing for the number of members in each ensemble. The error bars about zero at the location of the observed trends are uncertainties of the trend estimates due to nature internal climate variability, as simulated by the models. They are 5-95% confidence intervals for a single realization, estimated from control simulations from the ECHAM4, HadCM2 and PCM models, which are the only one with DTR data available. From Karoly et al. (2004).

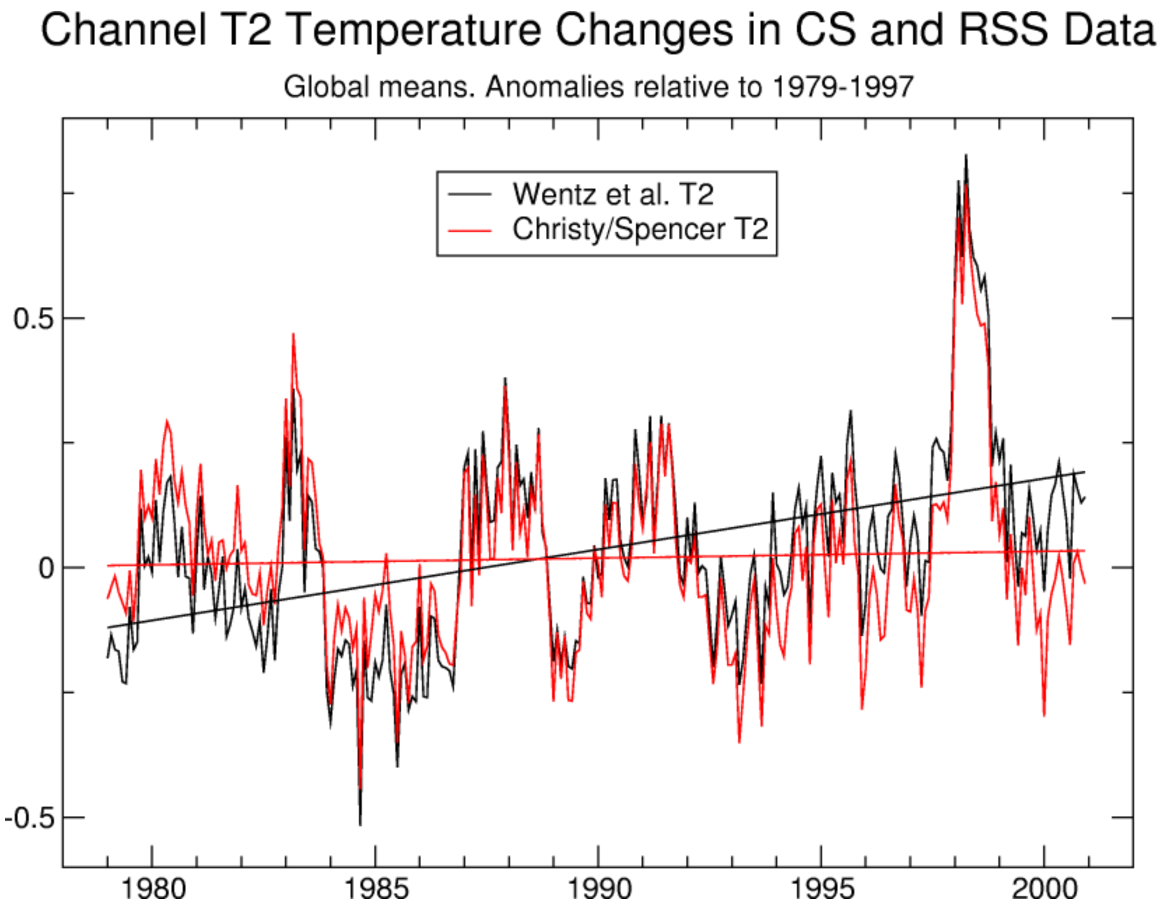


Figure 8: Changes in global-mean monthly-mean mid- to upper tropospheric temperatures (MSU channel 2) over 1979 to 2000. The Christy et al. (2000) version of the MSU channel 2 data shows virtually no overall warming (+0.014°C/decade). Independent reprocessing of the raw channel 2 radiances by Wentz et al. (2001) yields a warming of +0.142°C/decade over 1979-2000. Anomalies in both data sets are defined relative to their respective climatological monthly means over 1979-1997.

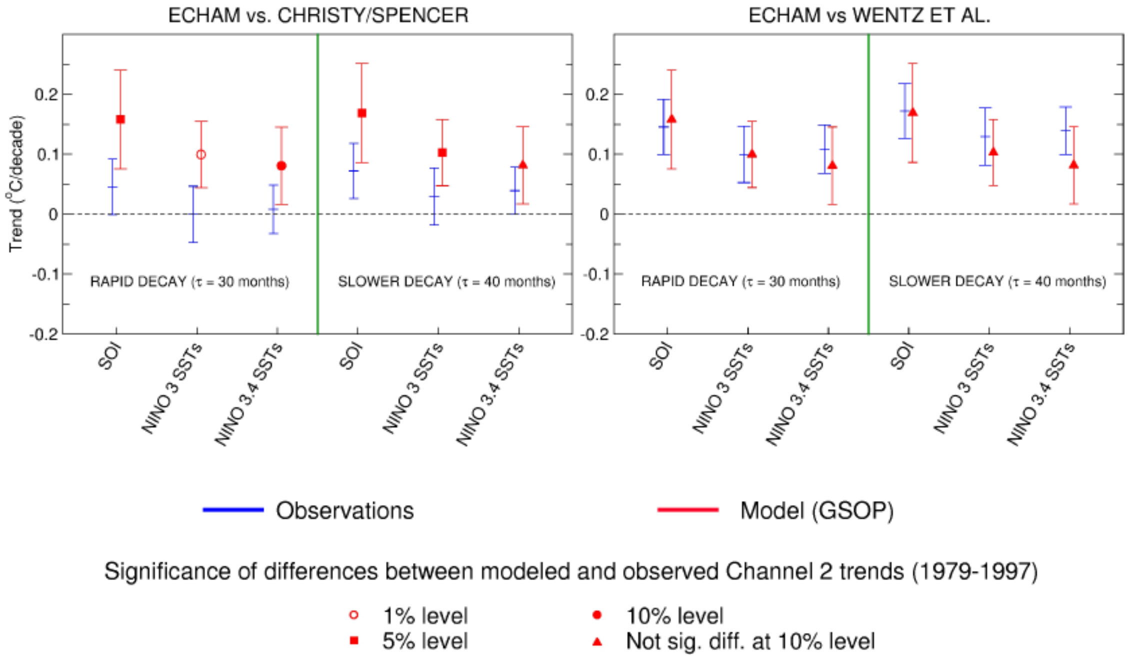


Figure 9: Statistical significance of differences between simulated and observed channel 2 temperature trends. Observed trends are from two different sources (Christy et al. 2000, and Wentz et al. 2001). Model equivalent channel 2 trends are from the "GSOP" experiment performed by Bengtsson *et al*. (1999) with the ECHAM4/OPYC coupled model. The comparison involves "residual" trends after removal of estimated ENSO and volcano effects from model and observed data (Santer et al. 2001). The range of residual trends arises from uncertainties in τ (the assumed decay time for a volcanic signal) and from the choice of index used for removal of ENSO influences. All trends were computed with global-mean monthly-mean data spanning the 228-month period January 1979 through December 1997, the period of the GSOP integration. Model channel 2 trends are not significantly different from the Wentz *et al*. (2001) results, but are generally inconsistent with Christy et al. (2000).

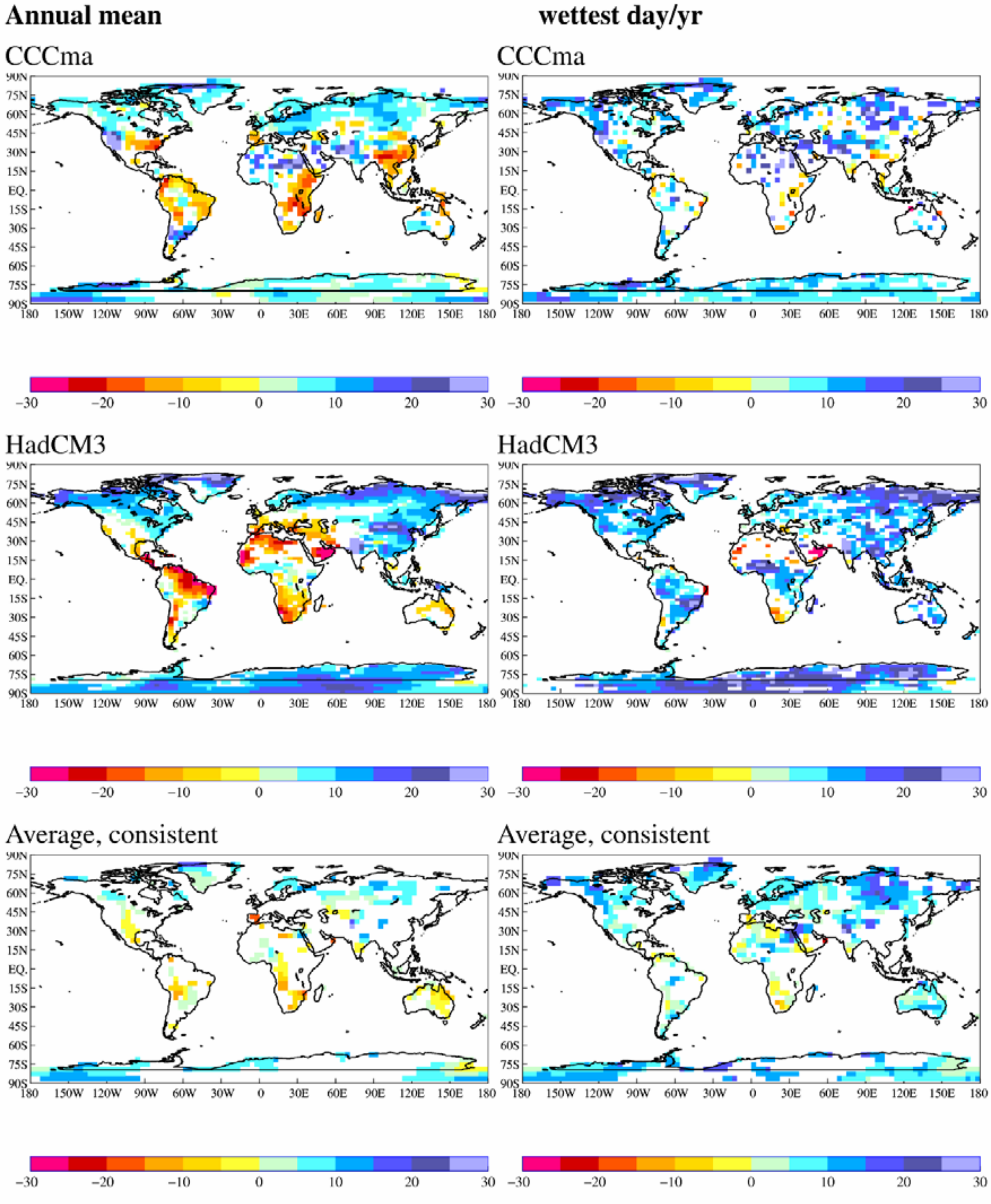


Figure 10: Change in annual mean precipitation (left panels) and the wettest day per year (right panels) in CGCM1 (top) and HadCM3 (middle) at the time of CO2 doubling. Changes are expressed as a percentage of the present day climatological value.The scale ranges from -30% to +30%. Changes are only plotted where they are significant at the 10% level according to a Mann-Whitney test. The bottom two panels show the average of climate change patterns from both models where the large-scale (smoothed) changes are consistent between the models (inconsistency at the 10% level determined by a Mann-Whitney test using the ensemble member simulations). (See Hegerl et al. 2004).

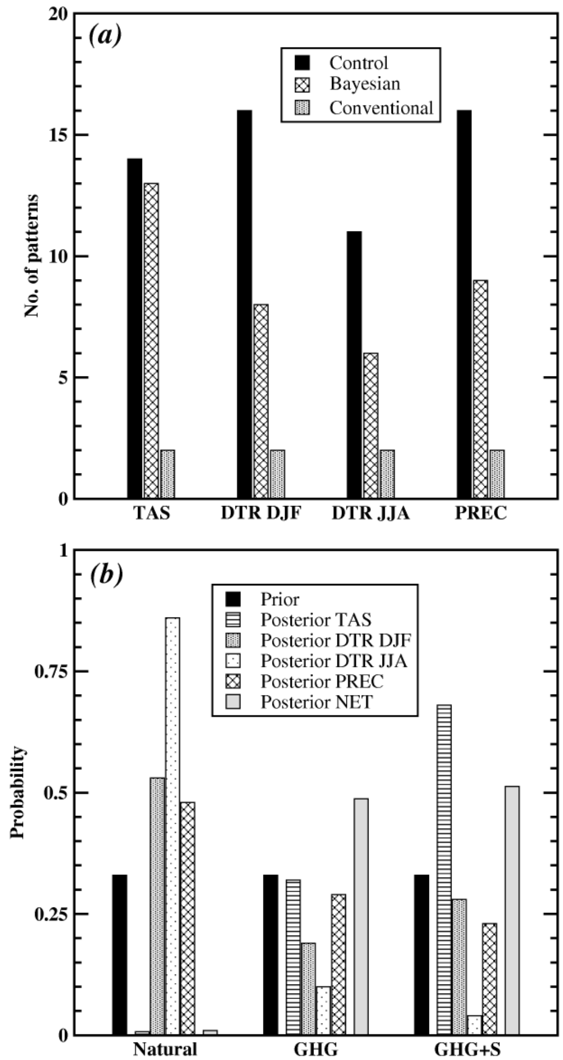


Figure 11: Optimal Bayesian detection and attribution analysis for annual means of surface temperature (TAS), precipitation (PREC), and summer (JJA) and winter (DJF) means of diurnal temperature range (DTR). (a) The number of EOFs retained from the ECHAM3/LSG control run (spanning the initial phase space), and the number of patterns retained after applying the optimal filtering in the Bayesian and conventional detection and attribution analysis, respectively.(b) The prior and posterior probabilities for each of the climate change hypotheses (climate change can be explained by natural internal variability ("Natural"), greenhouse gases alone ("GHG"), greenhouse gases plus sulfate aerosols ("GHG+S")). Uniform priors were assumed giving each hypothesis a probability of 1/3. The net posteriors refer to the posterior probabilities if evidence from all four variables is considered simultaneously. The two small values for "Natural" are actually almost zero and have been inflated for display purposes only. See Schnur and Hasselmann (2004), for details.