EXPERT ASSESSMENT OF UNCERTAINTIES IN DETECTION AND ATTRIBUTION OF CLIMATE CHANGE

BY JAMES S. RISBEY AND MILIND KANDLIKAR

The problem of detection of climate change and attribution of causes of change has been formalized as a series of discrete probability judgements by experts

The study of detection and attribution of climate change addresses the issue of whether, and to what extent, human-induced increases in greenhouse gases have caused climate changes. Communication of the science and role of uncertainties on this issue has been hindered by the lack of explicit formal approaches for making *overall* conclusions on detection and attribution. We have developed a protocol to quantify uncertainties in each component of the detection and attribution process and to provide a structured way to make overall conclusions (Risbey et al. 2000). Here we describe results from use of the protocol with a set of climate experts.

Studies of detection and attribution of climate change have focused first on "detecting" climate change against the background of natural variability, and further on "attributing" any detected signal to increases in greenhouse gases or other possible causes (Hasselmann 1998; Hegerl et al. 1997; Santer et al. 1996a; Zwiers 1999). Summary assessments of the detection and attribution issue have tended to take a qualitative approach to characterizing uncertainties (Santer et al. 1996a; Barnett et al. 1999; Mitchell et al. 2001). This study outlines some of the major uncertainties in detection and attribution, uses expert judgements to quantify them, and gives an overall quantitative assessment of detection and attribution of climate change.

In making overall assessments on detection and attribution of climate change, a variety of scientific judgements are called for. These relate to the quality of the underlying data needed to monitor climate changes, to the quality of models used to assess possible causes of those changes, to the ability to monitor the forcing of the earth system, and to model future consequences. The probability-based protocol employed here makes many of these underlying judgements explicit. The protocol breaks the detection and attribution problem up into its component steps and allows for uncertainties on each of these judgements to be represented quantitatively via probability density functions (pdfs). The protocol was completed by a set of 19 experts working in the field

AFFILIATIONS: RISBEY—School of Mathematical Sciences, Monash University, Melbourne, Australia; KANDLIKAR—Faculty of Graduate Studies, University of British Columbia, Vancouver, Canada
CORRESPONDING AUTHOR: Dr. James S. Risbey, School of Mathematical Sciences, P.O. Box 28M, Monash University, Clayton, Vic 3800, Australia
E-mail: ris@cmu.edu

In final form 31 May 2002
©2002 American Meteorological Society

Table 1. List of experts and their institutional affiliations.

Expert	Affiliation
Myles Allen	Oxford University
Robert Balling	Arizona State University
Curt Covey	Lawrence Livermore National Laboratory
Chris Folland	Hadley Centre
Klaus Hasselmann	Deutsches Klimarechenzentrum
Gabi Hegerl	Texas A&M University
Phil Jones	University of East Anglia
David Karoly	Monash University
Haroon Kheshgi	Exxon Corporate Research
Michael MacCracken	U.S. Global Change Research Program
Nathan Mantua	University of Washington
Patrick Michaels	University of Virginia
John Mitchell	Hadley Centre
Neville Nicholls	Australian Bureau of Meteorology
Stephen Schneider	Stanford University
John Wallace	University of Washington
Warren Washington	National Center for Atmospheric Research
Tom Wigley	National Center for Atmospheric Research
Francis Zwiers	Canadian Climate Center

of detection and attribution studies (see Table 1). Note that the order of experts has been scrambled in presenting results to prevent identification with particular experts. Twenty-three experts were approached to complete the protocol and four declined to do so. The experts were selected to provide depth and diversity in a range of different areas. For example, some were data specialists, some modelers, some focused on methods in detection and attribution, some were from the Intergovernmental Panel on Climate Change (IPCC) detection and attribution groups (Santer et al. 1996b; Mitchell et al. 2001), and some were not in those groups. Further, they entertained a range of views about the likely magnitude of greenhouse climate change. This sample of 19 experts is considered to be indicative, but not exhaustive.

The first step in the detection and attribution process is the *determination* of a set of lines of evidence, E_i, $i = 1, I$, which serve as signifiers of climate change. For each expert the protocol elicits lines of evidence that are relatively independent in order to minimize redundant information.[1] The expert must then *evaluate* the magnitude and uncertainty in each line of evidence, since each will be associated with some uncertainty deriving from incomplete and biased datasets. This is done by eliciting the pdf, $f(E_i)$, for the observed change over a specified period for each line of evidence from the expert. The *detection* step entails discriminating forced climate changes (Shine and Forster 1999) from internal natural variability (Pelletier 1997). This requires an elicitation of the pdf, $f(N_i)$, for natural variability in each line of evidence over the same time period. The distributions, $f(E_i)$ and $f(N_i)$ can then be compared using statistical tests to see how similar they are.

Following (Bell 1986) and (Zwiers 1999), the detector, D_i, is defined as $D_i = N(0, \sigma_{ni}^2)$ when $S = 0$ (null hypothesis); $D_i = N(S, \sigma_{si}^2)$ when $S \neq 0$ (alternative hypothesis), where $D \sim N(\mu, \sigma^2)$ indicates that D is normally distributed with the mean, μ, and variance, σ^2. The null hypothesis of no climate change is rejected at the $p \times 100\%$ significance level using a one-sided z test if the value of the detector exceeds a critical value $D_c = Z_{1-p}\sigma_{ni}^2$, where Z_{1-p} is the $(1 - p)$ quantile of the standard normal distribution. The signal is "detected" when the null hypothesis is rejected. The power of the test or probability of detecting the signal when it is present is given by

$$P(D_i > D_c) = \frac{1}{\sqrt{2\pi}\sigma_{si}} \int_{D_c}^{\infty} \exp^{-\frac{1}{2}(D-S)^2/\sigma_{si}^2} dD. \quad (1)$$

The conventional form of detection described above uses only information on the critical percentile of $f(N_i)$, and may be sensitive to the choice of critical percentile. For that reason we perform detection tests for a range of values of the critical percentile (1%, 5%, 10%, 20%). Detection implies rejection of the null hypothesis at the critical percentile. This does not identify the cause of any climate change or the contribution of natural variability to observed changes. It merely identifies the likelihood that the change exceeds the critical threshold.

[1] For example, changes in global mean sea level are not considered because they are a fairly direct consequence of changes in global mean temperature, which is a chosen line of evidence.

For the attribution step, responsibility in explaining observed trends (E_i) is partitioned out among competing causes. These causes can include natural variability as well as a range of various forcing processes. The attribution step requires the experts to identify this set of forcings, and to characterize their patterns and magnitude. The set of forcings considered include changes in greenhouse gas forcing, aerosol forcing, solar forcing, volcanic forcing, and ozone forcing. The uncertainty in each of the *j* forcings relevant to a line of evidence, *i,* can also be represented by a pdf, $f(F_{i,j})$. Once the forcings have been characterized, the next step is to elicit pdfs *attributing* fractional responsibilities from 0 (no responsibility) to 1 (complete responsibility) for each forcing in explaining the detected signal. Note that any given forcing could explain a vanishingly small fraction of the detected signal with high probability or almost all of the signal with vanishingly low probability. Calculation of the attribution to each forcing requires a convolution of the forcing pdf and attribution pdf. This step can be simplified for the expert elicitation process by discretizing the range of values that each forcing, $f(F_{i,j})$ can take, and by eliciting just the expected fraction, $Q_{i,j}$, of the detected signal for line of evidence, *i,* accounted for by each forcing, *j.* These fractions provide a breakdown of the contributions of only those forcings that could have caused a change in E_i in the direction expected. For example, when the change in E_i is positive, the $Q_{i,j}$ fractions provide a breakdown of the total positive forcing ($\Sigma_j Q_{i,j} = 1$).

The attribution model described here presupposes a linear-additive model of the forcing, which is not necessarily the case (Ramaswamy et al. 2001). Our approach differs from more conventional approaches to attribution described in (Zwiers 1999) in that we do not explicitly compare model outputs to observations to make attributions to given forcings. That approach measures consistency between forced model runs and observations, but could still yield high attributions in the case where there was no signal present—that is, the forced run is similar to observations, but the observations are dominated by natural variability. In the present approach, the experts provide assessments of contributions from different forcings, conditioned on the uncertainties in those forcings. That is, the signal is partitioned on the basis of estimates of the relevant forcing.

The final share of responsibility accorded to each of the forcings ($A_{i,j}$) in partitioning changes in E_i in the expected direction is adjusted to allow for contributions from natural variability ($A_{i,\mathrm{nv}}$), such that $\Sigma_j A_{i,j} + A_{i,\mathrm{nv}} = 1$. In a hypothetical world in which we had many realizations of climate change over the recent period, the expected contribution from natural variability to the observed change would be zero. But since we have only one real observed period to draw from, it is possible that there were long-term excursions of natural variability during this particular period that contributed to the observed change. As for the forcing, natural variability can only contribute when it has the same sign as the expected change. The fractional contribution of natural variability is calculated from those cases where natural variability makes a contribution to the expected change and is given by the ratio N_i/E_i. Sampling from $f(E_i)$ and $f(N_i)$ in this way leads to an expected contribution from natural variability, $A_{i,\mathrm{nv}}$, which can range from 0 to 1. Note that when $f(E_i)$ and $f(N_i)$ are identical distributions, $A_{i,\mathrm{nv}} = 1$, and natural variability is responsible for all of the observed change.

The results presented here for each step in the above process were obtained in a set of detailed personal interviews (lasting between 4 and 8 h each) where the experts offered probabilistic judgements as well as the underlying rationale for their judgements. Though results included attributions to all potentially significant forcings, we isolate only outcomes for greenhouse forcing here. Greenhouse forcing is potentially most problematic among the forcings because of its magnitude and expected persistence (Hansen and Lacis 1990; Hansen et al. 1998; Kasting 1998).

RESULTS. *Lines of evidence.* Four primary lines of evidence resulting from the protocol are the following:

M: The century-long trend in global mean surface temperature (Jones et al. 1999);
V: The past 30-yr trend in vertical pattern of temperature (Santer et al. 1996a; Tett et al. 1996; Folland et al. 2001);
G: The past 30-yr trend in geographical pattern of surface temperature (Santer et al. 1996b);
D: The past 30-yr trend in diurnal temperature range over land (Karl et al. 1993; Folland et al. 2001).

The latter three lines of evidence use the shorter time period because that is the period over which more reliable observations are available and on which most detection studies of these lines of evidence have been carried out (Santer et al. 1996b; Mitchell et al. 2001). For the detection and attribution exercise each line of evidence is quantified using a single metric. The vertical pattern of temperature is quantified as the

temperature change at 50 hPa. This choice was made because the stratospheric component is the aspect of the vertical pattern that discriminates it from the surface temperature change, and because we found that experts had difficulty operating on the difference of temperature between the stratosphere and lower troposphere. The geographical pattern was quantified as the temperature difference between land and ocean areas (Karoly and Braganza 2001). This choice does not capture all features of the geographical pattern, but it is one of its more salient components. These lines of evidence are fairly distinct (as judged by expert assessments of the dependence among them) and rely on a range of underlying physical processes in response to greenhouse and other climate forcing.

Most experts are comfortable with the selection of M, V, G, and D as core lines of evidence. While there is a high level of agreement among the experts on the appropriate lines of evidence to use to detect climate change and differentiate greenhouse forcing, the selection of lines of evidence is still a potential source of bias. That is because lines of evidence were chosen partly on the basis that they would be useful for finding evidence of greenhouse climate change. Experts tend to select lines of evidence that are already in the literature, and such lines of evidence become established in the literature because they are considered indicative of greenhouse forcing. Were lines of evidence chosen more generically to differentiate nonspecified sources of climate change, a different set might have been selected and results may be less indicative of a greenhouse forcing signal.

Some experts did add other lines of evidence, but the additional lines are rarely ranked high in terms of utility for detection and attribution of greenhouse climate change. The global mean, M, is generally ranked first on the basis that it provides a good fit to a priori estimates of greenhouse and aerosol-induced climate change for a relatively long data record. It is often coupled with the millennial timescale proxy temperature record (Mann et al. 1998; Briffa and Osborn 1999) on the basis that the latter highlights the uniqueness of the trend in the century-scale record. The vertical pattern, V, is generally ranked second on the basis that it provides some ability to discriminate between greenhouse and solar forcing at upper levels where the expected signal direction is opposed. The importance of V is diminished on the basis that the data record is short (less than 40 yr), and the stratospheric temperature signal is strongly confounded by changes in ozone concentration over the period. Some point out that this confound is less important in the Tropics, where ozone-induced cooling is likely to be less important than that due to greenhouse gases. While many experts note that the geographical pattern, G, could be an important discriminator of greenhouse climate change, it is usually ranked below V. This is because the observational record for which broadscale global coverage exists is relatively short, there is an overlap of data with the global mean temperature, and there are few model runs exploring the geographical pattern response to a wide range of different forcing combinations. The diurnal cycle, D, generally ranks last among the primary lines of evidence. This is because there is no clear expectation about what signal to expect in response to greenhouse or other forcing. Virtually all experts cite the critical role of cloudiness in controlling changes in D, and point to the fact that little is known about controls on long-term (low frequency) cloudiness changes.

Among the additional lines of evidence cited by only a few experts, no single selection dominates. Some cite changes in the annual cycle over the last century, which shows more warming in winter than summer (Karoly and Braganza 2001). This may help discriminate between greenhouse and solar forcings, since greenhouse forcing tends to reduce the annual cycle, whereas solar forcing is concentrated in summer and would act to enhance it. Some experts suggested that extreme events may eventually provide opportunity for increasing signal-to-noise ratios for detection, but that the existing observational records are too short (Wigley 1999). Examples cited included increases in precipitation intensity in the upper decile of the precipitation distribution (Karl and Knight 1998). Another line of evidence cited is the concentration of warming in dry, cold anticyclones in northern winter (Michaels et al. 2000), which act as a marker of greenhouse warming in regions with reduced water vapor feedback.

Sources and assessments of natural variability. Experts tended to cite similar processes when asked to account for the major sources of natural internal variability in each line of evidence. For example, the major sources of variability cited for M were: upper-ocean air–sea exchanges (e.g., ENSO); deep-ocean circulation including the thermohaline circulation; large-scale oscillations (natural modes) in the atmosphere [e.g., the North Atlantic Oscillation (NAO)]; land surface processes and changes (soil moisture, vegetation, land use, hydrology); and snow and sea ice cover changes. Experts were also asked to assess confidence in understanding and modeling of each of these processes

as they relate to characterization of each line of evidence. Responses were made on a qualitative five-point scale following Moss and Schneider's (2000) characterization from "very high confidence" to "very low confidence." There was little agreement among experts, with responses (for each process) ranging from high to low confidence on assessments of understanding and the ability to model these processes. In some cases experts noted that while understanding of processes was reasonable, the appropriate mechanisms had not yet been incorporated in climate models, thereby diminishing confidence in the model simulations.

For each line of evidence, respondents were asked to make judgements about whether the natural variability was typically underestimated in models (less than half the true natural variability), about right in models (between half and one and a half times the true natural variability), or overestimated in models (greater than one and a half times the true natural variability). For the global mean, M, little probability mass was given to possible model overestimates of natural variability. Experts typically assigned about seven-tenths of the probability mass to "about right" and the balance to "model underestimated." The exceptions were one expert who reversed this allocation, and one who assigned almost all the probability mass to about right. For the vertical pattern, V, experts were less consistent with one another in assigning probability masses, and tended to increase the allocation of probability mass to the model underestimate category. Some experts placed the majority of the probability mass in this category, citing the fact that climate models have truncated stratospheres (reduced in vertical extent and capped by a lid) with coarse resolution. These experts had low confidence in models simulating sufficient variability near the lid, and cited the failure to generate a quasi-biennial oscillation (QBO) in the models. For the geographical pattern, G, most experts tended to allocate nearly equal probability masses to models underestimating and being about right. In general, the responses indicate that model estimates of natural variability would be biased on the underestimate side of about right (consistent with Barnett 1999). Depending on the expert, this bias may be moderate, or quite severe. The more severe this bias, the more the detection studies would overestimate levels of detection when drawing estimates of natural variability from climate model simulations.

Detection. The probability distributions characterizing century-long changes in global mean surface temperature and natural variability for each expert are shown in Fig. 1a. Experts are in good agreement on the size of the century-long trend, with the smallest 95% confidence band expressed encompassing the mean of almost all the other experts. The agreement on the expected value of the trend in internal natural variability over the period is not meaningful as that would be close to zero by definition—the expected value of any large ensemble of 100-yr periods should tend toward zero in the absence of external forcing.

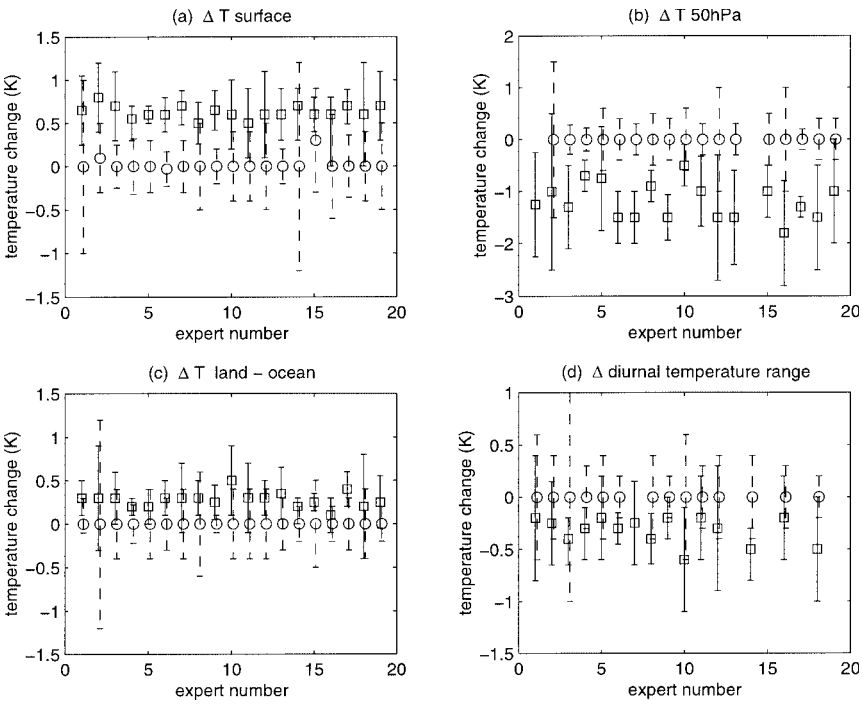

Fig. 1. The 95% confidence bands for each expert on the trend and natural variability in (a) global mean surface temperature, (b) vertical temperature pattern (50 hPa), (c) geographical pattern, and (d) diurnal temperature range. The solid lines refer to the trend and the dashed lines refer to natural variability. Some distributions are missing because the expert felt the relevant available data was too sparse to allow a reasonable assessment of the distribution.

The variation in expert assessment of the size of the 95% confidence band in natural variability spans a factor of 5.

The probability of detection for the global mean line of evidence M [denoted $P(D_M)$] calculated from these distributions [$f(E_M)$ and $f(N_M)$] is given in Table 2 for the 5% level of significance. Detection probabilities were also calculated for the 1%, 10%, and 20% significance levels (not shown). Results are somewhat sensitive to the significance level selected, with detections increasing slightly as the level relaxes, as would be expected. Bearing that in mind, we discuss results for the 5% significance level. For the global mean, M, a majority of experts reject the null hypothesis of no climate change, thereby suggesting a near consensus among experts for detection of climate change as measured by the change in global mean surface air temperature. The median probability of detection for experts is high (~0.95) and the spread across experts is relatively low. For the vertical pattern, V, (Fig. 1b) detection probabilities are generally higher (than for M) with even less spread across experts, though two experts did not provide data for this line of evidence, citing low confidence in existing observations and model results.

Table 2. Probability of detection for each line of evidence [$P(D_i)$]. In cases where the null hypothesis is rejected (detection), the probability of detection is given. Otherwise, acceptance of the null hypothesis is indicated by "ϕ." A "—" indicates that the expert declined to provide an assessment. For the diurnal cycle (D) this was typically because the expert did not feel sufficiently well acquainted with the data and/or did not have clear expectations about what signal to expect.

Expert	$P(D_M)$	$P(D_V)$	$P(D_G)$	$P(D_D)$
1	ϕ	—	ϕ	ϕ
2	0.97	ϕ	ϕ	ϕ
3	0.99	0.99	ϕ	ϕ
4	0.71	0.99	0.81	0.71
5	0.99	0.70	ϕ	ϕ
6	0.99	0.99	0.70	0.95
7	0.99	0.99	ϕ	—
8	0.77	0.99	ϕ	0.72
9	0.99	0.99	0.95	0.64
10	0.91	ϕ	0.80	ϕ
11	0.80	0.99	ϕ	ϕ
12	0.78	0.87	ϕ	ϕ
13	0.99	0.99	0.75	—
14	ϕ	—	ϕ	0.89
15	ϕ	0.99	ϕ	—
16	0.64	0.98	ϕ	ϕ
17	0.98	0.99	0.94	—
18	0.82	0.99	ϕ	0.75
19	0.93	0.91	0.72	—
median	0.95	0.99	0.80	0.73
c.v. $\left(\frac{\sigma}{\mu}\right)$	0.13	0.08	0.12	0.15

Results for the geographical pattern, G, and diurnal cycle, D, show far less support for detection of climate change based on these lines of evidence. For G, less than half the experts reject the null hypothesis (detection) at 5% and 10% levels of significance. For the diurnal cycle there is even less support for detection. Five of the experts declined to provide an assessment on the diurnal cycle. The rest are almost evenly split between accepting the null hypothesis and rejecting it. For the experts who do reject the null hypothesis, the probabilities of detection are lower than for M and V. For the geographical pattern and diurnal cycle lines of evidence, expert judgements tend to reflect both a lack of confidence in the available data and uncertainty on their role in detection of climate change.

Attribution. The first part of the attribution process entailed an assessment of uncertainty in the key forcings for each line of evidence. The pdfs representing forcing uncertainty are discretized and then attributions are made for each of the combinations of discrete forcing cases. To keep the number of such combinations manageable, solar and aerosol forcing pdfs are represented by a "low" and "high" range only. In the case of the global mean temperature line of evidence, M, solar forcing low and high are defined by less or more than 1 W m^{-2} over the century and aerosol low and high are defined by less or more negative forcing than 1.5 W m^{-2}. Experts produced a wide range of estimates of the probability mass in each category for each of the forcings. There is more expert agreement in the case of solar forcing, with only five experts allocating more than 30% of the probability mass to solar forcing > 1 W m^{-2}, and all experts allocating 50% or more probability mass to solar forcing < 1 W m^{-2}. For aerosol forcing, all of the experts allo-

cate 50% or more of the probability mass to aerosol forcing < 1.5 W m^{-2}, but fewer than half of the experts allocate more than 70% of the probability mass to this range. The low ranges of aerosol and solar forcing above encompass most of the uncertainty range for estimates of these forcings in (Shine and Forster 1999) and (Ramaswamy et al. 2001)—that is, the so-called low ranges here correspond to the more conventional views of the forcings. The considerable spread of probability mass across the low and high forcing ranges for most experts is consistent with the "low confidence" afforded aerosol and solar forcing estimates in (Shine and Forster 1999) and (Ramaswamy et al. 2001).

The fraction of the signal attributed to greenhouse forcing, $Q_{i,\mathrm{gh}}$, for each line of evidence, *i,* is shown in Table 3. The fraction $1-Q_{i,\mathrm{gh}}$ is the fraction attributed to all other forcings. For both the global mean, M, and geographical pattern, G, greenhouse forcing explains more than half of the signal for most experts. For the vertical pattern, V, and diurnal temperature range, D, the fraction of responsibility attributed to greenhouse forcings is much lower and the spread across experts is higher. The median fraction of greenhouse responsibility ascribed for V is only 0.2, reflecting the belief that changes in ozone are more directly responsible for recent stratospheric temperature changes.

The fractional attributions to greenhouse gases reported in Table 3 represent averages over the uncertainty in the relevant forcings. For example, the attribution of greenhouse forcing to global mean temperature, M, is a probability weighted average of the attributions for low and high forcing combinations of solar and aerosol forcing (described above). The effect of forcing uncertainty on attribution can be determined by placing the entire probability mass in either the low or high forcing case (i.e., assuming one of these cases is true) and comparing the resulting attribution values to those in Table 3. The results of such an analysis for M are discussed below.

The effect of solar forcing uncertainty on global mean temperature, M, is consistent across all experts. Fractional attribution values assessed for the low solar forcing assumption are within a few percent of the weighted-average values in Table 3 for most experts. For the high solar forcing case, most experts increased the fractional attribution to solar forcing and decreased the attribution to greenhouse forcing, typically by about 10%. The results for aerosol forcing uncertainty are more subtle and depend on whether the expert considers attribution to be more sensitive to uncertainty in forcing or to climate sensitivity. In the former case, higher aerosol forcing (negative) must be balanced by greater positive forcing. This tends to increase attribution to nongreenhouse forcings because they are typically less well known than greenhouse forcing, and many believe they are more plausible candidates for adjusting upward to balance the increased negative forcing. If uncertainty in climate sensitivity is considered more important than uncertainty in forcing, then higher negative forcing does not need to be balanced by higher positive forcing to explain the observed trend. In this case the trend can be explained by lower net positive forcing and higher climate sensitivity. Experts making this argument typically did not adjust attributions for the high aerosol case. In either case the effect of aerosol

Table 3. Expected fraction attributed to greenhouse forcing for each line of evidence ($Q_{i,\mathrm{gh}}$). The $Q_{i,j} = \Sigma_k Q_{i,j,k} P_k$, where $P_k = \prod_{j=1}^{J} P_{j,k}$, and $P_{j,k}$ are the probability masses assigned to each of the *k* discrete forcing cases.

Expert	$Q_{M,\mathrm{gh}}$	$Q_{V,\mathrm{gh}}$	$Q_{G,\mathrm{gh}}$	$Q_{D,\mathrm{gh}}$
1	0.84	0.04	0.77	0.78
2	0.75	0.15	0.77	0.23
3	0.75	0.25	0.72	0.50
4	0.63	0.20	0.44	0.40
5	0.71	0.40	0.71	0.78
6	0.70	0.20	0.70	0.39
7	0.72	0.20	0.61	—
8	0.65	0.20	0.56	0.30
9	0.82	0.25	0.60	0.30
10	0.75	0.20	0.70	0.60
11	0.60	—	0.60	0.30
12	0.61	0.15	0.47	0.59
13	0.61	0.45	0.50	—
14	0.50	—	0.50	0.80
15	0.80	0.20	0.80	—
16	0.40	0.40	0.42	0.60
17	0.74	0.20	0.75	—
18	0.77	0.50	0.49	0.30
19	0.74	0.50	0.77	—
median	0.72	0.20	0.61	0.45
c.v. $\left(\frac{\sigma}{\mu}\right)$	0.16	0.51	0.20	0.41

forcing uncertainty on greenhouse attributions is modest. For these experts then, the attributions set by weighting the fractional attributions by the probability masses assigned to each forcing range (as in Table 3) are not substantially different from those that would be obtained if the high or low forcing cases alone were considered.

Overall attributions to each cause take into account the role of natural variability as well as climate forcings. Figure 2 provides a bar plot summary of the expected contributions of greenhouse forcing, all other forcing, and natural variability to the change in each line of evidence. It shows the consistent (across experts) high attributions to greenhouse forcing for M (and to a lesser extent, G), the dominant role of "all other forcings" in explaining changes in V, and the large spread of expert assessments of the greenhouse forcing contribution to changes in D. The figure also indicates a typically small role for contributions of natural variability to explain changes in M and V, with more substantial contributions across experts for G and D. There are some exceptions to this pattern in each case. While most experts attribute nearly all the responsibility for changes in M (and V) to forcing, a few of the experts apportion larger roles for natural variability. These experts allow broader distributions for natural variability (see Fig. 1). For both G and D there are substantial contributions from each of greenhouse forcing, all other forcing, and natural variability across experts such that no one cause dominates.

After levels of attribution have been determined for each line of evidence, it is desirable to assess the implications of that for a more general picture of climate change than afforded by each line of evidence alone. One such picture of climate change is obtained by considering the set of lines of evidence.[2] These can be aggregated in a variety of different ways. One way is to set thresholds for the overall attributions to each potential cause. For example, a threshold of 0.5 would imply that the cause must be responsible for half or more of the change in each given line of evidence. With that threshold, inspection of Fig. 2 indicates that greenhouse forcing exceeds the threshold for *some* lines of evidence for *some* experts. For the global mean surface temperature line of evidence the threshold is exceeded for all but three experts, while for the vertical temperature pattern it is not satisfied for any of the experts (see Table 4). The level at which such thresholds are set can be modified to match the level of responsibility deemed important, which is in part a value judgement (Schneider 1989). Attribution studies can evaluate the potential contributions from different causes, but they cannot set the burden of responsibility required, since that entails broader social judgements.

CONCLUSIONS. Among a varied sample of 19 experts engaged in studies of detection and attribution of climate change, there are a number of issues on which there is general agreement and a number on which there is a broad spread of expert judgement. Caution must be used in interpreting these results, since expert agreement does not necessarily imply low uncertainty (all experts may be wrong), just as expert disagreement does not necessarily imply high uncertainty (some experts may simply know better than others). Nonetheless, on a contentious issue like de-

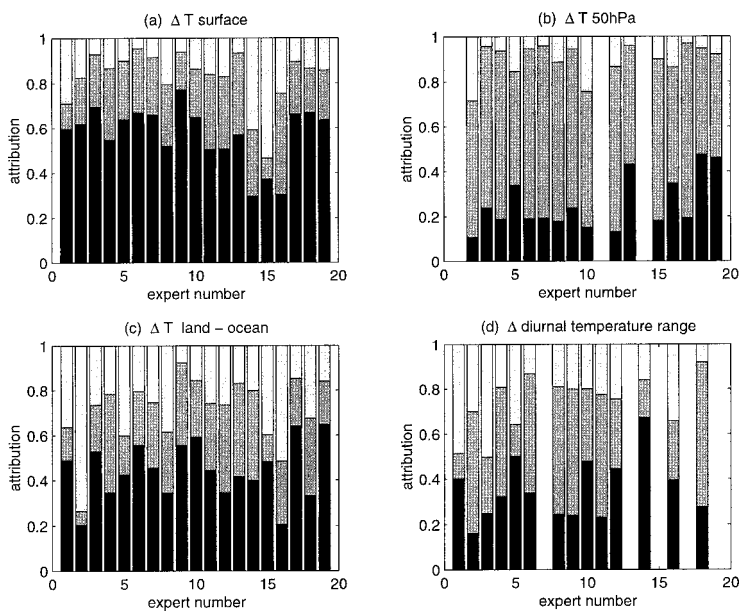

Fig. 2. Responsibility attributed to greenhouse forcing (black segment) [$A_{i,gh}$], all other forcing (medium gray segment) [$A_{i,oth}$], and natural variability (light gray segment) [$A_{i,nv}$] in explaining the change in (a) global mean surface temperature, (b) vertical temperature pattern (50 hPa), (c) geographical pattern, and (d) diurnal temperature range for each expert. Note that $A_{i,gh} = (1 - A_{i,nv})Q_{i,gh}$, where $A_{i,nv} = m/n\Sigma_{I N>0 E>0} N_I/E_I$, where I represents random draws from $f(N_i)$ and $f(E_i)$, and m and n are the number of draws where N and E are greater than 0, respectively, for positive trends.

[2] The process of combining lines of evidence is implicit in statements such as the "balance of evidence suggests . . ." (Santer et al. 1996b).

tection and attribution, it is useful to know where experts agree or disagree and why, along with some indication of whether the disagreements are consequential or not.

Expert agreement is strongest on some aspects of the detection phase of the detection and attribution problem. The characterization of natural variability is the largest source of uncertainty at the detection step, and there is a general belief that it has been underestimated in model-based detection studies. Experts are in agreement on selection of the key lines of evidence that signify climate change, and there is broad agreement that changes in the global mean (M) and vertical pattern (V) are unlikely to be the result of natural variability alone. For the geographical pattern and diurnal cycle lines of evidence, there is much more disagreement across experts.

Expert assessments of uncertainty in key forcings of the different lines of evidence vary quite a bit. For the global mean temperature (M) line of evidence, the main forcing uncertainties assessed were aerosol and solar forcing. One consistency that emerges is that most experts allocate substantial probability mass to values of the forcings outside the conventional uncertainty ranges for these forcings. The assessed effect of uncertainty in these forcings on greenhouse attribution values is fairly modest, however.

In attributing causes to the observed changes, experts allocate the lion's share (typically about 60%) of the responsibility to greenhouse forcing for the global mean surface temperature. Solar and all other forcings and natural variability together account for the minority of responsibility for this line of evidence. For both the vertical pattern of temperature change and the change in diurnal temperature range the responsibility allocated to greenhouse forcing is typically much lower. However, for these lines of evidence the spread across experts is higher and several experts declined to make assessments. In the case of the vertical pattern, the low greenhouse forcing attributions are largely due to ozone forcing in the stratosphere. While these results suggest that the vertical pattern is not very useful for discriminating a greenhouse signal at present, it may play a bigger role in the future as the confounding effects of ozone forcing in the stratosphere gradually diminish. For the diurnal temperature range, experts point to higher uncertainty in the expected response to greenhouse forcing, citing key unknowns related to changes in cloudiness. This high uncertainty may allow for a greater role for the diurnal temperature in discriminating among different forcings as more is known about the processes governing its response.

Table 4. Cases where the fractional contribution of greenhouse forcing to changes in each given line of evidence exceeds an arbitrary threshold of 0.5—indicated by "◊."

Expert	M $\Delta T_{surface}$	V ΔT_{50hPa}	G $\Delta T_{land\text{-}ocean}$	D Δ DTR
1	◊	—	—	—
2	◊	—	—	—
3	◊	—	◊	—
4	◊	—	—	—
5	◊	—	—	◊
6	◊	—	◊	—
7	◊	—	—	—
8	◊	—	—	—
9	◊	—	◊	—
10	◊	—	◊	—
11	◊	—	—	—
12	◊	—	—	—
13	◊	—	—	—
14	—	—	—	◊
15	—	—	—	—
16	—	—	—	—
17	◊	—	◊	—
18	◊	—	—	—
19	◊	—	◊	—

Some of the details of the results reported here will undoubtedly change as the evidence and science underlying detection and attribution progresses. Expert levels of confidence will change, new lines of evidence will be added, and some will fall from favor. However, it is interesting to note that as of circa 2000, there is a high level of confidence in detection of climate change based on changes in global mean temperature and in the vertical pattern of temperature change. There is also general agreement that points to a substantial role for greenhouse forcing in contributing to global mean surface temperature changes, and scattered support for a greenhouse role for the other lines of evidence examined. The change in global mean surface air temperature appears to be the single most important line of evidence in the detection of climate change and attribution to anthropogenic causes.

ACKNOWLEDGMENTS. This work was funded by the NOAA Climate and Global Change Program. We are grateful to the set of experts listed in Table 1 for their participation and feedback. Thanks are also due to Peter Reichert, Peter Reinelt, Urmila Diwekar, Benoit Morel, and Karl Braganza.

REFERENCES

Barnett, T., 1999: Comparison of near-surface air temperature variability in 11 coupled global climate models. *J. Climate,* **12,** 511–518.

——, and Coauthors, 1999: Detection and attribution of recent climate change: A status report. *Bull. Amer. Meteor. Soc.,* **80,** 2631–2659.

Bell, T. L., 1986: Theory of optimal weighting of data to detect climate change. *J. Atmos. Sci.,* **43,** 1694–1710.

Briffa, K. R., and T. J. Osborn, 1999: Seeing the wood from the trees. *Science,* **284,** 926–927.

Folland, C., and Coauthors, 2001: Observed climate variability and change. *Climate Change 2001: The Scientific Basis,* J. T. Houghton et al., Eds., Cambridge University Press, 99–181.

Hansen, J., and A. Lacis, 1990: Sun and dust versus greenhouse gases: An assessment of their relative roles in global climate change. *Nature,* **346,** 713–719.

——, M. Sato, A. Lacis, R. Ruedy, I. Tegen, and E. Matthews, 1998: Perspective: Climate forcings in the industrial era. *Proc. Natl. Acad. Sci.,* **95,** 12 753–12 758.

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

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

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

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

——, and Coauthors, 1993: Asymmetric trends of daily maximum and minimum temperature. *Bull. Amer. Meteor. Soc.,* **74,** 1007–1023.

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

Kasting, J., 1998: The carbon cycle, climate, and the long-term effects of fossil fuel burning. *Consequences,* **4,** 15–27.

Mann, M. E., R. S. Bradley, and M. K. Hughes, 1998: Global-scale temperature patterns and climate forcing over the past six centuries. *Nature,* **392,** 779–787.

Michaels, P. J., P. C. Knappenberger, R. C. Balling Jr., and R. E. Davis, 2000: Observed warming in cold anticyclones. *Climate Res.,* **14,** 1–6.

Mitchell, J., and Coauthors, 2001: Detection of climate change and attribution of causes. *Climate Change 2001: The Scientific Basis,* J. T. Houghton et al., Eds., Cambridge University Press, 695–738.

Moss, R., and S. H. Schneider, 2000: Towards consistent assessment and reporting of uncertainties in the IPCC TAR: Initial recommendations for discussion by authors. Intergovernmental Panel on Climate Change Tech. Report, IPCC Guidance Paper, 35 pp.

Pelletier, J., 1997: Analysis and modeling of the natural variability of climate. *J. Climate,* **10,** 1331–1342.

Ramaswamy, V., and Coauthors, 2001: Radiative forcing of climate change. *Climate Change 2001: The Scientific Basis,* J. T. Houghton et al., Eds., Cambridge University Press, 349–416.

Risbey, J. S., M. Kandlikar, and D. J. Karoly, 2000: A protocol to articulate and quantify uncertainties in climate change detection and attribution. *Climate Res.,* **16,** 61–78.

Santer, B., and Coauthors, 1996a: A search for human influences on the thermal structure of the atmosphere. *Nature,* **382,** 39–46.

——, T. M. L. Wigley, T. Barnett, and E. Anyamba, 1996b: Detection of climate change and attribution of causes. *Climate Change 1995: The Science of Climate Change,* J. T. Houghton et al., Eds., Cambridge University Press, 407–443.

Schneider, S. H., 1989: The greenhouse effect—Science and policy. *Science,* **243,** 771–781.

Shine, K., and P. de F. Forster, 1999: The effect of human activity on radiative forcing of climate change: A review of recent developments. *Global Planet. Change,* **20,** 205–225.

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

Wigley, T. M. L., 1999: The science of climate change: Global and U.S. perspectives. Tech. Report, Pew Center on Global Climate Change, Arlington, VA, 48 pp.

Zwiers, F. W., 1999: Climate Change Detection: A review of Techniques and Applications. *Anthropogenic Climate Change. Proceedings of the First GKSS Spring School on Environmental Research,* H. Storch, E. Raschke, and G. Flöser, Eds., Springer Verlag, 161–203.