Origins of Model–Data Discrepancies in Optimal Fingerprinting

GABRIELE C. HEGERL

Department of Earth and Ocean Sciences, Duke University, Durham, North Carolina

MYLES R. ALLEN

Space Science and Technology Department, Rutherford Appleton Laboratory, Chilton, Didcot, Oxon, United Kingdom

(Manuscript received 26 February 2001, in final form 11 December 2001)

ABSTRACT

Two approaches to distinguishing anthropogenic greenhouse gas and sulfate aerosol signals in the observed surface temperature record are compared. Both rely on a variant of general regression called ''optimal fingerprinting.'' One approach is equivalent to a stepwise regression procedure estimating, first, a greenhouse gas signal and, in a second step, the sulfate aerosol signal. This is different from multiple regression, under which both signals are estimated simultaneously and treated symmetrically. The stepwise regression approach is a more powerful means of detecting greenhouse gas influence in the presence of a small and possibly poorly simulated sulfate aerosol signal. However, when both signals are of comparable size, multiple regression provides estimates of the amplitude of the greenhouse and sulfate responses that are, in principle, independent of each other, making it generally simpler to interpret. It is shown that there is a simple linear transform relating the stepwise and multiple regression approaches. Application of this transform to previous results of stepwise regression illustrates that estimated responses to anthropogenic greenhouse gas forcing are very similar between different climate models and are generally consistent with the signal estimated from the observations. The sulfate component of the anthropogenic signal appears to be responsible for the most prominent discrepancies between observations and some of the model simulations considered. The estimated contribution of anthropogenic greenhouse gases to the observed warming over the period of 1949–98 lies in the range of 0.39–1.29 K (50 yr)$^{-1}$ or 0.28–1.16 K (50 yr)$^{-1}$ (5%–95% range), depending on the model used to estimate the signal. These ranges depend only on the accuracy of the spatial pattern and the sign of the modeled sulfate forcing and response, not on its amplitude.

1. Estimating several anthropogenic signals in observations

The detection and attribution of anthropogenic climate change attempts to distinguish anthropogenic contributions to the observed climate record from those due to internal climate variability and natural external forcings such as changes in solar irradiance. To distinguish between different forcings, methods are applied that are based on multiple regression (Hasselmann 1979, 1997; Allen and Tett 1999). The observations $\mathbf{y}$ are generally composed of several (n) climate change ''signals,'' $\mathbf{g}_i$, such that

$$\mathbf{y} = \sum_{i=1}^{n} b_i \mathbf{g}_i + \text{noise}, \qquad (1)$$

where the b_i are unknown ''pattern amplitudes'' to be estimated from the data. In the present paper, the observations $\mathbf{y}$ are global patterns of surface temperature trends.

Several differently phrased but philosophically identical variants of this method have been used in publications (Hegerl et al. 1997; Allen and Tett 1999; Tett et al. 1999; Allen et al. 2001; North and Stevens 1998). Since different signal patterns, such as those corresponding to greenhouse gas and sulfate aerosol forcing, may be highly correlated with each other, Hegerl et al. (1997) ''orthogonalized'' the sulfate aerosol signal relative to the greenhouse gas signal prior to the regression. This is equivalent to a stepwise regression procedure, in which first a greenhouse gas signal is estimated and then a sulfate aerosol signal is estimated from the residual. Such an approach is more powerful for the detection of a ''greenhouse gas–type'' signal from observations, but the amplitude estimate will be influenced by aspects of the sulfate aerosol signal that are colinear with the greenhouse gas signal (in practice, the estimated greenhouse gas signal amplitude will be reduced by aerosol-induced global cooling). The orthogonal sulfate fingerprint represents a signal that is unique to the sulfate aerosol forcing and distinguishes whether such

Corresponding author address: Gabriele Hegerl, Dept. of Earth and Ocean Sciences, Duke University, Durham, NC 27708.
E-mail: hegerl@duke.edu

a signal can also be detected. For surface temperature trends, for example, the orthogonal sulfate fingerprint shows the Northern Hemisphere midlatitudes (where sulfate aerosol forcing is stronger) cooling relative to the rest of the globe. Similarly, North and Stevens (1998) attempted to detect climate change signals that were orthogonal to other signals, for example, a greenhouse gas signal that was orthogonal to the combined response to solar, volcanic, and aerosol forcing.

In contrast, Leroy (1998), Allen and Tett (1999), Tett et al. (1999), Stott et al. (2001), and Allen et al. (2001) apply the multiple regression model in (1) to the ''raw'' model-simulated signals, accounting for correlations between signals directly in the regression without prior orthogonalization. The advantage is that all signals are treated symmetrically and that each estimated signal amplitude is, in principle, independent of the amplitude of the other ''confounding'' signal(s). We stress ''in principle'' here, because this independence only holds if the statistical model used is actually correct, that is, there are no errors in the simulated pattern of response to either forcing. In practice, with imperfect model-simulated signals, some cross contamination will always occur, but the multiple regression approach minimizes its effects. Because of this independence, the physical interpretation of the result is more straightforward. Therefore, full multiple regression is preferable for deriving meaningful estimates of anthropogenic signals. We stress, however, that both approaches are strongly related and yield the same result in attribution tests that test the consistency between observed and model-simulated signal amplitudes. The choice of stepwise or multiple regression depends on the nature of the problem and the expected size and reliability of signals relative to each other. Also, beyond the difference between stepwise and multiple regression, all optimal detection approaches are mathematically identical, differences being only a matter of interpretation (Hegerl and North 1997; Allen and Tett 1999).

The purpose of this paper is to explain how independent estimates of the amplitude of climate signals (particularly, an estimate of the amplitude of the greenhouse gas signal that is independent from the sulfate aerosol signal) can be extracted from the results of the detection and attribution of climate change as implemented by Hegerl et al. (1997, 2000). This is done by a transform of the results from the stepwise regression to a multiple regression. The so-derived signal amplitudes simplify the comparison of results from Hegerl et al. (1997) with the other studies based on multiple regression. Also, an application of the stepwise regression approach to several model simulations shown in Barnett et al. (1999) detected inconsistencies between some model simulations and observations. The transform of these results to independent amplitudes of a greenhouse gas and sulfate aerosol signal helps to interpret these discrepancies, as will be demonstrated below.

In the present paper, the specific case of estimating two independent climate change signals, one for greenhouse gas–only forcing and one for anthropogenic sulfate aerosol forcing, is considered. The extension of the present analysis to more than two signals is straightforward. To facilitate comparison with the results of Hegerl et al. (1997, 2000), signals from natural external climate influences such as volcanism and solar irradiance changes not directly considered. Tett et al. (1999) and Allen et al. (2001) estimate that the net influence of natural forcing on trends over the past 50 years is likely to be small. However, incorporating natural forcing estimates into the anthropogenic estimates will increase uncertainty ranges (Allen et al. 2001).

The paper is structured as follows: In the second section, the stepwise regression approach and its results are outlined. The third section introduces multiple regression and a transform from stepwise regression to multiple regression. An application of that transform to earlier detection results in the fourth section demonstrates the differences between both approaches. Conclusions are drawn in the final section.

2. Stepwise regression

In Hegerl et al. (1997, 2000), the amplitudes of the climate response to anthropogenic greenhouse gas forcing and anthropogenic sulfate aerosol forcing are estimated from a set of observations $\mathbf{y}$ using a stepwise linear regression. First, the amplitude of the response to greenhouse gases $\mathbf{g}_G$ is estimated using the single-pattern variant of optimal fingerprinting (Hasselmann 1993); thus,

$$a_G = \mathbf{g}_G \cdot \mathbf{y} \equiv \mathbf{g}_G^{\mathrm{T}} \mathbf{C}^{-1} \mathbf{y}. \tag{2}$$

Throughout this paper, the scalar product is defined in terms of the inverse noise covariance matrix $\mathbf{C}^{-1}$, yielding the lowest-variance, linear, unbiased estimator of the signal amplitude (Hasselmann 1979, 1993). An equivalent way of expressing this is to assume that both observations and signals are first transformed to a coordinate system in which equal and uncorrelated levels of climate noise are expected in all components of $\mathbf{g}_G$ and $\mathbf{y}$, and then the conventional scalar product is used in (2)—see Allen and Tett (1999). Note that a_G in (2) is an unbiased estimator of b_G in (1) if and only if $\mathbf{g}_G$ is normalized under the metric defined by $\mathbf{C}^{-1}$, or $\mathbf{g}_G \cdot \mathbf{g}_G = 1$, and all other signals contributing to $\mathbf{y}$ are orthogonal to $\mathbf{g}_G$ under this metric (Hasselmann 1997). The estimation of the inverse noise covariance matrix $\mathbf{C}^{-1}$ is nontrivial. In general, the estimate is based on samples of climate variability from ''control simulations'' of climate models after reducing the dimension of the problem to a tractable size (Hegerl et al. 1996; Allen and Tett 1999).

It is then established whether a second response pattern, in this case the response to sulfate aerosol forcing, can be detected in addition to the greenhouse gas response. In Hegerl et al. (1997), this is achieved by pro-

jecting the residuals of regression, $\mathbf{y} - a_G\mathbf{g}_G$, onto a second signal $\mathbf{g}_{GS}$, which includes the influence of both greenhouse gas and sulfate aerosol forcing. These residuals are, by construction, orthogonal to the greenhouse-only pattern under the metric defined by $\mathbf{C}^{-1}$; that is,

$$\mathbf{g}_G \cdot [\mathbf{y} - (\mathbf{g}_G \cdot \mathbf{y})\mathbf{g}_G] \equiv 0. \tag{3}$$

Hence, it does not matter that the $\mathbf{g}_{GS}$ signal also contains a contribution due to greenhouse gases. An equivalent way of expressing this is that the original observations $\mathbf{y}$ are projected onto a signal $\mathbf{g}^*_{GS}$ that has been orthogonalized relative to the greenhouse gas signal $\mathbf{g}_G$ as follows:

$$\mathbf{g}^*_{GS} = s^{-1}(\mathbf{g}_{GS} - r\mathbf{g}_G), \tag{4}$$

where r is the correlation coefficient between $\mathbf{g}_G$ and $\mathbf{g}_{GS}$, s is a normalization factor

$$s = \sqrt{1 - r^2}, \tag{5}$$

and all signal patterns are normalized thus:

$$\mathbf{g}_G \cdot \mathbf{g}_{GS} = r, \tag{6}$$

$$\mathbf{g}_G \cdot \mathbf{g}_G = 1, \quad \text{and} \tag{7}$$

$$\mathbf{g}_{GS} \cdot \mathbf{g}_{GS} = 1, \tag{8}$$

and hence

$$\mathbf{g}^*_{GS} \cdot \mathbf{g}^*_{GS} = 1 \quad \text{and} \quad \mathbf{g}^*_{GS} \cdot \mathbf{g}_G = 0. \tag{9}$$

Thus, the combined estimate can be written

$$\begin{pmatrix} a_G \\ a^*_{GS} \end{pmatrix} = s^{-1}\begin{pmatrix} s & 0 \\ -r & 1 \end{pmatrix}\begin{pmatrix} \mathbf{g}^T_G \\ \mathbf{g}^T_{GS} \end{pmatrix} \cdot \mathbf{y}, \tag{10}$$

with a_G and a^*_{GS} denoting regression coefficients as in (1), except that the signals applied are orthogonal to each other. Note that the amplitude of the greenhouse gas signal a_G reflects the amplitude of both a greenhouse gas signal and aspects of the sulfate aerosol signal that are colinear with it [since the climate response to sulfate aerosol–only forcing, which is different from $\mathbf{g}^*_{GS}$, involves large-scale cooling that projects negatively onto $\mathbf{g}_G$, e.g., Hegerl et al. (1997)]. In contrast, a^*_{GS} is by construction not influenced by the greenhouse gas forcing and reflects a pattern that is purely in response to sulfate aerosol forcing.

We illustrate this approach using greenhouse gas and orthogonalized sulfate aerosol fingerprints from spatial patterns of 50-yr trends in Northern Hemispheric summer [June–July–August (JJA)] near-surface temperatures. In Hegerl et al. (1997), fingerprints were derived from simulations with the coupled climate model ECHAM3/LSG (version 3 of the ECHAM atmosphere model coupled to the Large-Scale Geostrophic ocean model). This model is denoted ''model 2'' for consistency with Barnett et al. (1999). This approach was subsequently extended to the Hadley Centre model HadCM2 (denoted as ''model 1'') in Hegerl et al. 2000. The results from the stepwise regression have been used to estimate the significance level of anthropogenic signals in observations using each model's signal patterns. Figure 1 reiterates the results from the Hegerl et al. (2000) analysis, using fingerprints from both models 1 and 2 to estimate the signals a_G and a^*_{GS}. In both cases, a significant greenhouse gas signal is detected (i.e., the horizontal one-dimensional uncertainty bar does not intersect zero), but only with model 2 do we detect a significant orthogonal aerosol signal. The ellipses give the joint uncertainty level of both signals. [Very attentive readers will note that if the covariance matrix $\mathbf{C}$ in (2) were perfectly estimated, the uncertainty ellipse would be the unit circle. The ellipses grow and deviate from a circle because of sampling uncertainty and additional uncertainty reflecting sampling error in observations (Hegerl et al. 2001), and because of the chunk of control simulation used to estimate the covariance matrix being different from that used to estimate uncertainty to avoid artificial skill.] In the case of model 1, the joint uncertainty around both signals encompasses 0, indicating that the case of both signals being 0 cannot be rejected although the greenhouse gas signal by itself and, even more so, the combined greenhouse gas and sulfate aerosol signal is significant (Barnett et al. 1999). This seeming discrepancy reflects the fact that uncertainties are inevitably larger if we are estimating two signals rather than one.

Figure 1 also shows the amplitudes of greenhouse gas and orthogonal sulfate patterns derived from model simulations. To derive those, model-simulated trend patterns for the same 50-yr period of 1946–95 were substituted for the observations $\mathbf{y}$[(1)]. For results using the same model as provided the fingerprints (e.g., model-1 fingerprints, Fig. 1b, and ''GS1'') this reflects where the observed signal would be expected (disregarding noise uncertainty) if the model were perfect. For the respective other model and a range of additional models, the comparison between the model amplitudes and those from observations focuses on the pattern represented by the fingerprints used in the comparison (e.g., the strength of a pattern represented by model 1's fingerprint is compared between model simulations 2–6 and the observations in Fig. 1b). The additional models were the Geophysical Fluid Dynamics Laboratory (GFDL) R30 model (model 3), two versions of the Canadian Climate Center model (model 5 and 6), and simulations with the version 4 ECHAM/Ocean isopycnal (ECHAM4/OPYC) model (model 4). Details about the simulations and references for the models are given in Barnett et al. (1999), from which the present results are derived. Note that the forcing of the model simulations, particularly those with greenhouse gas–plus–sulfate aerosol forcing, differs between simulations. In particular, model 4 has a simulation with additional indirect aerosol and tropospheric ozone forcing (''GSI''), while the model-4 greenhouse gas–plus–direct sulfate aerosol simulation (denoted by ''GS'' for all models) is forced with a small-

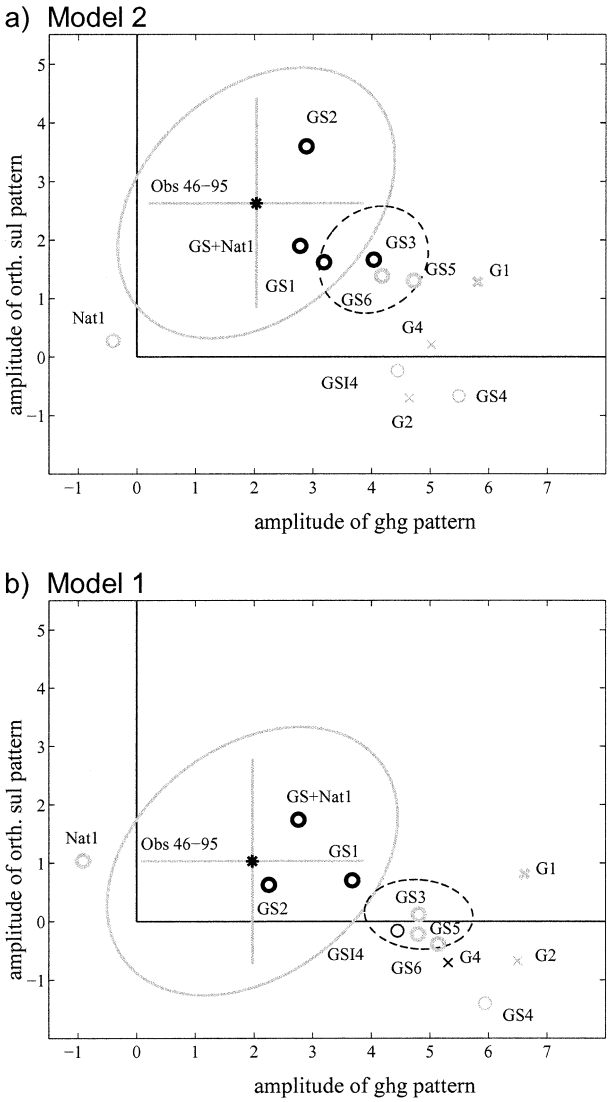

FIG. 1. Estimated amplitude of a greenhouse gas signal (horizontal axis) and an orthogonalized sulfate aerosol signal (vertical axis) in observed JJA surface temperature trends for 1945–95 for (a) ECHAM3/LSG fingerprints and (b) HadCM2 fingerprints. Note that in this stepwise regression, sulfate-induced global cooling will project on the greenhouse warming signal and reduce its amplitude. In contrast, the orthogonal sulfate pattern is unique to sulfate aerosol forcing. All units are std devs of the HadCM2 control simulation (which provides the largest estimate of internal variability from the models used), otherwise, results are identical to those shown in Hegerl et al. (2000) and Barnett et al. (1999). The gray bars indicate the 90% uncertainty range for a greenhouse gas and an orthogonal sulfate aerosol signal estimated from observations; the gray ellipse indicates the uncertainty in the combined estimate of both. The uncertainty ellipse in (b) encompasses zero, indicating that both signals cannot be detected simultaneously with confidence if HadCM2 fingerprints are used. However, the probability that the greenhouse gas signal is zero or smaller is less than 5% (independent of which fingerprints are used) as indicated by its uncertainy bar, and a combined greenhouse gas and sulfate aerosol signal is also clearly detectable (Hegerl et al. 2001). A significant sulfate signal can only be detected with ECHAM3/LSG fingerprints. The amplitudes of both signals from the observations are compared with those from model simulations forced with various forcing histories and using different climate models.Circles refer to simulations using greenhouse gas and sulfate aerosol forcing without (GS) or with indirect sulfate aerosol effect (GSI); crosses refer to simulations with greenhouse gas forcing only (*G*). The numbers behind the kind of forcing refer to the model used in the simulation [1: HadCM2, 2: ECHAM3/LSG, 3: GFDL, 4: ECHAM4/OPYC, 5: CCCmal, and 6: CCCma2; for more details see Barnett et al. (1999)]. A consistency test establishes whether the difference between a model estimate and the observed amplitude estimate is significantly larger than the combined uncertainty in the observations (due to internal variability and observational uncertainty, indicated by the gray ellipse) and the model simulation (due to internal variability; an example for GS3, which is an ensemble average of five simulations, is given by the dashed ellipse). Simulations shown in black are consistent with observations; simulations that are inconsistent at the 10% significance level are shown in gray. Model simulations for which only a single ensemble member is available are illustrated by thin symbols; those based on ensembles of simulations are shown by fat symbols. Results from consistency tests indicate that most greenhouse gas–only simulations are inconsistent with observations. The simulation of natural climate forcings only (nat, solar + volcanic) is also inconsistent with observations, indicating that natural signals alone cannot be mistaken as anthropogenic signals. About one-half of the GS simulations are in agreement with observed trend patterns.

er estimate of the direct aerosol forcing than is the case for most of the other models.

Amplitude estimates from observations are subject to uncertainty in the observations due to internal climate variability and differences in sampling of observational and model data (Hegerl et al. 2001), those from model simulations due to sampling uncertainty in a small ensemble (see dashed ellipse in Fig. 1). The attribution method outlined in Hasselmann (1997) tests whether the difference of the $a_G\, a^*_{GS}$ amplitudes in observations and in a particular model simulation is significant relative to (i.e., is unlikely to be accounted for by) these uncertainties. Uncertainty due to systematic model and forcing errors is not taken into account directly but is demonstrated by the spread of estimates from the different model simulations, some of which are forced by different sulfate aerosol forcing, and one of which incorporates natural forcing (''GS + nat 1''). Note that most greenhouse gas–only simulations are rejected since their greenhouse gas amplitudes are too strong, and some of the combined greenhouse gas–plus–sulfate aerosol simulations are also rejected. A simulation forced with natural forcing only (''nat,'' model 1) is also inconsistent, but superimposing natural forcing onto anthropogenic signals improves the agreement (Barnett et al. 1999; Hegerl et al. 2001; Tett et al. 1999).

As in any stepwise regression problem, this approach is not symmetric, since a different result would be obtained if the analysis started with the combined greenhouse gas–plus–sulfate aerosol response pattern $\mathbf{g}_{GS}$ and then introduced an orthogonalized version of the greenhouse pattern $\mathbf{g}_G$. The justification of using stepwise regression is that the second signal is less known or less important than the first. Our confidence in the nature and reliability of the sulfate response is definitely much lower than our confidence in the greenhouse gas response: see, for example, Houghton et al. (2001). How-

ever, stepwise regression does not truly separate contributions from both forcing agents due to the above-mentioned asymmetry. While the amplitude estimate of the orthogonal sulfate aerosol signal is independent of the greenhouse gas signal, that of the greenhouse gas signal contains contributions of the sulfate aerosol signal that are colinear with the greenhouse gas signal. Therefore the stepwise approach is fully justified and its results are easily interpreted only if the sulfate response were to represent a relatively small and poorly characterized perturbation on the main greenhouse signal. However, the magnitude of the sulfate signal appears to be comparable to the greenhouse response.

Figure 1 illustrates this problem: since the sulfate-induced cooling pattern is negatively correlated with the greenhouse gas warming pattern, the amplitude of sulfate cooling will influence a_G, the total amplitude of the greenhouse gas signal. Thus this total greenhouse gas signal amplitude in simulations run with anthropogenic sulfate forcing decreases relative to the simulations using greenhouse gas forcing alone, as seen when comparing the greenhouse gas simulations G1 and G2 with the simulations with additional sulfate forcing GS1 and GS2 (note that model 4 is an exception, but since this model has only single-member ensembles, the uncertainties are very large). Note also how locations of models driven with GS forcing range from upper left to lower right, meaning greenhouse amplitude is anticorrelated with orthogonal sulfate amplitude. Model simulations for which orthogonal sulfate amplitude is smaller tend to show larger overall warming, which is consistent with smaller sulfate cooling. Therefore this does not necessarily indicate there is anything wrong with the models' greenhouse gas responses.

3. Multiregression using correlated signals

An alternative to stepwise regression described above is to set the problem up as a multiple regression, estimating both greenhouse and sulfate signals simultaneously from the data, as was done by Allen and Tett (1999), Tett et al. (1999), and Stott et al. (2001). This approach treats both response patterns ''equally'' as if they were subject to the same intrinsic level of uncertainty. If $\mathbf{G}$ refers to the matrix $[\mathbf{g}_G, \mathbf{g}_{GS}]$ and matrix products are also defined in terms of $\mathbf{C}_N^{-1}$ as above, then an estimate of the regression coefficients b_i in (1) is given by

$$\tilde{b}_i = (\mathbf{G}^T \cdot \mathbf{G})^{-1}\mathbf{G}^T \cdot \mathbf{y}, \tag{11}$$

or, equivalently (given the normalization),

$$\begin{pmatrix} \tilde{b}_G \\ \tilde{b}_{GS} \end{pmatrix} = s^{-2}\begin{pmatrix} 1 & -r \\ -r & 1 \end{pmatrix}\begin{pmatrix} \mathbf{g}_G^T \\ \mathbf{g}_{GS}^T \end{pmatrix} \cdot \mathbf{y}. \tag{12}$$

The results presented in Hegerl et al. (1997, 2000, 2001) and Barnett et al. (1999) can be represented in terms of a regular multiple regression (to within a constant factor of s^{-1} that is not relevant for the outcome of detection and attribution tests) by orthogonalizing *both* the G and GS fingerprints with respect to each other, applying (4) and (5) to introduce

$$\mathbf{g}_G^* = s^{-1}(\mathbf{g}_G - r\mathbf{g}_{GS}) \tag{13}$$

$$= s\mathbf{g}_G - r\mathbf{g}_{GS}^*. \tag{14}$$

Hence the amplitude of the ''orthogonalized greenhouse fingerprint'' is given by a simple linear combination of the coordinates of the points shown in Fig. 1, with

$$a_G^* = sa_G - ra_{GS}^*, \tag{15}$$

and with a_{GS}^* remaining as it was. We stress that this transformation affects all points and uncertainty ellipses in the same way, so while it alters the visual impression of the figure and helps to understand whether the model–data discrepancies originate from the greenhouse gas or sulfate aerosol signal, it has no impact on formal attribution results (whether a particular model simulation is consistent with the observations).

One further step is required to ensure that greenhouse gas and sulfate aerosol signals are being treated as equally as possible, taking into account the fact that they were not treated equally in the design of the original model simulations: both $\mathbf{g}_G$ and $\mathbf{g}_{GS}$ were estimated from simulations subject to greenhouse forcing, whose magnitude was identical in the greenhouse gas–only and combined greenhouse gas–plus–sulfate aerosol simulation. If we assume $\mathbf{g}_{GS}$ is a sum of of pure greenhouse and sulfate signals (there is no evidence on the timescales applied here of nonlinear interactions between greenhouse and sulfate responses), $\mathbf{g}_{GHG}$ and $\mathbf{g}_{Sul}$, then

$$\mathbf{g}_{GS} = \mathbf{g}_{GHG} + \mathbf{g}_{Sul} \quad \text{and} \tag{16}$$

$$\mathbf{g}_G = \rho\mathbf{g}_{GHG}. \tag{17}$$

An additional normalization constant ρ becomes necessary to reconstruct the original amplitude ratio of the greenhouse gas and combined forcing simulations, since the fingerprints have been normalized differently [(7) and (8)].[1]

We substitute $\mathbf{g}_G$ and $\mathbf{g}_{GS}$ in (1) by (16) and (17) and

[1] Hegerl et al. (1997, 2000) applied an additional normalization of $\mathbf{g}_G$ and $\mathbf{g}_{GS}$ prior to (7) and (8) since both signals were derived from empirical orthogonal functions of data from simulations of the twentieth and twenty-first century. The latter normalization constants were no longer available, and a recomputation was unpractical. Therefore, we evaluate the amplitude of $\mathbf{g}_G$ and $\mathbf{g}_{GS}$ in the simulations from which they were derived, taking an average of five consecutive amplitude estimates from trends ending around the year analyzed in the observations. There is clearly some arbitrariness here, but results shown are very insensitive to changes in the averaging length. The constant ρ is given by the ratio between these modeled signal amplitudes, $\rho = \check{a}_{GS}/\check{a}_G$. It will generally be smaller than 1, since the amplitude of a normalized greenhouse signal in a greenhouse-only simulation will be larger than the amplitude of a normalized combined greenhouse-plus-sulfate signal in a greenhouse-plus-sulfate simulation (i.e., trends are larger in greenhouse-only simulations).

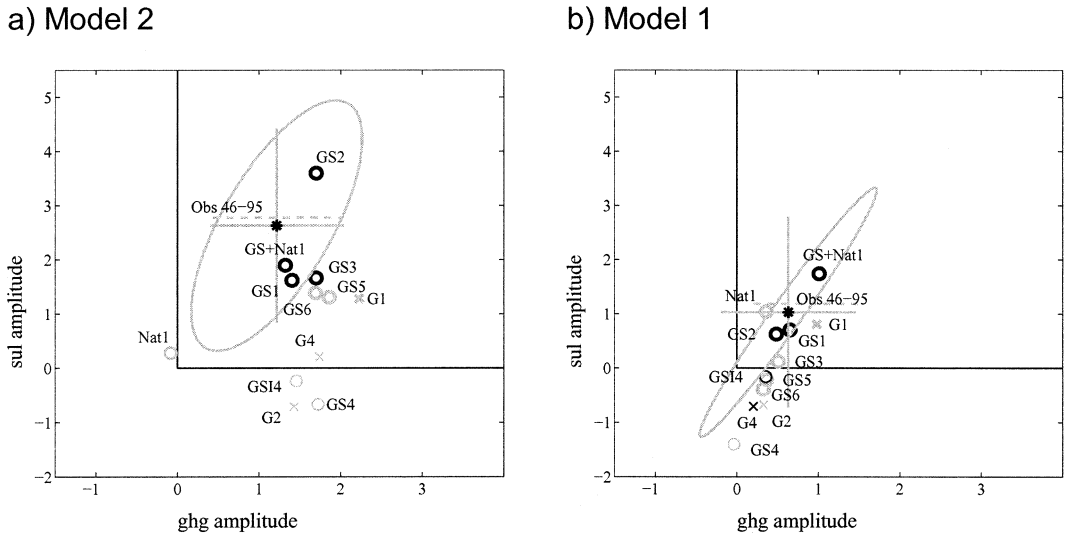

FIG. 2. Greenhouse and sulfate response amplitudes, a_{GHG} and a_{Sul}, after transformation rendering them mutually independent (notation as in Fig. 1; the horizontal axis refers now to the amplitude of a greenhouse gas fingerprint; the vertical axis refers to the amplitude of a sulfate aerosol fingerprint). Note that the same model simulations are consistent with observations as in Fig. 1, but that the separation between a greenhouse-type signal and a sulfate signal is now different. In particular, we fail to detect both a greenhouse gas and a sulfate signal in (b), since now the two signals are very highly correlated, making a separation of the signals difficult [note that processing time information also helps to separate both signals in that model, Tett et al. (1999)]. If the part of the uncertainty range that encompasses negative amplitudes of sulfate aerosol forcing (''sulfate warming'') is excluded, a significant greenhouse gas signal is detected (dashed gray bar). The wide spread of model results along the sulfate axis shows that estimates of a sulfate aerosol signal are model dependent and very uncertain and that discrepancies between models and observations originate mainly from errors in the sulfate amplitude.

rearrange to obtain an estimate of the amplitude of the net response to greenhouse gases $\mathbf{g}_{\mathrm{GHG}}$ that is physically independent from the response to sulfate aerosols $\mathbf{g}_{\mathrm{Sul}}$

$$a_{\mathrm{GHG}} = \rho a^*_G + a^*_{\mathrm{GS}} \tag{18}$$

with the response to sulfates given, as before, by $a_{\mathrm{Sul}} = a^*_{\mathrm{GS}}$. There is, of course, an element of conventionality in our definition of a_{GHG}: we have chosen to leave the vertical (sulfate) amplitude unaffected by the transform to make comparison between figures easier. Thus horizontal (greenhouse) amplitudes in Fig. 2 are given by

$$a_{\mathrm{GHG}} = \rho s a_G + (1 - \rho r) a^*_{\mathrm{GS}}, \tag{19}$$

with vertical (sulfate) amplitudes unchanged from Fig. 1.

4. Example: Transformation to multiregression results

Figures 2a,b show a_{GHG} and a_{Sul} for the same cases as shown in Fig. 1; the coefficients used in the transformation are given in Table 1. In Fig. 2a, model-estimated signal amplitudes from GS-forced runs show a wide spread along the vertical axis—while the same number of models are found to be inconsistent with the observations as before (consistency tests are unaffected by the transformation), it is now clear that the origin of the discrepancies lies primarily in the model-simulated sulfate responses rather than their simulated greenhouse responses. The latter, although generally slightly stronger than the observed greenhouse response, are all sim-

TABLE 1. Data for the transform from stepwise regression to multiple regression using signal patterns from ECHAM3/LSG (″MPI″) and HadCM2 (″Had″) (see equations). The transform is derived for trends for 1946–95 (MPI95, Had95) and 1949–98 (MPI98, Had98). Amplitude estimates from the observed trend patterns are given in a_G, a_{GS}, a^*_G, a^*_{GS}, and a^*_{GHG} (units of std dev of HadCM2 noise). Differences in transform coefficients between 1995 and 1998 are due to differences in normalization and to slight changes in coverage, affecting particularly r and ρ. Here $\boldsymbol{\delta}T_{\mathrm{GHG}}$ and $\boldsymbol{\delta}T_{\mathrm{Sul}}$ gives the best estimate of global warming attributable to greenhouse gas and sulfate aerosol forcing respectively [K (50 yr)]$^{-1}$.

Signal	a_G	a_{GS}	r	s	a^*_G	a^*_{GS}	ρ	a^*_{GHG}	$\boldsymbol{\delta}T_{\mathrm{GHG}}$	$\boldsymbol{\delta}T_{\mathrm{Sul}}$
MPI95	2.03	2.90	0.921	0.390	−1.63	2.63	0.78	1.22	0.46	−0.22
Had95	1.97	2.11	0.986	0.164	−0.69	1.03	0.57	0.63	0.52	−0.34
MPI98	3.11	3.65	0.928	0.374	−0.75	2.06	0.51	1.61	0.80	−0.44
Had98	2.97	3.12	0.987	0.163	−0.68	1.18	0.64	0.75	0.59	−0.30

ilar to the observed greenhouse signal (except for the simulation forced without greenhouse gas forcing ''nat''). Figure 2b shows the very strong correlation between a_{GHG} and a_{Sul} using the HadCM2 fingerprints—a consequence of the correlation between the HadCM2 *G*- and GS-simulated response patterns being 0.98 in this diagnostic. Uncertainty ranges are increased because of this degeneracy, so that only the combined anthropogenic signal can now be detected (see Barnett et al. 1999), but greenhouse gas and sulfate signals can not be individually detected. A true separation between greenhouse gas and sulfate signal in multiple regression, which makes it simpler to interpret results physically, comes at the cost of losing the detection of a greenhouse gas–type pattern. Again, model-simulated greenhouse responses are consistent with the observed response.

Given the high correlation between the *G* and GS signal patterns and the high level of uncertainty in the individual models' sulfate signals, the way the different models project onto the sulfate signal is largely a matter of chance: the best-guess estimated amplitudes of this signal in models 4, 5, and 6 are negative without physical justification (although, given the sampling uncertainties in these small ensembles, not significantly so). Also, the observed warming in Fig. 2b could be anything between a combination of high greenhouse and sulfate amplitudes or even negative amplitudes of both signals in that diagnostic. It can be argued that, despite the considerable uncertainty in its amplitude, the *sign* of sulfate aerosol forcing and response is known ahead of time and is unlikely to be incorrectly simulated by a climate model (Houghton et al. 2001). Note that the patterns and signs of other types of aerosol forcing, such as soot and mineral dust, are poorly known. We assume here that their projection on greenhouse gas or sulfate aerosol response patterns is small; we note, however, that this assumption introduces some uncertainty into our reasoning. Under the assumption of the correct sign of sulfate aerosol forcing and model response, we exclude all estimates of the joint (a_{GHG}, a_{Sul}) distribution for which a_{Sul} is negative. The uncertainty ranges are then based on the sampling distribution of a large number of Monte Carlo simulations (covering the range of signal estimates given the uncertainty due to climate noise and observational sampling error), where all samples with a negative sulfate signal are excluded. After thus renormalizing the distribution above the horizontal axis in Fig. 2, we detect a significant and positive greenhouse gas signal (dashed horizontal line) at the 5% significance level.

The discrepancies between the estimated amplitudes of the sulfate response in the different model simulations and observations suggest that the inconsistencies between modeled and observed amplitude estimates reported in Barnett et al. (1999) originate mainly from errors in the sulfate aerosol forcing and modeled climate response. Direct radiative forcing by sulfate aerosols is difficult to measure and depends on assumptions about transport and lifetime of sulfate aerosols generated from sulfate emissions. The processes involved in indirect sulfate aerosol forcing (through sulfate aerosol affecting cloud droplet availability) are even more uncertain (Houghton et al. 2001). Recent measurement campaigns illustrate that we still know very little about the magnitude and mechanisms of forcing through sulfate aerosols and other related anthropogenic forcing agents such as soot (e.g., Satheesh and Ramanathan 2000; Ackerman et al. 2000). Not all GS simulations shown in the figures apply the same forcing; in particular, ECHAM4/OPYC has weaker (but possibly more realistic) direct forcing, while the simulation GSI has additional indirect sulfate aerosol forcing. The fact that these simulations show particularly strong disagreement with the observations and the other model's signals illustrates this high uncertainy in the sulfate forcing. In addition to the large uncertainty in sulfate aerosol forcing, model response to this forcing is also highly uncertain and differs between models. This uncertainty is illustrated by the difference in results due to estimating sulfate aerosol signals using signal patterns $\mathbf{g}_{GS}$ from two different models with nearly identical sulfate forcing, shown in Figs. 2a,b.

The results from optimal detection can also be used to estimate how much of the observed warming is attributed to which external forcing (e.g., Tett et al. 1999; Allen et al. 2001). We do this by computing the spatial means of the greenhouse gas and sulfate aerosol signals, $\mathbf{g}_{GHG} = \mathbf{g}_G/\rho$ and $\mathbf{g}_{Sul} = \mathbf{g}_{GS} - \mathbf{g}_{GHG}$, and multiplying the results by a_{GHG}/s and a_{Sul}/s, respectively. The best guess for the anthropogenic contribution for trends ending in 1995 is given in Table 1.

Since more data have become available, we have updated the results for JJA trend patterns ending in 1998 (note that the naturally forced simulations end in 1996 and are therefore not shown). Figure 3 shows the results in units of global mean warming/cooling attributable to greenhouse gases and sulfate aerosols (note that consistency results do not depend on units). The most notable changes are that now a positive greenhouse gas signal is detected independent from which model's signal pattern is used, and the possibility of both signals simultaneously being zero is rejected (since the ellipse no longer encompasses the origin). Some of the models that were found to be inconsistent with observations (at a 10% significance level) for trends ending in 1995 are no longer rejected, and the greenhouse gas amplitude is no longer systematically larger in model simulations than is observed. This difference in results demonstrates the influence of climate variability: trends ending in 1995 were somewhat smaller than those ending in 1998, due both to conditions at the beginning of the trend being warmer in the mid-1940s than in the late 1940s and to 1998 being influenced by a strong El Niño event (note, however, that the estimate of climate variability is based on a climate model with El Niño events of fairly realistic amplitude; therefore the uncertainty rang-

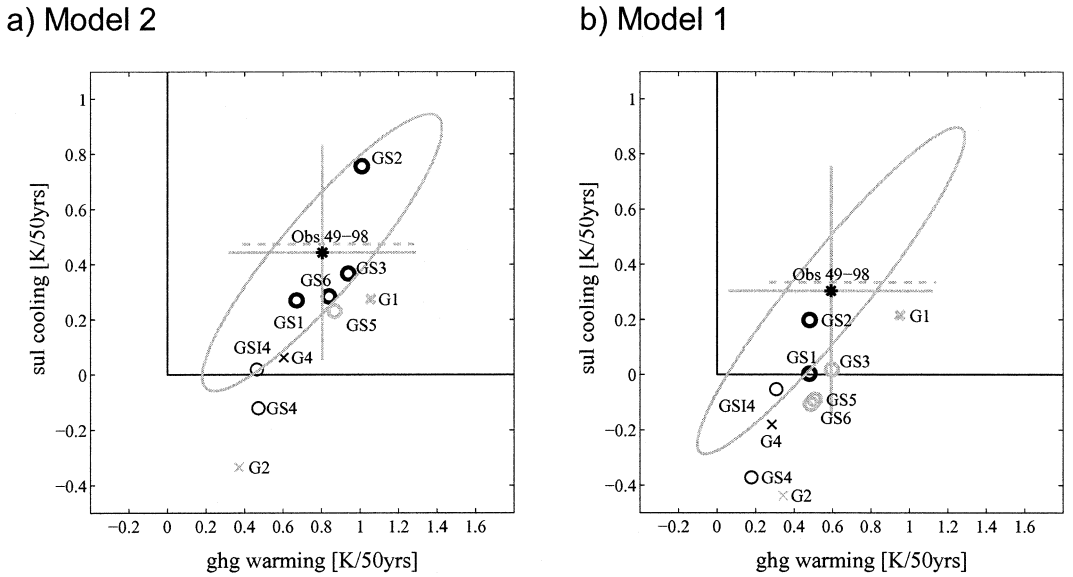

FIG. 3. Results of a multiregression using JJA trend patterns for 1949–98 in both models and observations. The units are estimated anthropogenic contributions of greenhouse gas and sulfate aerosol forcing to observed and simulated JJA trend patterns for 1949–98 [(K 50 yr)$^{-1}$; note that this change in units does not affect consistency results]. The sulfate axis is inverted for better comparison with other figures (i.e., units denote cooling rather than warming). (a) Results using ECHAM3/LSG fingerprints. (b) Results using HadCM2 fingerprints. Note that now a positive greenhouse gas signal can be detected independent of which model's fingerprints are used and of assumptions on the sign of the sulfate aerosol signal.

es should generally account for variations due to El Niño events). Fluctuations in the significance levels of results from one period to the next are to be expected because of the statistical nature of the attribution assessment. As expected, results are particularly uncertain for the model for which only a single realization was available (ECHAM4/OPYC). The overall impression of the results remains similar to that of Fig. 2.

The estimated anthropogenic contribution to global warming by sulfate aerosols and greenhouse gases is comparable to those shown in Allen et al. (2001), with anthropogenic signals being slightly larger here. The latter is the case because we consider only the most recent 50 years (during which the warming was particularly strong), while Allen et al.'s results are derived from anomalies for the entire twentieth century. The best-guess estimates suggest that most of the recent observed warming trend is anthropogenic (note that the JJA trend over the recent 50 years is 0.26 K for trends ending in 1995 and 0.37 K for trends ending in 1998; see Table 1 for estimated anthropogenic contributions). The difference between results based on model 1 and model 2 (Figs. 2b, 3b versus 2a, 3a) suggests that presently a separation of the observed warming into a greenhouse gas and sulfate aerosol component is uncertain. Nevertheless, the 5%–95% range in the trend attributable to anthropogenic greenhouse gases over the period of 1948–98 is 0.32–1.29 K (50 yr)$^{-1}$ (model 2) or 0.06–1.12 K (50 yr)$^{-1}$ (model 1). The larger range obtained with signals from model 1 is primarily due to the strong correlation between greenhouse and sulfate signals simulated by this model when expressed in terms of 50-yr trends. If the amplitude of the sulfate signal is constrained to be positive (i.e., if we exclude the possibility of sulfate-induced warming), we find that the 5%–95% range in the 1948–98 trend attributable to anthropogenic greenhouse gases is 0.28–1.16 K (50 yr)$^{-1}$ (model 1) and 0.39–1.29 K (50 yr)$^{-1}$ (model 2).

Note that with a spatiotemporal diagnostic, the separation between sulfate aerosol and greenhouse gas signal is less ambiguous (e.g., Tett et al. 1999; and Allen et al. 2001) and uncertainty ranges become substantially smaller. Gillett et al. (2002) demonstrate that this is due to the use of space–time diagnostics using annual data in the latter work [which had become feasible since, in general, ensembles of a larger size are available at the present than were available at the time of the Hegerl et al. (1997) analysis] and that, if the Allen et al. (2001) algorithm is applied to boreal summer trends, the results closely resemble the multiregression results shown in the present paper. Nevertheless, the spread of results in a joint fingerprint analysis (here, derived from two different models) emphasizes that presently modeled and observed responses to sulfate aerosol forcing both remain uncertain.

5. Conclusions

Attempts to separate greenhouse gas and sulfate aerosol signals in observations have relied on both stepwise and multiple regression approaches, making a comparison of results difficult. The difference between both is

that in multiple regression all signals are treated equally, while in stepwise regression the signals that are searched for in data first are considered more reliable or important than the other signals. While the preference of one method over the other is a matter of choice dependent on the nature of a problem, results from multiple regression are easier to interpret physically. It is demonstrated that results from stepwise regression can be easily transformed to multiple regression results and that attribution results, which test the consistency between model simulations and observations, are not affected by this transform. We have applied the transform to stepwise regression results from Hegerl et al. (1997, 2001) and Barnett et al. (1999), where first a greenhouse gas signal and second a sulfate aerosol signal are estimated from observations. After the transform it becomes clear that inconsistencies between model simulations and observations discussed in Barnett et al. (1999) originate from different amplitudes of the sulfate aerosol signal in different models and in observations. This emphasizes the importance of uncertainty both in the sulfate forcing and the model response to it.

Results from multiple regression also allow for attributing fractions of the observed warming to the considered forcings. Results indicate that, independent of which model's fingerprints were used, most of the recent warming is likely due to increasing greenhouse gases.

Acknowledgments. The authors thank David Ritson for many helpful questions and suggestions leading to this work. We also thank Francis Zwiers, Stephen Leroy, Tim Barnett, John Mitchell, Jeff Kiehl, and an anonymous reviewer for helpful suggestions. Different components of the work were sponsored by (for GH) NOAA's Office of Global Program's and DOE's Office of Energy Research in conjunction with the Climate Change Data and Detection Program Element, and NSF (Grants ATM-0096017 and ATM-0002206), by (MR) the Hadley Centre, under UK-DETR and UK-NERC, and by the EC Project QUARCC.

REFERENCES

Ackerman, A. S., O. B. Toon, D. E. Stevens, A. J. Heymsfield, V. Ramanathan, and E. J. Welton, 2000: Reduction of tropical cloudiness by soot. *Science,* **288,** 1042–1047.

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

——, and Coauthors, 2001: Quantifying anthropogenic influence on recent near-surface temperature change. *Rev. Geophys.,* in press.

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

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

Hasselmann, K., 1979: On the signal-to-noise problem in atmospheric response studies. *Meteorology over the Tropical Oceans,* B. D. Shaw, Ed., Roy. Meteor. Soc., 251–259.

——, 1993: Optimal fingerprints for the detection of time dependent climate change. *J. Climate,* **6,** 1957–1971.

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

Hegerl, G. C., and G. R. North, 1997: Statistically optimal methods for detecting anthropogenic climate change. *J. Climate,* **10,** 1125–1133.

——, H. von Storch, K. Hasselmann, B. D. Santer, U. Cubash, and P. D. Jones, 1996: Detecting greenhouse gas–induced climate change with an optimal fingerprint method. *J. Climate,* **9,** 2281–2306.

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

——, P. A. Stott, M. R. Allen, J. F. B. Mitchell, S. F. B. Tett, and U. Cubasch, 2000: Optimal detection and attribution of climate change: Sensitivity of results to climate model differences. *Climate Dyn.,* **16,** 737–754.

——, P. D. Jones, and T. P. Barnett, 2001: Effect of observational sampling error on the detection of anthropogenic climate change. *J. Climate,* **14,** 198–207.

Houghton, J. T., Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, and C. A. Johnson, Eds., 2001: *Climate Change 2001: The Scientific Basis.* Cambridge University Press, 881 pp.

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

North, G. R., and M. Stevens, 1998: Detecting climate signals in the surface temperature record. *J. Climate,* **11,** 563–577.

Satheesh, S. K., and V. Ramanathan, 2000: Large differences in tropical aerosol forcing at the top of the atmosphere and Earth's surface. *Nature,* **405,** 60–63.

Stott, P. A., S. F. B. Tett, G. S. Jones, M. R. Allen, W. J. Ingram, and J. F. B. Mitchell, 2001: Attribution of twentieth century temperature change to natural and anthropogenic causes. *Climate Dyn.,* **17,** 1–21.

Tett, S. F. B., P. S. Stott, M. R. Allen, W. J. Ingram, and J. F. B. Mitchell, 1999: Causes of twentieth century temperature change. *Nature,* **399,** 569–572.