The Potential Effect of GCM Uncertainties and Internal Atmospheric Variability on Anthropogenic Signal Detection

TIM P. BARNETT

Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California

GABRIELE C. HEGERL

Max-Planck Institute for Meteorology, Hamburg, Germany

BEN SANTER AND KARL TAYLOR

Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, Livermore, California

(Manuscript received 6 May 1996, in final form 14 July 1997)

ABSTRACT

When long integrations of climate models forced by observed boundary conditions are compared against observations, differences appear that have spatial and temporal coherence. These differences are due to several causes, the largest of which are fundamental model errors and the internal variability inherent in a GCM integration. Uncertainties in the observations themselves are small in comparison. The present paper constitutes a first attempt to compare the time dependence of these spatial difference patterns with the time dependence of simulated spatial patterns of climate change associated with anthropogenic sources.

The analysis procedure was to project the model minus observed near-surface temperature difference fields onto estimates of the anthropogenic ''signal'' (in this case the response to greenhouse-gas and sulfate-aerosol forcing). The temporal behavior of this projection was then compared with the estimated temporal evolution of the anthropogenic signal. Such comparisons were performed on timescales of 10, 20, and 30 yr. For trends of only 10 yr in length, the model minus observed spatial difference patterns are of the same magnitude and have the same time rate of change as the expected anthropogenic signal. In the case of 20- and 30-yr trends, the prospects are favorable for discriminating between temperature changes due to anthropogenic signal changes and changes associated with model minus observed difference structures. This suggests that attempts to quantitatively detect anthropogenic climate change should be based on temporal samples of at least several decades in length. This study also shows the importance of distinguishing between purely statistical detection and what the authors term practical prediction. It is found that the results of the detection analysis are sensitive to the spatial resolution at which it is performed: for the specific case of near-surface temperature, higher spatial resolution improves ability to discriminate between an anthropogenic signal and the type of model error/internal variability ''noise'' considered here.

1. Introduction

The past few years have seen a rapid improvement in the methodology used to detect greenhouse-gas (GHG) signals and other combinations of anthropogenic effects in observed global climate data [see, e.g., Santer et al. (1996a) for a summary]. The first practical applications of a pattern recognition approach to the GHG signal detection problem were by Barnett (1986), Barnett and Schlesinger (1987), and Barnett et al. (1991) for equilibrium response signals and Barnett (1991) for transient response signals. Some aspects of the pattern correlation statistic used in this method were discussed and elucidated by Santer et al. (1993). This approach has been advanced through the work of Hasselmann (1979, 1993, 1997), North et al. (1995), and North and Kim (1995). These studies made the major advance of casting the detection problem in an ''optimal search framework'' that maximizes the chance of detecting a known signal if it is present in observed data (see Hegerl and North 1997).

The recent theoretical advances noted above have been put into practical application most notably by Hegerl et al. (1996, 1997), Hasselmann et al. (1995), North and Stevens (1998), Santer et al. (1995a), and Santer et al. (1996b). Although the methodologies used by these authors differ in important ways, all of the above-mentioned studies suggest that, in the case of atmospheric

Corresponding author address: Dr. Tim P. Barnett, Scripps Institution of Oceanography, University of California, San Diego, 9500 Gilman Dr. Dept. 0224, La Jolla, CA 92093-0224.
E-mail: tbarnett@ucsd.edu

temperature observations, an anthropogenic signal is now beginning to emerge from the background noise of natural climate variability. The recent studies of Santer et al. (1996b) and Tett et al. (1996), building on earlier work by Karoly et al. (1994), suggest that a signal due to the combined effect of GHGs and aerosols may now be detectable in the temperature structure of the global atmosphere, although there are numerous caveats associated with these investigations.

Our goal in this paper is to examine the potential impact that some of the uncertainties in different components of the general circulation model (GCM) signal estimates (and in the observations themselves) may have on a detection strategy. The effect of such errors, if considered at all, has been treated in a highly simplified fashion in the studies noted above. There are four types of uncertainty that are relevant for signal detection strategies.

1) The first is associated with errors in the forcing fields used to drive the models. These errors may result from either omission of certain forcings and/or from inaccurate specification of forcings that are included in the model calculations. This category encompasses both natural forcings, such as changes in the solar irradiance or the volcanic dust loading of the atmosphere [these effects have been generally neglected in pattern-based detection work, with one or two exceptions (see, e.g., Cubasch et al. 1997)] and anthropogenic forcings. The major uncertainty in terms of anthropogenic forcing relates to the magnitude, patterns, and time evolution of sulfate aerosol direct and indirect effects.

2) The second type of uncertainty is associated with errors in the response of GCMs to specified forcing—that is, errors in the basic physics or in the parameterization of physical processes whose spatial scales cannot be directly resolved. Examples include neglected feedbacks, inadequate representation of sea ice, errors in modeling cloud radiative properties, etc.

3) The third type of uncertainty is associated with model representation of natural internal climate variability. This may arise from model errors in simulating the patterns, amplitude, and timescales of internal variability. However, even a model that perfectly represented the dominant modes of natural internal climate variability and had *no* signal uncertainties of types 1 or 2 would show model-observed differences over time and space due to inherently unpredictable atmospheric and oceanic variability (cf. Barnett 1995), which occurs in observations as well as in climate models.

4) The fourth source of uncertainty is introduced by random and systematic errors in the observations used to verify GCM simulations of present-day and historical climate [see Nicholls et al. (1996) for further details of such errors].

In practice, it is difficult to partition the individual components contributing to model-observed differences (see, e.g., Zwiers 1996). The current work concentrates on uncertainties of types 2, 3, and 4. Since this paper is a first attempt to estimate the impact of these uncertainties on anthropogenic signal detection, there are a number of caveats associated with our work. These will be noted at appropriate places in the text.

The main question we pose is, does the space–time structure of the differences between observations and the models' representation of the climate of the last few decades resemble the pattern of climate change expected to occur due to anthropogenic influences? If it does, the detection problem has associated with it a higher level of uncertainty than it would if the predicted signal patterns were orthogonal to the patterns characteristic of model versus observed differences. This is so because the difference fields then could alternatively, erroneously strengthen the anthropogenic signal, making it appear real when it is not, or obscure the signal if they differ in sign from it.

Another way to pose the question is to ask if the model-observed difference field projects well onto the expected anthropogenic signal pattern. If this is the case, then it becomes harder to say with confidence that an anthropogenic signal has been detected until it is substantially larger than the difference field signature.

In the following section we describe briefly the data and techniques used in this study. Subsequent sections evaluate the degree of correspondence between the model-observed difference structure and several anthropogenic signals and then summarize the results of this study.

2. Data: Observed and model generated

We computed differences in near-surface air temperature between a long GCM integration (forced by prescribed SST and sea-ice distributions) and observations. These are referred to hereafter as difference fields. Calculations were performed on a global grid. Note that the difference fields reflect all four sources of uncertainty mentioned above: missing or inaccurately specified forcing, real-model errors, the effects of internal variability, and observational uncertainties.

a. Observed data

The observed data came from two sources.

(i) Our analysis of model-observed difference fields primarily utilized the observed near-surface temperature dataset compiled by Jones and Briffa (1992) at the Climatic Research Unit (CRU) of the University of East Anglia. These data have global coverage and were available in the form of monthly means on a 5° lat × 5° long grid. To facilitate comparison with model results (see section 2b), we used observed data for the period 1950–94 and interpolated these to the T42 grid of the ECHAM3 atmospheric GCM (AGCM) (approximately 3° × 3° horizontal resolution).

(ii) We also made use of near-surface temperatures from a 63-station network of radiosonde data (Angell

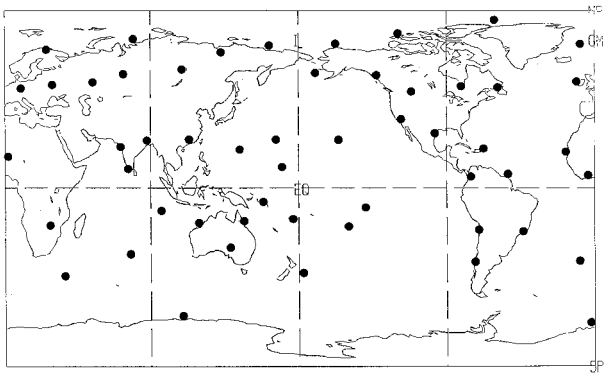


FIG. 1. The location of the 53 radiosonde stations used in this study.

and Korshover 1983). We employ a subset of 53 of these data for two purposes: to study the effects of observational uncertainty on anthropogenic signal detection, and to consider how signal detection might be influenced by the spatial resolution at which the detection analysis is performed (section 5). The second issue is particularly relevant in the light of previous detection studies, such as those by Barnett (1986) and Barnett and Schlesinger (1987), that have been performed at the spatial resolution of the Angell station network (see Fig. 1). While this resolution may have been adequate to capture the salient features of a large-scale GHG signal, there is some concern as to whether it can fully represent the most important components of smaller-scale anthropogenic signals due to combined GHG + aerosol effects (see, e.g., Trenberth and Olson 1991). Furthermore, using a more complete spatial network allows small-scale climate variations to cancel out over large spatial scales (Bell 1986), providing a clearer view of the anthropogenic signal.

The radiosonde estimates of monthly mean near-surface air temperature, originally obtained from R. Jenne (National Center for Atmospheric Research), were carefully quality controlled prior to their use here. This quality checking was accomplished by plotting the twice daily data from the stations and then inspecting these plots for outliers. Monthly mean values were then estimated from the twice daily data that satisfied quality control checks, provided that at least one-half of the data for a specific month were present. Gaps in monthly values were filled by linear interpolation.

Note that the analysis of the effects of observational uncertainty was performed at the coarse spatial resolution of the radiosonde network. This required subsampling the CRU data at grid points closest to the upper-air station locations. Data from the Angell radiosonde network were not used in the construction of the CRU dataset, so the two estimates of near-surface temperature change are quasi-independent observationally (P. D. Jones 1996, personal communication).

b. Simulation of recent climate: Difference fields

The model data used for estimating the model-observed difference fields were taken from an extended integration with ECHAM3, a state-of-the-art atmospheric GCM (Roeckner et al. 1992), whose use here is courtesy of L. Bengtsson [Max Planck Institute for Meteorology (MPI)]. We will refer to this simulation as the RC (recent climate) run. The model version used in this study was run at T42 resolution and forced with observed time-varying global fields of monthly mean sea surface temperatures (SSTs) and climatological sea-ice distribution over the period 1950–94. The SSTs were from the GISST dataset (see acknowledgments). This integration did not involve explicit natural forcing by volcanoes, solar variations, etc., although some effects of these factors may be implicitly incorporated in the time-varying SST. Similarly, any anthropogenic signal that may have existed during this period would also be implicitly included, at least in part, via the SST field. Due to the use of a climatological sea-ice distribution in RC, any component of an anthropogenic signal associated with changes in sea ice is excluded from the model experiment.

The perfomance of the ECHAM3 in such prescribed SST simulations is generally quite good, as has been documented in numerous studies (e.g., Bengtsson et al. 1996; Graham et al. 1994; Barnett et al. 1993; Barnett et al. 1997). In the present demonstration study, we concentrate only on the near-surface air temperature. In the GCM, this is the diagnosed temperature interpolated to a height of 2m.

c. Model simulations of anthropogenic signals

We used estimates of expected anthropogenic signals from coupled model simulations performed at MPI. A T21 version of the ECHAM3 atmospheric model used in section 2b was coupled to the Hamburg large-scale geostrophic (LSG) global ocean model (Maier-Reimer et al. 1993) for transient anthropogenic scenario runs (Voss and Sausen 1996). The coupling procedure, including flux correction, is similar to that of Cubasch et al. (1992). This model is referred to below as HAM3L. It has been used for a multicentury control simulation and for a number of climate change simulations (e.g., Hasselmann et al. 1995; Cubasch et al. 1997; Hegerl et al. 1997).

The transient experiments start with anthropogenic forcing conditions as estimated for the year 1880 and continue up to year 2049. Two simulations (A and B) were carried out, each with time-varying forcing by a combination of greenhouse gases (expressed as equivalent CO_2 concentrations) and direct aerosol effects (modeled as changes in surface albedo). The integrations differ slightly in terms of the applied aerosol forcing (Hasselmann et al. 1995; Hegerl et al. 1997). Henceforth we use the average of the two runs in our calcu-

lations, referring to this average as the GA run. Additional details of the forcing and analysis of the simulations may be found in Hasselmann et al. (1995) and Hegerl et al. (1997).

The time-evolving anthropogenic signals were estimated by subtracting from the above experiments the respective 2-m temperature values from a control simulation with no anthropogenic forcing. This instantaneous differencing reduces the impact of any drift problems the experiments and control run may have in common (Cubasch et al. 1994). Note, however, that the use of instantaneous differences increases the temporal standard deviation of each individual signal estimate by a factor of $\sqrt{2}$. Since the same control run is subtracted from both A and B simulations, this increase in noise is not fully offset by averaging these two integrations.

The procedure for computing an anthropogenic signal (GA minus control run) also minimized the impact of errors in the model's mean field and seasonal cycle on our subsequent analysis. Thus, the anthropogenic signal and the model-observed difference structures (RC minus CRU) have been computed in a similar manner.

3. Analysis

a. Difference field estimation

The first step in estimating the model-observed difference field was to compute model anomaly fields of seasonal mean 2-m air temperatures with respect to ECHAM3's climatological seasonal cycle. This filters from subsequent analysis possible errors in the model's mean field and seasonal cycle. The next step was to compute observed anomalies for the radiosonde and CRU datasets relative to their respective time-averaged seasonal cycles.

The difference field that we study here is simply the difference (at each of roughly 5200 spatial points) between the anomaly time series from the ECHAM3 RC run described in section 2b and the CRU observed dataset. This field, denoted $e(\mathbf{x}, t)$, has seasonal values for each of the 45 yr in the original integration, so it is possible to see if the differences have trends in time, scatter randomly, etc.—behavior that is important in the context of the detection problem. By subtracting the fields as we have done, we have tended to remove at least part of any anthropogenic signal from the e field that may have been present over the last decades. Any anthropogenic signal component that is *not* implicitly incorporated in the time-varying SST forcing will have some signature in $e(\mathbf{x}, t)$.

As noted in section 1, $e(\mathbf{x}, t)$ is influenced by uncertainties arising from at least four different sources, and it is difficult to partition the individual contributions from forcing errors, response errors, internal variability, and observational uncertainty. Previous work has yielded some insights into the size and structure of the internal variability component of $e(\mathbf{x}, t)$. Thus Zwiers (1996), using an ensemble of six AMIP (Atmospheric Model Intercomparison Project; Gates 1992) integrations with an AGCM forced by observed SST fields, showed that the bulk of the interannual variability was over the landmasses and was associated with internal atmospheric dynamics, as opposed to boundary forcing. The variability also increased with latitude. Barnett (1995) found that the natural internal variability inherent in ECHAM3 had similar large-scale spatial coherence. Such coherence and latitudinal amplification are also evident in the model-predicted temperature changes in response to GHG forcing, although simulations that additionally incorporate aerosol effects show greater spatial heterogeneity in their surface temperature response (see, e.g., Kattenberg et al. 1996).

In the following analysis, we are assuming that the difference field estimated from a stand-alone AGCM forced by observed SST is representative of those produced by the CGCM used in the transient scenario runs. This assumption requires that the characteristics of natural internal variability should be similar in the AGCM and the CGCM, and that errors in both models should have similar signatures. We were unable to test this assumption rigorously in the current study. However, recent work by Manabe and Stouffer (1996), who contrasted the characteristics of temperature variability in AGCM runs with three different representations of the ocean (fixed SST, coupling to a mixed-layer ocean, and coupling to a full OGCM), suggests that the assumption of similar internal variability characteristics may not be unreasonable.

Our goal was to see how well the model difference structure projects onto the anthropogenic signal. If it does project well, then there is a similarity between the spatial structure of the difference field and the spatial structure of the anthropogenic signal. Since this difference field has some contribution from bona fide model errors, a large projection of $e(\mathbf{x}, t)$ onto the signal, sustained for several decades or longer, would be of concern. This result could make it difficult to discriminate between detection of a real anthropogenic signal and detection of spurious climate change that may have been partially induced by model errors.

b. HAM3L scenarios

The changes in temperature for the GA run are shown in Fig. 2 for two different seasons and two different averaging periods. These data were analyzed as follows: the temperature change time series at each grid point for the scenario run was area-weighted by the cosine of its latitude, yielding a signal dataset, $s(\mathbf{x}, t)$. Empirical orthogonal functions (EOFs) were then computed from $s(\mathbf{x}, t)$. This was done separately for the December–February (DJF) and June–August (JJA) data. In both cases, the leading eigenmode $f_1(\mathbf{x})$ represented the anthropogenic signal of interest, with the remaining EOFs being indistinguishable from noise processes in time; a

result found previously for this simulation by Hasselmann et al. (1995) and Hegerl et al. (1997).

In the following, the temporal behavior of the dominant signal EOF (i.e., the leading principal component) will be denoted by $A_1(t)$ and is defined as

$$A_1(t) = \int_{\mathbf{x}} s(\mathbf{x}, t) f_1(\mathbf{x}) \cos[\theta(\mathbf{x})]\, d\mathbf{x}. \quad (1)$$

The time-dependent difference amplitude is defined in an analogous way

$$A'_e(t) = \int_{\mathbf{x}} e(\mathbf{x}, t) f_1(\mathbf{x}) \cos[\theta(\mathbf{x})]\, d\mathbf{x} \quad (2)$$

and represents the projection of the time-dependent model-observed difference field, $e(\mathbf{x}, t)$, onto the dominant (and time-independent) scenario response pattern. The inclusion of the cosine (latitude) weight ensures that the model-observed difference field $e(\mathbf{x}, t)$ is treated in exactly the same manner as the anthropogenic signal field $s(\mathbf{x}, t)$. The net result of the above procedure is that $A_1(t)$ is directly comparable to $A'_e(t)$. Additional studies not reported here show that the error structure of the T21 atmospheric model used in the HAM3L GA run were quite similar to those of the T42 atmospheric model, so the fact that the resolution was different in the comparison should not be a first-order problem.

4. Results

a. Description of scenario

The HAM3L GA signal is shown as a difference between two 20-yr means centered on 1890 and 1990 and a similar difference centered on 1980 and 2040. In winter of the earlier period (Fig. 2a, upper panel), temperature increases of order 0.5°–1.0°C occur over most continental areas. Maximum increases of 2°–3°C appear in very high latitudes of the Northern Hemisphere, especially over the Barents Sea and Greenland. There is little or no signal over the bulk of the oceans. During summer (Fig. 2b, upper panel), the pattern of response over Northern Hemisphere land areas is generally weak and of mixed sign. The Southern Hemisphere landmasses show a similar response, except for the strong SST anomalies around the Weddell Sea area. Over the rest of the oceans, the HAM3L GA signal is generally less than 0.5°C.

The projected HAM3L GA signal over roughly the next 60 yr is shown in the lower panels in Fig. 2 (lower). The signals are much stronger (as expected) over all landmasses and in all seasons. The signal now exceeds 1°C over much of the Northern Hemisphere oceans. In short, this global signal, if it occurred, would be apparent without any sophisticated detection scheme.

b. HAM3L projection

The model minus observed difference field was projected onto the HAM3L GA signal patterns as described in Eq. (2). The associated projection time series $A'_e(t)$ (solid line), which spans the years 1950–94, is shown in Fig. 3 for the DJF and JJA seasons. Also shown is the principal component time series $A_1(t)$ (dashed line) for the projection of the GA signal onto the dominant EOF of the GA run. This time series spans the full period 1880–2049. For the purposes of visual comparison, the mean of $A'_e(t)$ was adjusted to be identical to the mean of $A_1(t)$ over the period 1950–94 so as to better compare the two.

Careful inspection of Fig. 3 shows that the projection of the model minus observed difference field $e(\mathbf{x}, t)$ onto the dominant GA signal pattern has fluctuations in projection amplitudes in DJF that are slightly larger than the inherent variability in the signal principal component $A_1(t)$. The same is not true in JJA. Furthermore, there is some evidence in the DJF results that the difference field projection $A_1(t)$ varies on timescales that are at least of order a decade. This suggests that the model minus observed difference field could have some characteristics that are similar to the GA anthropogenic signal over significant lengths of time, thus complicating signal detection.

To address this question, we next compare the rates of change in the $A'_e(t)$ and $A_1(t)$ projection time series given in Fig. 3. We did this by computing linear trends from blocks of both time series, as described below. Our aim is to determine the threshold timescale (block length) beyond which the signal $A_1(t)$ increases much more rapidly than the fluctuations in the difference field projection $A'_e(t)$.

The rate of climate change in $A'_e(t)$ was computed for 5, 6, . . . , 30-yr blocks of data. We allowed the blocks to slide along the 45-yr long $A'_e(t)$ time series, overlapping by all but 1 yr, and computed trends for each block, stipulating only that the current block should fit within the 45-yr data span. This procedure yielded 41 different estimates of trends over 5-yr periods, 40 estimates of trends over 6-yr blocks, and 16 estimates of trend over a 30-yr block. Obviously these estimates are not independent due to temporal autocorrelation in $A'_e(t)$, but the relatively short length of the RC run (and of the observations themselves) makes it difficult to obtain more reliable estimates of the sampling distributions of trends in $A'_e(t)$.

Figure 4 shows the scatter of trends in the difference field projection $A'_e(t)$ as a function of block length and season. For each block length, each computed trend in $A'_e(t)$ is represented by a dot, so that there are 41 dots for 5-yr trends, 40 dots for 6-yr trends, etc.

We next calculated the ''signal trends'' of interest—that is, the trends in the $A_1(t)$ projection time series for the GA experiment. These are the trends whose signif-

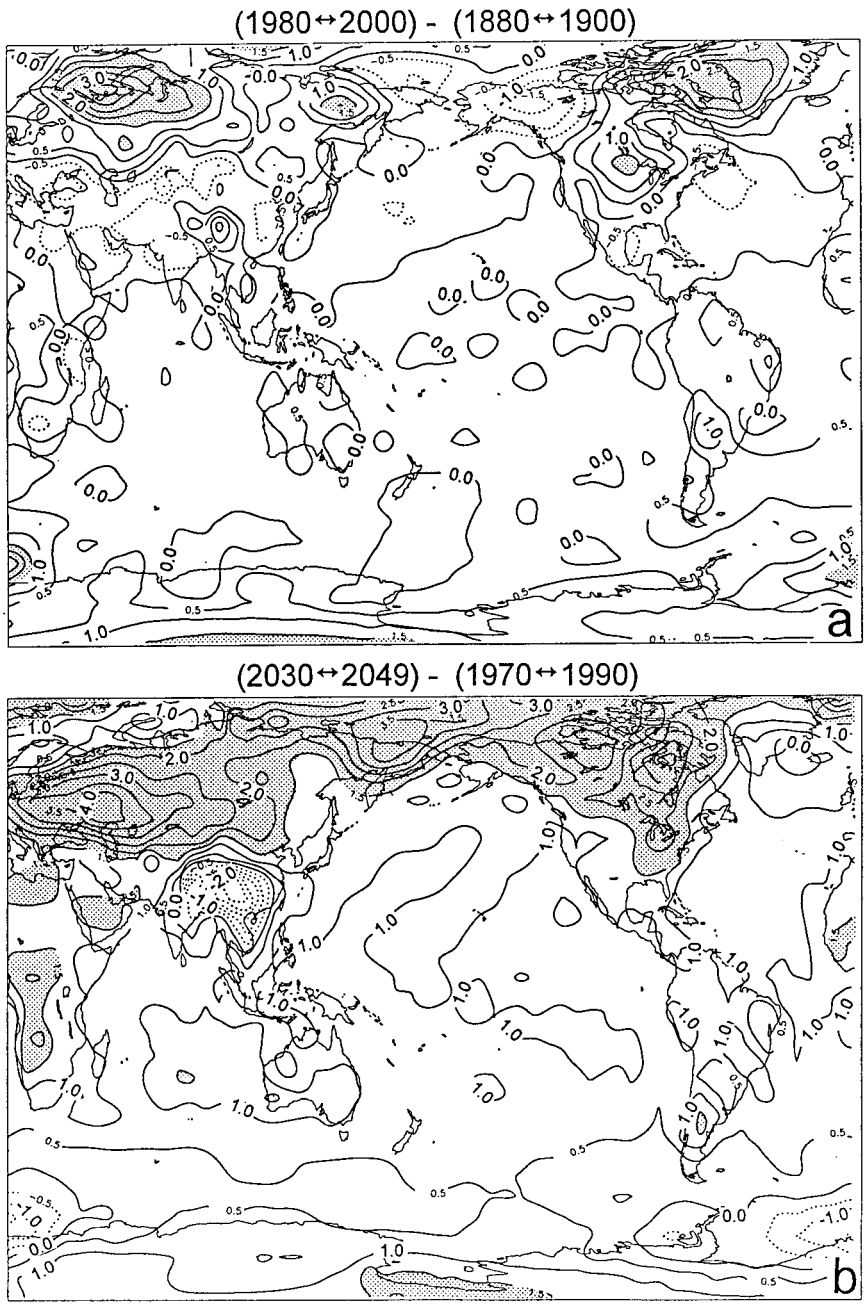


FIG. 2a. The DJF 2-m temperature change signal (°C) projected by HAM3L in response to combined changes in equivalent CO_2 and anthropogenic sulfate aerosols. The signal is given as the difference in mean states over two time periods: 1980–2000 minus 1880–1900 (upper panel), and 2030–49 minus 1970–90 (lower panel). Positive values greater than 1.5°C have been stippled.

icance we would like to assess. We first selected four arbitrary base years, 1990, 2000, 2010, and 2020, and then computed trends for data blocks extending backward from the base years. For example, the trend over a 5-yr period from the base year 1990 was computed with $A_1(t)$ data for the years 1986–90. This value was then plotted on Fig. 4 (as a plus sign) at the 5-yr block length coordinate. The 5-yr trends depend on the selected base year, so they were also estimated and entered on Fig. 4 for the base years 2000, 2010, and 2020. This was done for trend estimates for all possible block lengths between 5 and 30 yr. The point at which the signal trends in $A_1(t)$ emerge from the funnel-shaped swarm of trends in $A'_e(t)$ is a good indication of the length of record one needs to use in trend estimation to avoid mistaking temperature changes associated with

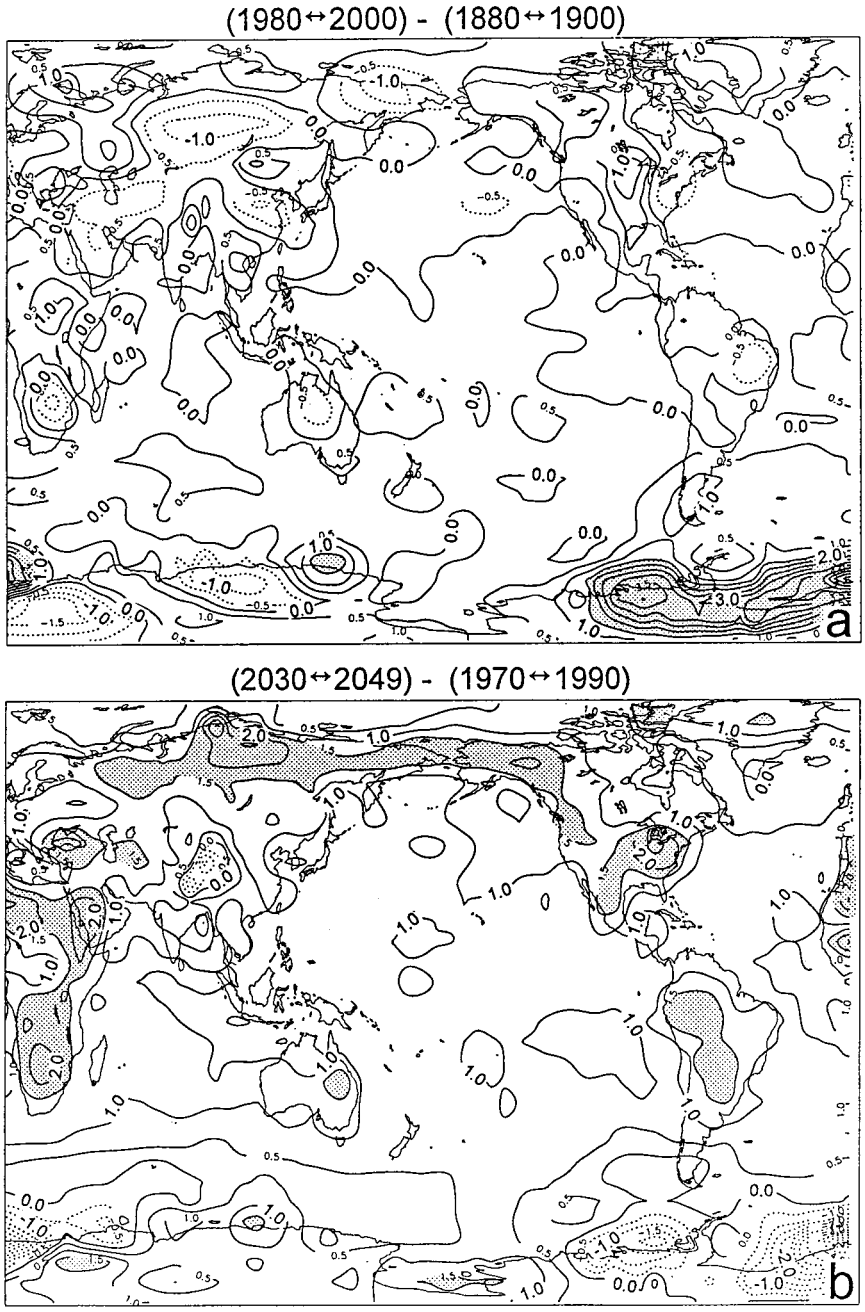
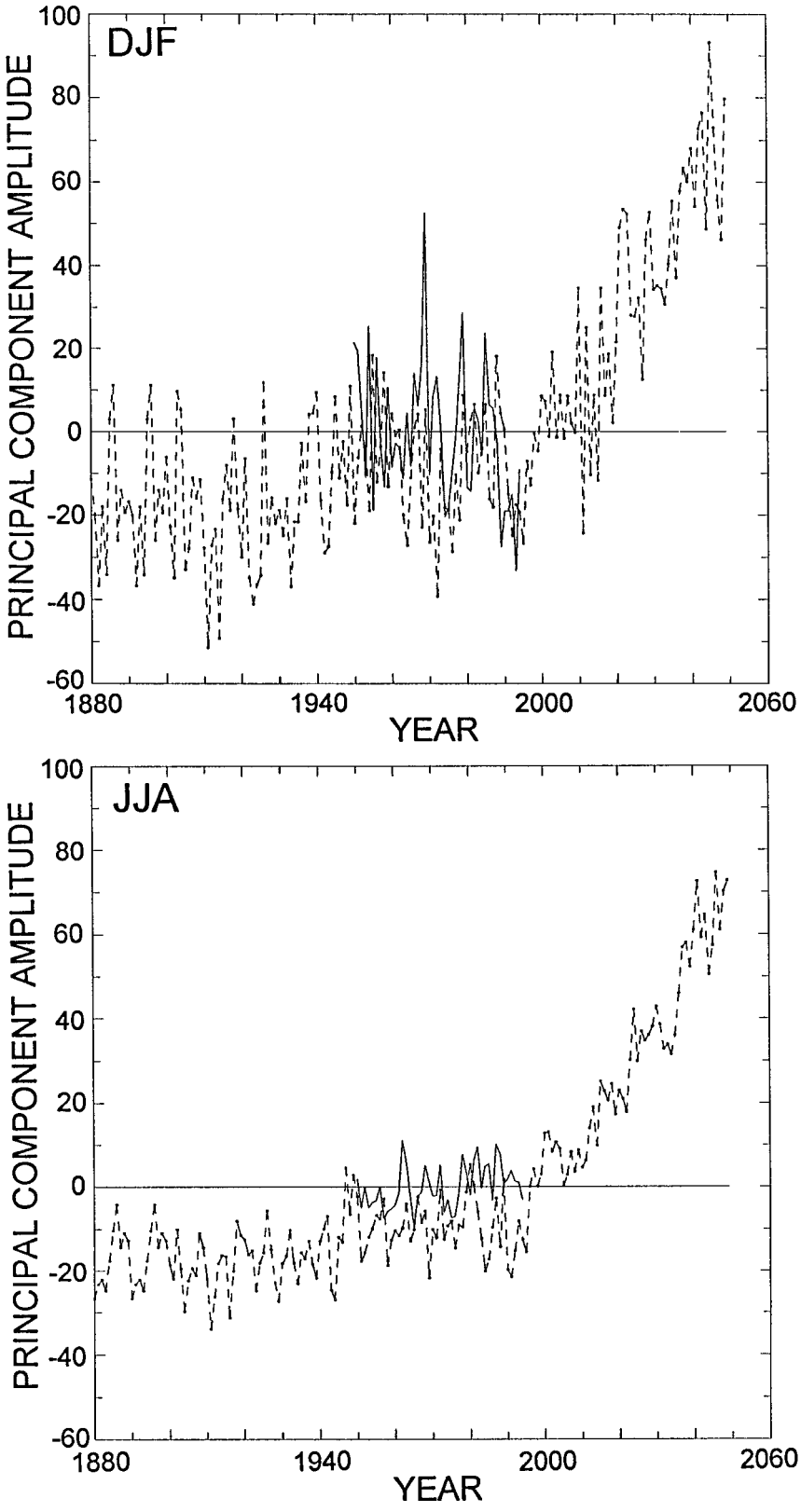


FIG. 2b. Same as Fig. 2a except for JJA.

model minus observed differences for temperature changes due to anthropogenic effects.

Examination of Fig. 4 yields three conclusions. First, the magnitude of trends in the difference field projection $A'_e(t)$ generally decrease with increasing block length. This behavior is similar to that found for trends due to internal variability alone (Santer et al. 1995a). For the base years 1990 and 2000, the signal trends in $A_1(t)$ also tend to decrease with increasing block length, primarily due to the large interannual variability in $A_1(t)$ (making trend estimates sensitive to beginning and end points) and the fact that $A_1(t)$ does not increase dramatically until after the year 2000. For later base years (2010, 2020), the signals are larger and trends in $A_1(t)$ may even increase as block length increases.

Second, Fig. 4 suggests that data records of at least 20–30 yr in length are required in order to obtain anthropogenic signal trends that can be discriminated from trends due to the combined effects of model forcing and response errors, internal variability, and observational



uncertainty. In some cases (signals estimated from base years of 1990 and 2000 in JJA), even 30-yr data records are not adequate to distinguish between signal and ''difference field'' trends. Since our RC run is only 45 yr long, the statistics derived from it, that is, the $A'_e(t)$ trend estimates, do not support investigation of averaging periods longer than about 30 yr. Longer RC simulations would be required to remove this shortcoming.

Third, the results of Fig. 4 show a perplexing problem: some of the signal trends from the GA run emerge from the difference field noise envelope and then, for longer averaging periods, dive back into the noise, for example, the trends based on the base year 2000 during JJA. Again, this shows that for relatively short timescale trend estimates (5–30 yr), results are sensitive to the choice of base period for defining the signal and to the large interannual-to-decadal timescale variability in the model RC and GA experiments. This illustrates the need for caution to avoid premature claims of positive detection of an anthropogenic signal. From a purely statistical point of view, the test of presence/absence of the anthropogenic signal is made at some point in time. The test is either satisfied at some significance level (say 95%) or it is not. If the test is satisfied, the signal is said to be present. If the same test is conducted at many different points in time, then we would eventually expect it to be satisfied simply by chance. Given that we are working with trends estimated from moving blocks of data, the chance significance would happen in a number of successive years. We will see an example of this below. For the present, the results of Fig. 4 demand that ''practical'' detection requires that the significance threshold be succeeded consistently for some period of time. During this interval, one should expect the signal significance to *increase* with time, eventually obtaining values so unusual they cannot reasonably be ascribed to chance.

c. HAM3L projection: Another view

The problem can be viewed in a slightly different perspective by comparing the time variations in GA anthropogenic trend estimates [i.e., in $A_1(t)$] with the trend estimates derived from the model-observed difference field (Fig. 5). Data blocks of length 10, 20, and 30 yr were allowed to slide down the $A_1(t)$ signal time series (Fig. 3) in 1-yr increments, commencing in the year 1880. The trend of the data in each block was estimated and plotted in Fig. 5 on the last year of data in the current block. For example, the 10-yr trend in $A_1(t)$ over 1880–89 was assigned to the year 1889, while the next 10-yr trend (over 1881–90) was plotted on the year 1890. In this way, all possible signal trends of lengths 10, 20, and 30 yr are considered over the full period of the GA experiment, an approach similar to that used by Hegerl et al. (1996). This has the advantage of yielding results that do not depend (as in Santer et al. 1996b) on some arbitrary choice of base year for computing signal trends.

The time series of trend values for the three different block lengths and summer/winter seasons are given in Fig. 5. The large internal variability within the GA run is evident, particularly for block sizes of 10 yr. Also shown in Fig. 5, as solid lines that have no time dependence, are the 10-, 20-, and 30-yr maximum trend estimates derived from $A'_e(t)$, the projection of the difference field onto the dominant GA response pattern. These trend values correspond to the extreme values of the ''dots'' in Fig. 4, for block sizes of 10, 20, and 30 yr and as such need not be symmetric about zero.

In the following, we define detection time as the date at which the trend in $A_1(t)$ first exceeds and then remains continuously above the ''significance threshold,'' which for our purposes is the maximum 10-, 20-, or 30-yr trend in $A'_e(t)$. We stress that our threshold incorporates the combined effects of model forcing and response errors, natural internal variability, and observational uncertainty, while previous detection studies have used significance thresholds that reflect trend fluctuations due to natural variability alone (e.g., Hegerl et al. 1996; Santer et al. 1996b; Tett et al. 1996).

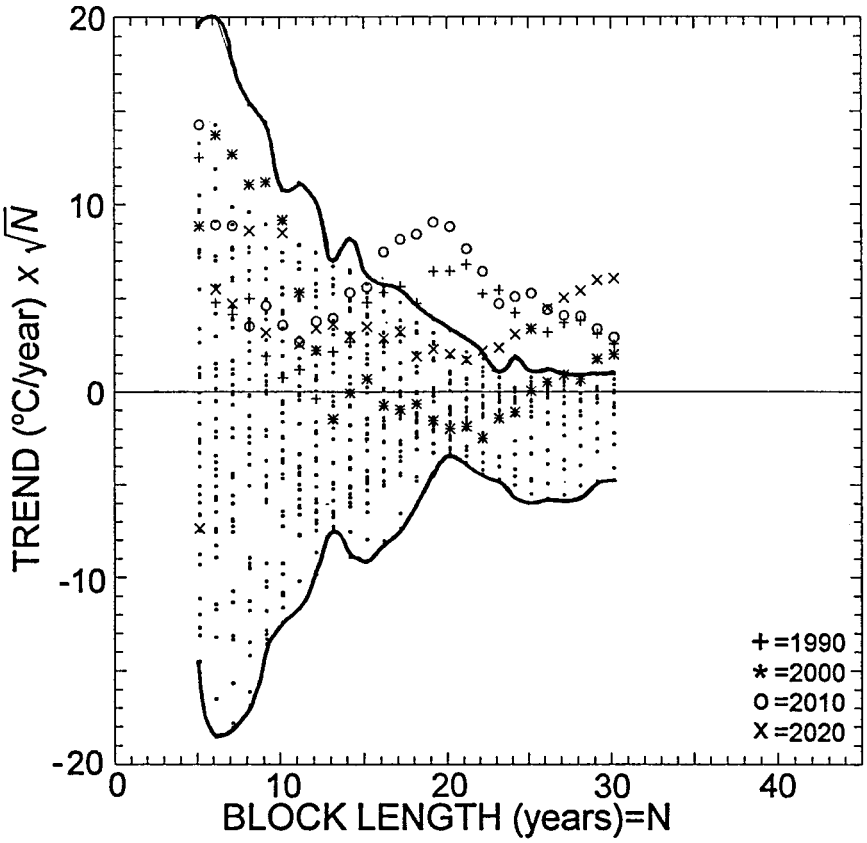
Examination of Fig. 5 leads to the following conclusions.

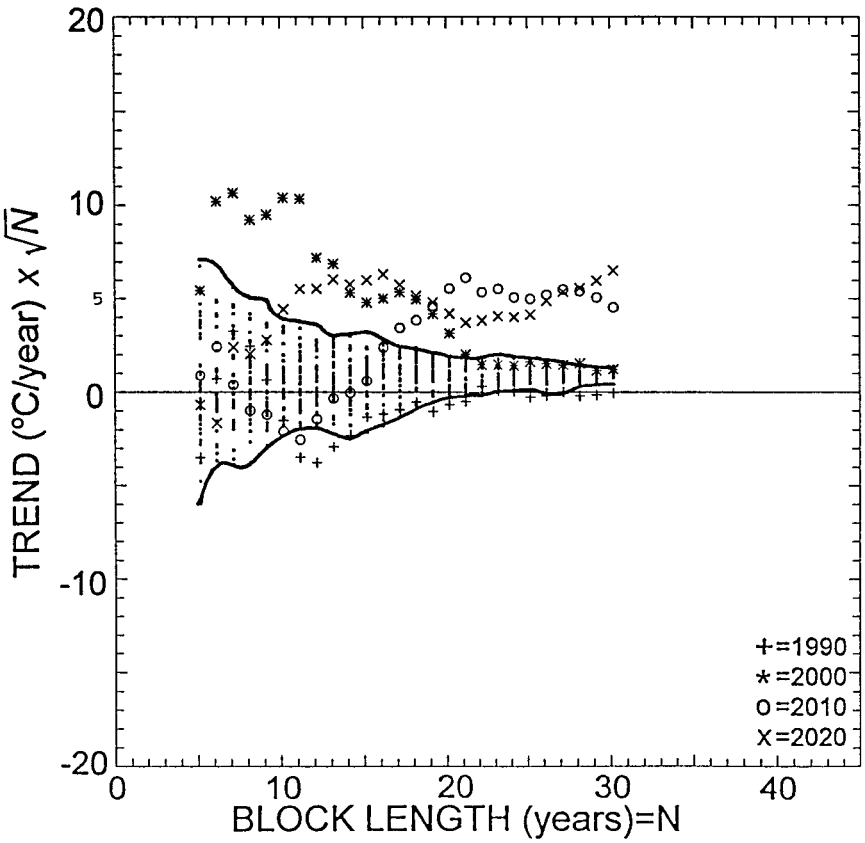
(i) Even out to the year 2050, anthropogenic signal trends for block sizes of only 10 yr cannot be clearly distinguished from difference field trends. While there are times when the 10-yr trends in $A_1(t)$ exceed the significance threshold, the final trend values are *below* this threshold. This highlights the difficulty of using very short record lengths to detect an anthropogenic signal (Christy and McNider 1994) and relates to results for 15-yr trends in Hegerl et al. (1996).

(ii) For signal trends of 20 and 30 yr in length, we *can* identify times at which the signal trend exceeds and remains above the maximum difference field trend. For DJF data, the first ''detectable'' 20- and 30-yr signal trends are for the periods 2002–21 and 1979–2008, respectively. Signal-to-noise (S/N) characteristics are more favorable in JJA, in part due to the smaller difference field trends in this season of small noise: the

←

FIG. 3. Evolution of the projection time series $A'_e(t)$ (solid line) and $A_1(t)$ (dashed line). $A'_e(t)$ is the projection of the model minus observed 2-m temperature difference field, $e(\mathbf{x}, t)$, onto the first EOF of the HAM3L transient experiment with combined GHG + aerosol forcing (GA). The principal component time series $A_1(t)$ is the projection of the 2-m temperature signal in GA onto the first EOF of GA. Results are given for DJF (upper panel) and JJA (lower panel). Note that the mean of $A'_e(t)$ was adjusted to be identical to the mean of $A_1(t)$ over 1950–94.





first 20- and 30-yr signal trends to exceed and remain continuously above our stipulated noise threshold are now over 1981–2000 and 1972–2001. These seasonal differences in S/N are similar to those found in other investigations of near-surface temperature by Santer et al. (1995b) and Hegerl et al. (1997), albeit for rather different ''noise.''

(iii) Despite fundamental differences in the nature of the ''noise threshold'' used in the present study and in these earlier studies, the qualitative picture of an anthropogenic GA signal produced by HAM3L gradually rising above the noise threshold is very similar.

(iv) Another surprising correspondence between the present study and earlier work of Hegerl et al. (1996b) relates to the finding of a transient ''bump'' in S/N around the middle of this century (see Hegerl et al. 1996b). This bump marks an extended period of time when the signal trends (even for 20- to 30-yr block lengths) exceed the difference field threshold and then plunge beneath it. Whatever the cause of the bump, it is just the type of ''chance'' detection scenario discussed above. It shows the need for some caution before claiming positive detection of anthropogenic effects. We note, however, that the duration of the S/N bump in the mid-twentieth century is much shorter (by a factor 3–5) than the 30- to 50-yr period toward the end of the GA experiment, when the longer timescale GA signal trends remain continuously above the difference field threshold and the signal ratio (statistical significance) is growing with time. This is an excellent example of what we referred to above as ''practical'' detection.

In summary, if the GA signal used here is a reliable estimate of the near-surface temperature response to combined GHG + aerosol forcing, it should be possible to distinguish this signal from the background noise generated by model forcing and response errors, internal variability, and observational uncertainty. In the most favorable cases (20- and 30-yr trends in JJA), signal detection should be feasible now or within the next decade. As noted above, this estimate depends on how one defines the significance threshold and our definition is different than those used previously. For signal trends of only 10 yr in length, the prospects of discriminating between changes due to anthropogenic and nonanthropogenic effects are poor. Finally, it is likely that natural variability effects can lead to the detection metric exceeding, purely by chance, a detection threshold, sometimes for a duration of up to 10 yr. Such behavior could lead to erroneous claims of detection of anthropogenic effects, a possibility that seems unlikely in the final 50–75 yr of the GA experiment considered here. The results clearly demonstrate the difference between signal detection based on purely statistical grounds and that of practical utility.

5. Observational errors and spatial resolution

As noted in section 1, the model minus observed difference fields computed here have many components, one of which is observational uncertainty. Estimating the possible size and nature of errors in the observations is an important task that is only now beginning to be addressed (Jones et al. 1997). Here, we briefly consider the effect of observational errors on our results. We also investigate whether the results presented in section 4 are sensitive to the spatial resolution at which the analysis is performed.

The ''observational error'' field was computed by differencing the radiosonde station values of surface air temperature (see section 2a) and those from the CRU dataset, after interpolation of the CRU data to the radiosonde station locations. These differences, available for 53 spatial points, are denoted by $d(\mathbf{x}, t)$. Note that the act of subtraction likely leaves the d-field free of any anthropogenic signal that may have been in these datasets.

An inspection of the data showed some systematic errors, for example, shift of means, prior to 1974 in the radiosondes and so we ignore data prior to that time. The remaining information was subjected to a variety of analyses (most of which will not be reported on here). The standard deviation of the observational error field is shown in Fig. 6. Typical values are of order 0.5°C in the lower latitudes, increasing to order 1.0°C or higher toward the poles. The errors are clearly smaller in JJA than in DJF.

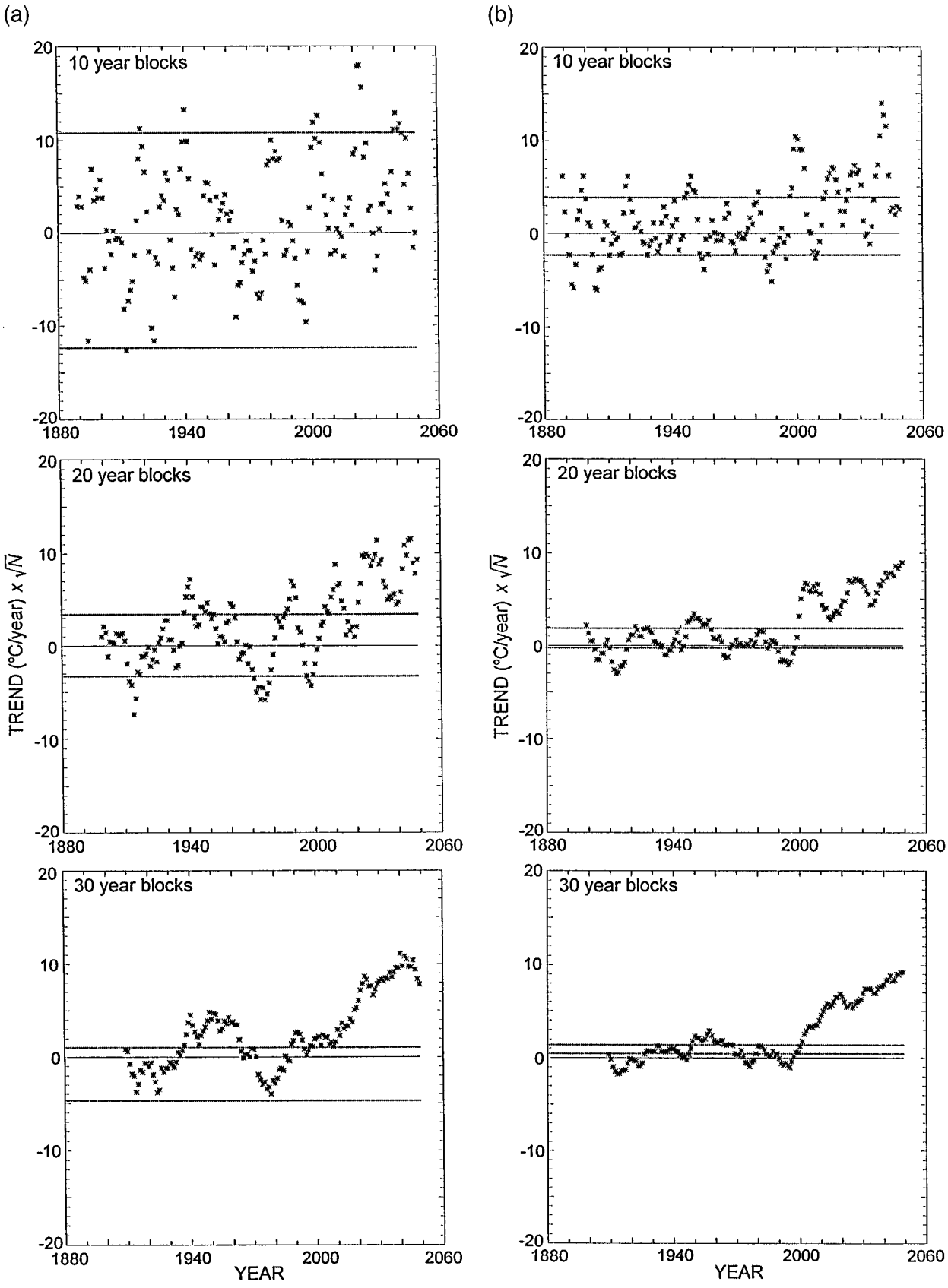
An EOF analysis (not shown) revealed there was some spatial correlation in the observational error field, especially over the Northern Hemisphere continents. The potential impact of this error characteristic was tested by projecting $d(\mathbf{x}, t)$ onto the HAM3L GA signal:

$$A'_d(t) = \int_{\mathbf{x}} d(\mathbf{x}, t) f_1(\mathbf{x}) \cos[\theta(\mathbf{x})]\, d\mathbf{x}. \qquad (3)$$

This yields a projection time series, $A'_d(t)$, analogous to the $A'_e(t)$ time series in Fig. 3. The observational error did not project well onto the anthropogenic signal: the

←

FIG. 4. Comparison of rates of changes in the difference field projections $A'_e(t)$ and signal projections $A_1(t)$ shown in Fig. 3. Rates of change in $A'_e(t)$ were estimated by fitting least squares linear trends to overlapping blocks of the time series. Block lengths ranged from 5 to 30 yr, in increments of 1 yr. Each dot represents an individual trend estimate derived from $A'_e(t)$, and the heavy solid lines enclose the region occupied by all difference trend values. Signal trends were estimated in the following way: after selecting four ''base years'' (1990, 2000, 2010, and 2020), linear trends of varying block length were computed from the $A_1(t)$ data, working backward in time from the base year. Thus the ''+'' sign entered at a block length of 5 yr represents an $A_1(t)$ trend over 1986–90, while the ''*'' at the same block length denotes a trend over 1996–2000, etc. Results are shown for DJF (upper panel) and JJA (lower panel).



magnitude of fluctuations in $A'_a(t)$ was a factor of approximately 7–10 below that of changes in either the signal or difference field projections.

In summary, the errors in the observations do not project well onto the predicted GA signals. Hence, they may not be a major hindrance in the detection process. However, the results discussed above were based on a relatively small time sample, thus precluding analysis of observational uncertainty effects on multidecadal timescales. They also ignored the effects of any systematic spatial and temporal biases that may be *common* to the radiosonde data and the CRU observations.

Would the primary conclusions of section 4 have changed if we had performed our entire analysis at the limited spatial resolution of the Angell radiosonde network? To address this question, we recomputed $A'_e(t)$ in Eq. (2), using an $e(\mathbf{x}, t)$ field calculated after interpolating the RC and CRU data to the spatial locations of the radiosonde stations. The same interpolation was performed before projecting the model signal, $s(\mathbf{x}, t)$, onto the dominant GA response pattern $f_1(\mathbf{x})$. Figure 7 shows the analysis of the resulting $A'_e(t)$ and $A_1(t)$ projections for 30-yr trends in DJF and JJA. Direct comparison of Fig. 7 with the lower panels in Figs. 5a and 5b isolates the impact of the reduced spatial resolution. It is clear from this that the results *are* sensitive to the degraded spatial resolution—the 30-yr signal trends are now less-easily distinguishable from the ''difference field'' trends than they were in the case of the high-resolution spatial fields. The same applies to the 20-yr signal trends (not shown).

This result has several interpretations. First, the radiosonde network, although probably adequate for capturing the structure of a large-scale GHG signal, is suboptimal for representing important features of a smaller-scale GHG + aerosol signal. Second, the spatial noise in the signal estimates derived from the network is strongly reduced by the inclusion of information from many thousands of (highly autocorrelated) grid points. Finally, the Angell radiosonde network is predominantly ''land only,'' whereas the CRU gridded near-surface temperature data have a higher proportion of ocean points than land points. Since the RC minus observed temperature differences over ocean areas are expected to be very small, the $e(\mathbf{x}, t)$ field computed with the CRU data is expected to have proportionally more values that are zero or very small than the $e(\mathbf{x}, t)$ field based on the Angell data. This has consequences for the relative sizes of difference field trends in the CRU-based and Angell-based $A'_e(t)$ projections. In other words, we would expect the difference field significance thresholds in Figs. 5a and 5b (lower panels) to be proportionately *lower* than the thresholds shown in Fig. 7. Assuming that the GA signal is equally well resolved in the coarse-resolution Angell data and fine-resolution CRU data, S/N should be more favorable if $e(\mathbf{x}, t)$ is computed with the CRU temperatures.

Our results therefore suggest that for this particular variable, increasing spatial resolution may help rather than hinder early detection of anthropogenic effects.

6. Conclusions

Previous detection studies have considered only the effects of natural climatic variability in attempting to assess whether an anthropogenic climate-change signal can be detected in observed records of temperature change. Our investigation is one of the first to consider not only natural variability, but also the effects of model errors and observational uncertainty. The noise that we consider here is therefore very different from that used in previous detection work.

The present study has focused on the model minus observed near-surface temperature ''difference structures,'' $e(\mathbf{x}, t)$, of a state-of-the-art atmospheric GCM (ECHAM3 from the MPI, Hamburg). The space–time properties of these difference structures, as estimated here, are influenced by missing or imperfectly represented natural and anthropogenic forcing, true model errors, internally generated model variability and errors in the observations themselves. We note that in our analysis, missing or inaccurate forcing and model errors primarily affect $e(\mathbf{x}, t)$, over land areas only. This is due to the differencing of the highly similar observed and simulated 2-m temperature datasets: the simulated 2-m temperature in the recent climate experiment is forced to be very similar to the 2-m temperature in the observations, since the RC experiment uses observed time-varying SSTs as a lower boundary condition. For the same reason, the effects of internally generated natural variability of the coupled atmosphere–ocean system are manifest in $e(\mathbf{x}, t)$, over land areas only. Note that in data from the anthropogenic signal experiments considered here, the effects of natural internal variability are manifest over *both* land and sea areas.

We have shown that $e(\mathbf{x}, t)$ has some temporal and spatial coherence. There are features of the difference structures that resemble the patterns of temperature

←

FIG. 5. (a) Trends in DJF anthropogenic signal strength as functions of time and block length. Results are shown for three different block lengths: 10, 20, and 30 yr (top, middle, and lower panels, respectively). The signal trends were computed from overlapping blocks of the $A_1(t)$ time series shown in Fig. 3, that is, from the projection of the 2-m temperature signal in GA onto the first EOF of GA. Signal trends are plotted on the final year of the overlapping block: for example, the first time point in the upper panel is plotted on 1889 and gives the trend in $A_1(t)$ over 1880–89, while the second time point in the upper panel gives the $A_1(t)$ trend over 1881–90, etc. The solid lines represent the maximum/minimum values of the $A'_e(t)$ trends estimated from the difference field projection onto the GA signal. (b) Same as 5a but for JJA.

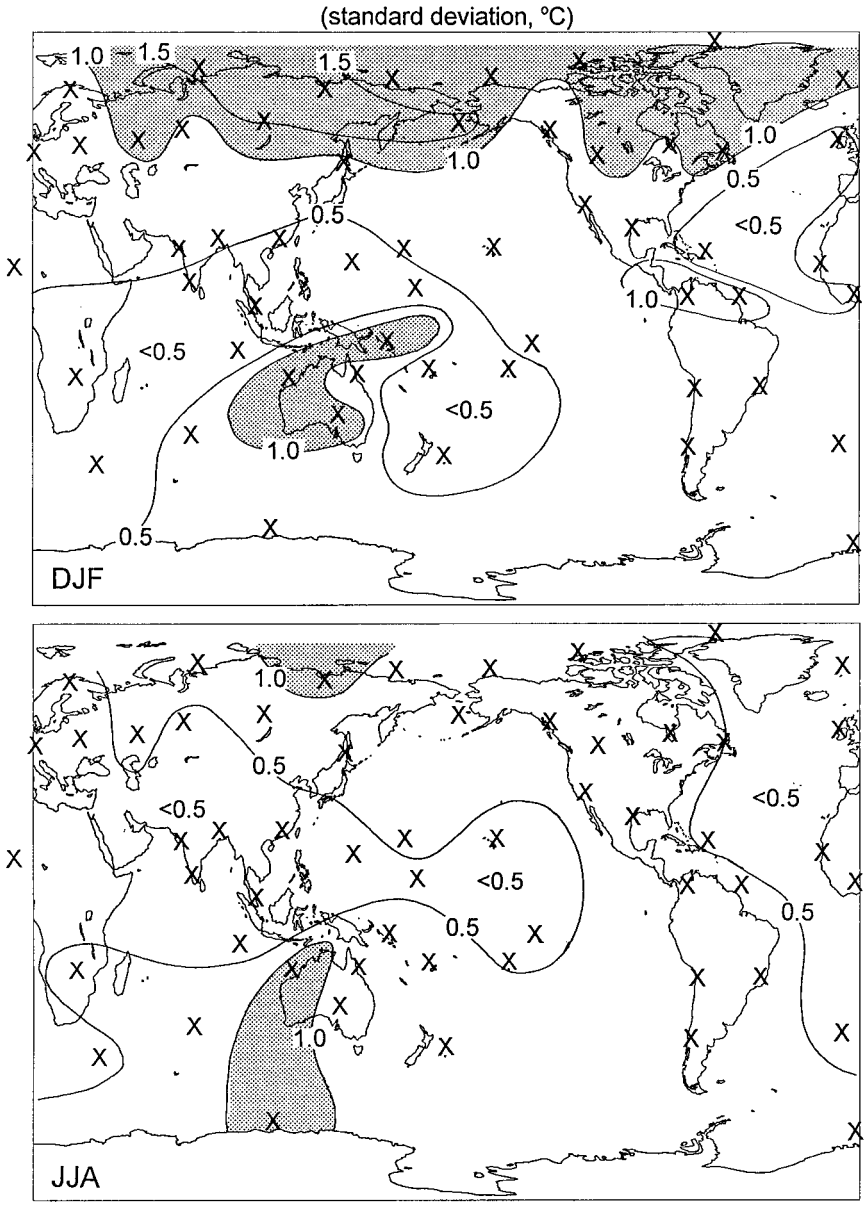


FIG. 6. Standard deviation of the difference, $d(\mathbf{x}, t)$, between two independent estimates of observed near-surface air temperature. Regions where values exceed 1.0°C are stippled. The heavy "×" denotes the location of radiosonde stations in the Angell network.

change predicted by a fully coupled version of the MPI model in response to combined changes in greenhouse gases (GHGs) and anthropogenic sulfate aerosols. On timescales of a decade or so, the difference structures that we computed project reasonably well onto the GHG+aerosol pattern predicted by the coupled model. On timescales of 2–3 decades, however, trends in this difference field projection are generally smaller than trends in the signal itself for time periods beyond 2000–2010.

These results suggest that it should, in principle, be feasible to separate a human-induced climate change signal from the type of noise considered here, at least if one considers timescales of 20 yr or greater for representing signal and noise, and if the signal is defined from a suitably late initial time (within the next several decades). In the most favorable cases—for 20- and 30-yr trends in JJA—we find that signal trends exceed and then remain continuously above our noise threshold starting in 1981–2000 (for 20-yr trends) and in 1972–2001 (for 30-yr trends). However, if signal and "noise" are considered over only 10-yr blocks of time, we would be unable to distinguish between them, not only at present, but also for a much larger signal 50–60 yr in the

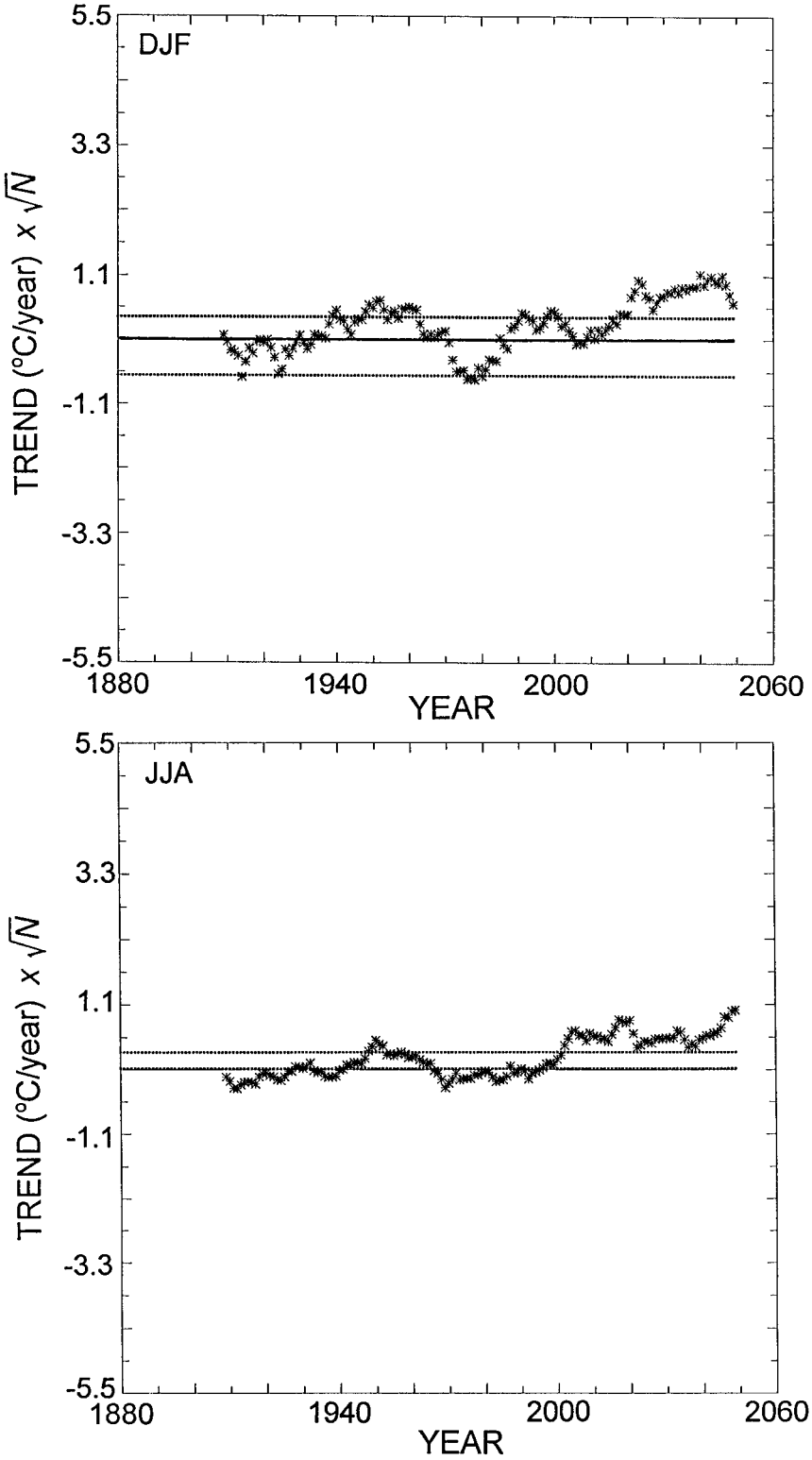


FIG. 7. Trends in DJF and JJA anthropogenic signal strength as functions of time. Results are for 30-yr trend lengths. This figure is directly comparable to the lower panels in Figs. 5a and 5b. The difference is that the analysis in this figure was performed at the coarse spatial resolution of the 63-station radiosonde network, while the analysis in Fig. 5 was carried out at very fine spatial resolution (see section 5 for further details).

future! This means that attempts to detect anthropogenic signals in surface temperature must focus on trends or averaging periods that substantially exceed decadal timescales. The recent detection studies of Santer et al. (1995b), Santer et al. (1996b), Hegerl et al. (1996), and Hegerl et al. (1997) satisfy this requirement by dealing primarily with temperature trends over 20–50 yr. This does not guarantee that these investigations are uninfluenced by low-frequency model errors, since the latter quantity is not determined by this study and is currently unknown.

Analysis of the HAM3L transient GHG + aerosol simulation showed that the natural variability in that run introduces uncertainties in estimates of detection time, in accord with earlier findings by Cubasch et al. (1994) and Santer et al. (1995a). In the current investigation, natural internal variability can manifest itself as periods when the anthropogenic signal exceeds the maximum difference field trend for a decade or more, for example, over 1945–55 in Fig. 5b (middle panel). Over this time the anthropogenic signal is clearly different from or ''unique'' relative to the difference field. But this period of uniqueness is transient—subsequent signal trends return to a level *below* the maximum difference field threshold for several decades and so are indistinguishable from it. If exceeding the difference field maximum threshold used here had been the only criterion for claiming detection of an anthropogenic signal, erroneous detection of anthropogenic effects would have been claimed in the 1940s. A similar example can be found in the recent detection study of Hegerl et al. (1996) where their chosen detection metric exceeds a significance threshold for some years in the middle of this century, only to fall back below the threshold until most recent times.

This points toward the need for some caution in the interpretation of detection results. It is clearly important to establish that a signal has to be significantly different from some noise threshold for some time, during which its significance ought to be increasing. This condition must be met before we can conclude that we have, for practical purposes, identified a signal and not simply a manifestation of natural variability. In our study, the Hamburg model results suggest that 20- to 30-yr signal trends should within the next few years to decade exceed and remain continuously above our significance threshold; a threshold that is different than used in previous detection studies. But our results also suggest that the signal trend must stay above the threshold for order of a decade before we should begin to believe an anthropogenic signal has indeed been detected for practical decision-making.

The purpose of this preliminary study has been to demonstrate methodology and explore how the combined effect of model errors, internal variability, and observational uncertainty might hamper attempts to detect an anthropogenic signal. Our results show that model-observed difference structures can be an important consideration in the detection problem, particularly on decadal timescales. Future work should focus on longer timescale model-observed differences and on better partitioning of the components of these difference structures.

Acknowledgments. This work was partially supported by NOAA's Office of Global Program's Climate Change Data and Detection Program Element, NSF Grant ATM93-14495, and the Scripps Institution of Oceanography (TPB). The EEC Environmental Program Contract ENV-CT95-0102 and the Max Planck Gesellschaft provided additional support for GCH. We thank Uli Cubasch, Phil Jones, Reinhard Voss, and Francis Zwiers for constructive discussions and suggestions on various aspects of the manuscript. We also gratefully acknowledge the use of the global air temperature dataset from the Climate Research Unit (courtesy of Phil Jones) and use of Version 2.2 of the Global Sea-Ice and Sea Surface Temperature dataset, 1903–94 by N. A. Rayner, E. B. Horton, D. E. Parker, C. K. Folland, and R. B. Hackett of the Hadley Centre for Climate Prediction and Research Meteorological Office, London Road, Bracknell, Berkshire, United Kingdom.

REFERENCES

Angell, J., and J. Korshover, 1983: Global temperature variations in the troposphere and stratosphere, 1958–1982. *Mon. Wea Rev.,* **111,** 901–921.

Barnett, T. P., 1986: Detection of changes in the global troposphere temperature field induced by greenhouse gases. *J. Geophys. Res.,* **91,** 6659–6667.

——, 1991: An attempt to detect the greenhouse-gas signal in a transient GCM simulation. *Greenhouse-Gas-Induced Climatic Change: A Critical Appraisal of Simulations and Observations,* M. E. Schlesinger, Ed., Elsevier, 559–568.

——, 1995: Monte Carlo climate forecasting. *J. Climate,* **8,** 1005–1022.

——, and M. E. Schlesinger, 1987: Detecting changes in global climate induced by greenhouse gases. *J. Geophys. Res.,* **92** (D12), 14 722–14 780.

——, ——, and X. Jiang, 1991: On greenhouse gas signal detection strategies. *Greenhouse-Gas-Induced Climatic Change: A Critical Appraisal of Simulations and Observations,* M. E. Schlesinger, Ed., Elsevier, 537–558.

——, N. Graham, M. Latif, S. Pazan, and W. White, 1993: ENSO and ENSO-related predictability. Part I: Prediction of equatorial pacific sea surface temperature with a hybrid coupled ocean–atmosphere model. *J. Climate,* **6,** 1545–1566.

——, K. Arpe, L. Bengtsson, M. Ji, and A. Kumar, 1997: Potential predictability and AMIP implications of midlatitude climate variability in two general circulation models. *J. Climate,* **10,** 2321–2329.

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

Bengtsson, L., K. Arpe, E. Roeckner, and U. Schulzweida, 1996: Climate predictability experiments with a general circulation model. *Climate Dyn.,* **12,** 261–278.

Christy, J. R., and R. T. McNider, 1994: Satellite greenhouse signal. *Nature,* **367,** 325.

Cubasch, U., K. Hasselmann, H. Höck, E. Maier-Reimer, U. Mikolajewicz, B. Santer, and R. Sausen, 1992: Time-dependent greenhouse warming computations with a coupled ocean–atmosphere model. *Climate Dyn.,* **8,** 55–69.

——, and Coauthors, 1994: Monte Carlo climate forecasts with a global coupled ocean–atmosphere model. *Climate Dyn.,* **10,** 1–19.

——, G. C. Hegerl, R. Voss, J. Waszkewitz, and T. J. Crowley, 1997: Simulation of the influence of solar radiation variations on the global climate with an ocean–atmosphere general circulation model. *Climate Dyn.,* **13,** 757–767.

Gates, W. L., 1992: AMIP: The Atmospheric Model Intercomparison Project. *Bull. Amer. Meteor. Soc.,* **73,** 1962–1970.

Graham, N. E., T. P. Barnett, R. Wilde, U. Schlese, and L. Bengtsson, 1994: On the roles of tropical and midlatitude SSTs in forcing interannual to interdecadal variability in the winter Northern Hemisphere circulation. *J. Climate,* **7,** 1416–1441.

Hasselmann, K., 1979: On the signal-to-noise problem in atmospheric response studies. *Meteorology Over the Tropical Ocean,* D. B. 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,** 601–611.

——, and Coauthors, 1995: Detection of anthropogenic climate change using a fingerprint method. Max-Planck-Institut für Meteorologie Rep. 168, 20 pp. [Available from MPI, Bundesstr. 55, D-20146 Hamburg, Germany.]

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

——, H. von Storch, K. Hasselmann, B. D. Santer, U. Cubasch, and P. D. Jones, 1996: Detecting greenhouse gas-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: On multi-fingerprint detection and attribution of greenhouse gas- and aerosol forced climate change. *Climate Dyn.,* **13,** 613–639.

Jones, P. D., and K. Briffa, 1992: Global surface air temperature variations over the twentieth century. Part 1: Spatial, temporal, and seasonal details. *Holocene,* **2,** 165–179.

——, T. J. Osborn, and K. R. Briffa, 1997: Estimating sampling errors in large-scale temperature changes. *J. Climate,* **10,** 2548–2568.

Karoly, D. J., J. A. Cohen, G. A. Meehl, J. F. B. Mitchell, A. H. Oort, R. J. Stouffer, and R. T. Wetherald, 1994: An example of fingerprint detection of greenhouse climate change. *Climate Dyn.,* **10,** 97–105.

Kattenberg, A., and Coauthors, 1996: Climate models—Projections of future climate. *Climate Change 1995: The Science of Climate Change,* J. T. Houghton, L. G. Meira Filho, B. A. Callander, N. Harris, A. Kattenberg, and K. Maskell, Eds., Cambridge University Press, 285–357.

Maier-Reimer, E., U. Mikolajewicz, and K. Hasselmann, 1993: Mean circulation of the Hamburg LSG OGCM and its sensitivity to the thermohaline surface forcing. *J. Phys. Oceanogr.,* **23,** 731–757.

Manabe, S., and R. J. Stouffer, 1996: Low frequency variability of surface air temperature in a 1000-year integration of a coupled ocean–atmosphere model. *J. Climate,* **9,** 376–393.

Nicholls, N., G. V. Gruza, J. Jouzel, T. R. Karl, L. A. Ogallo, and D. E. Parker, 1996: Observed climate variability and change. *Climate Change 1995: The Science of Climate,* B. A. Callander, N. Harris, A. Kattenberg, and K. Maskell, Eds., Cambridge University Press, 133–192.

North, G. R., and K. Y. Kim, 1995: Detection of forced climate signals. Part II: Simulation results. *J. Climate,* **8,** 409–417.

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

——, K. Y. Kim, S. P. Shen, and J. W. Hardin, 1995: Detection of forced climate signals. Part I: Filter theory. *J. Climate,* **8,** 401–408.

Roeckner, E., and Coauthors, 1992: Simulation of the present-day climate with the ECHAM model: Impact of model physics and resolution. Max-Planck-Institut für Meteorologie Rep. 93, 171 pp. [Available from MPI, Bundesstr. 55, D-20146 Hamburg, Germany.]

Santer, B. D., T. M. L. Wigley, and P. D. Jones, 1993: Correlation methods in fingerprint detection studies. *Climate Dyn.,* **8,** 265–276.

——, U. Mikolajewicz, W. Brueggemann, U. Cubasch, K. Hasselmann, H. Hoeck, E. Maier-Reimer, and T. M. L. Wigley, 1995a: Ocean variability and its influence on the detectability of greenhouse warming signals. *J. Geophys. Res.,* **100,** 10 693–10 725.

——, K. E. Taylor, T. M. L. Wigley, J. E. Penner, P. D. Jones, and U. Cubasch, 1995b: Towards the detection and attribution of an anthropogenic effect on climate. *Climate Dyn.,* **12,** 77–100.

——, T. M. L. Wigley, T. P. Barnett, and E. Anyamba, 1996a: Detection of climate change and attribution of causes. *Climate Change 1995: The IPCC Second Scientific Assessment,* J. T. Houghton and B. A. Callander, Eds., Cambridge University Press, 362 pp.

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

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

Trenberth, K. E., and J. G. Olson, 1991: Representativeness of a 63-station network for depicting climate changes. *Greenhouse-Gas-Induced Climatic Change: A Critical Appraisal of Simulations and Observations,* M. E. Schlesinger, Ed., Elsevier, 249–259.

Voss, R., and R. Sausen, 1996: Techniques of asynchronous and periodically synchronous coupling of atmosphere and ocean general circulation models 2. Impact of variability. *Climate Dyn.,* **12,** 605–614.

Zwiers, F. W., 1996: Interannual variability and predictability in an ensemble of AMIP climate simulations conducted with the CCC GCM2. *Climate Dyn.,* **12,** 825–847.