THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Detectability of anthropogenic changes in annual temperature and precipitation extremes

Citation for published version:
Hegerl, GC, Zwiers, FW, Stott, PA & Kharin, VV 2004, 'Detectability of anthropogenic changes in annual temperature and precipitation extremes', *Journal of Climate*, vol. 17, no. 19, pp. 3683-3700. https://doi.org/10.1175/1520-0442(2004)017<3683:DOACIA>2.0.CO;2

Digital Object Identifier (DOI):
10.1175/1520-0442(2004)017<3683:DOACIA>2.0.CO;2

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Journal of Climate

Publisher Rights Statement:
Published in Journal of Climate by the American Meteorological Society (2004)

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

OPEN ACCESS

Download date: 12. Sep. 2026

Detectability of Anthropogenic Changes in Annual Temperature and Precipitation Extremes

GABRIELE C. HEGERL

Nicholas School of the Environment and Earth Sciences, Duke University, Durham, North Carolina

FRANCIS W. ZWIERS

Canadian Centre for Climate Modelling and Analysis, Victoria, British Columbia, Canada

PETER A. STOTT

Hadley Centre for Climate Prediction and Research, Met Office, Exeter, Devon, United Kingdom

VIATCHESLAV V. KHARIN

Canadian Centre for Climate Modelling and Analysis, Victoria, British Columbia, Canada

(Manuscript received 20 June 2003, in final form 7 April 2004)

ABSTRACT

This paper discusses a study of temperature and precipitation indices that may be suitable for the early detection of anthropogenic change in climatic extremes. Anthropogenic changes in daily minimum and maximum temperature and precipitation over land simulated with two different atmosphere–ocean general circulation models are analyzed. The use of data from two models helps to assess which changes might be robust between models. Indices are calculated that scan the transition from mean to extreme climate events within a year. Projected changes in temperature extremes are significantly different from changes in seasonal means over a large fraction (39%–66%) of model grid points. Therefore, the detection of changes in seasonal mean temperature cannot be substituted for the detection of changes in extremes. The estimated signal-to-noise ratio for changes in extreme temperature is nearly as large as for changes in mean temperature. Both models simulate extreme precipitation changes that are stronger than the corresponding changes in mean precipitation. Climate change patterns for precipitation are quite different between the models, but both models simulate stronger increases of precipitation for the wettest day of the year (4.1% and 8.8%, respectively, over land) than for annual mean precipitation (0% and 0.7%, respectively). A signal-to-noise analysis suggests that changes in moderately extreme precipitation should become more robustly detectable given model uncertainty than changes in mean precipitation.

1. Introduction

Increasing evidence shows that the recent evolution of observed global-scale surface air temperature is significantly different from estimates of natural climate variability and that most of the global temperature increase that has been observed during the latter half of the twentieth century can be attributed to anthropogenic forcing (Mitchell et al. 2001; Hegerl et al. 1997; Tett et al. 1999; Hegerl et al. 2003a; and many others). This result has been extended to large regional scales (Zwiers and Zhang 2003; Stott 2003; Karoly et al. 2003). Evidence of an anthropogenic influence on climate has also been detected in other thermal indicators of the mean climate, such as in the vertical temperature structure of the atmosphere (e.g., Santer et al. 1996; Thorne et al. 2002) and of the deep ocean (Levitus et al. 2000; Barnett et al. 2001; Reichart et al. 2002). However, while the early detection of anthropogenic change in mean thermal indicators is of great interest, the impacts of change on society are more likely to be felt via changes in climate variability and extreme climatic events. It is therefore important to identify the changes in climatic extremes that are expected under climate change conditions and to determine whether such changes may already be detectable and consistent with simulated changes.

Global climate system models forced with scenarios of future greenhouse gas and sulfate aerosol emissions (reviewed in Cubasch et al. 2001) project changes in climate mean, variability, and extremes. For example, a greater frequency of extreme heat events is generally

Corresponding author address: Dr. Gabriele Hegerl, Nicholas School of the Environment and Earth Sciences, Duke University, Durham, NC 27708.
E-mail: hegerl@duke.edu

© 2004 American Meteorological Society

simulated along with a reduction in the number of extremely cold days, a decrease in the diurnal temperature range, and an increase in summer drought (reviewed in Meehl et al. 2000). An increase in intense precipitation is projected under greenhouse warming conditions over large parts of the globe by most models (see Cubasch et al. 2001; Meehl et al. 2000; Kharin and Zwiers 2000, 2004, manuscript submitted to *J. Climate*, hereafter KZ04; Semenov and Bengtsson 2002; Wehner 2004, manuscript submitted to *J. Climate*, hereafter WEH). The increase in intense precipitation is usually greater in magnitude than the change in mean precipitation and often goes along with regional decreases in the probability of moderate precipitation events (e.g., Hennessy et al. 1997; Cubasch et al. 1995; Semenov and Bengtsson 2002). Precipitation changes projected by models are consistent with changes expected in a warmer atmosphere due to an acceleration of the hydrological cycle (e.g., Trenberth 1999, 2003; Allen and Ingram 2002). A disproportionally stronger increase in extreme precipitation is expected due to this increase in mean rainfall based on a simple statistical model of daily rainfall and the variations in its parameters with observed climate variability (Groisman et al. 1999). However, a recent study of changes projected with a regional climate model suggests that local increases in temperature variability may result in larger future temperature extremes than would be anticipated from projected changes in mean temperature (Schär et al. 2004). In the present study, we limit ourselves to large-scale changes in the statistics of extreme events in order to be able to study, and in the future detect, these changes from global data.

Observations show some changes in extreme events that are consistent with projected changes. For example, the unusually hot summer of 2003 in Europe has some similarity to projected changes in temperature variability (Schär et al. 2004). Changes in observed extreme precipitation have also been reported (Karl et al. 1995a,b; Jones et al. 1999a; Easterling et al. 2000a,b; Groisman et al. 2001, 2004, manuscript submitted to *J. Climate*, hereafter G04; among others). Karl et al. (1995a,b) and Karl and Knight (1998) found indications that the proportion of precipitation occurring during extreme events has increased over the United States, while worldwide results are more ambiguous (Easterling et al. 2000a; Frich et al. 2002).

Formal detection studies are required to determine whether the observed changes in extremes are a response to anthropogenic forcing or whether they are simply within the range of natural internal climate variability. This paper aims at setting the stage for this by studying changes simulated by climate models and estimating the strength of climate change ''signals'' in extremes relative to the corresponding natural internal variability (''noise''). We use indices of climate extremes that are directly comparable with indices of annual and seasonal mean changes. As is discussed in the following paragraphs, this approach addresses several questions that arise in the detection of climate change in extremes.

The first issue that immediately arises with respect to extremes is that daily station data are not readily comparable with daily model output. This occurs because model data represent area averages while station data are local, yielding quite different extremes, particularly for fields with fine spatial structure such as precipitation. Gridding station data, or indices of extremes that are derived from station data, is not a trivial task (e.g., Kiktev et al. 2003; McCollum and Krajewski 1998). Satellite and reanalysis data represent larger scales and are therefore more comparable with model data. However, reanalysis data are generally considered not to be reliable enough for the study of long-term trends (e.g., Chelliah and Ropelewski 2000). Datasets that blend satellite and surface data are becoming available for averages of a few days, such as pentads (e.g., Xie et al. 2003), but they remain the subject of active research.

This scaling issue can be circumvented, at least to first order, if changes in extremes result only from a shift in the distribution without a change in shape. In that case, changes in extremes could be predicted by changes in longer-term monthly, seasonal, or annual means, which have longer length scales and are often available as gridded data (e.g., Jones et al. 1999b; Xie et al. 2003). We therefore investigate whether projected changes in extreme precipitation and temperature are significantly different from projected changes in their seasonal and annual means.

A second issue is that extremes have different aspects that can be characterized by many different indices, some of which will be more sensitive to anthropogenic change than others. Thus, our ability to detect anthropogenic change in climatic extremes will depend on the index that is used. Therefore, anthropogenic change in a range of indices needs to be studied, both to select indices that are sensitive to climate change and to document indices that will not respond strongly to anthropogenic change. In order to estimate which indices of extremes should show detectable changes in the future, we attempt to detect the model-simulated change in model data. Such ''perfect model'' studies serve to establish an upper limit for the ratio of signal to noise of climate change because detection is not hampered by possible errors in the model response to anthropogenic forcing or in estimates of climatic natural internal variability.

In the present study, we scan indices that describe the transition from annual/seasonal mean data to averages of 1–30 extreme days per year. It is generally expected that the rarer the event of interest, the longer it will take to reliably observe and detect an anthropogenic change in the frequency or intensity of that event, and that averaging several events should improve the signal-to-noise ratio. Our results will document whether averaging is necessary or helpful. Many other indices of extremes have been proposed and are being used, such

as indices that count exceedances of (climatological or absolute) thresholds, lengths of periods without rain, and rainfall due to heavy rain, etc. (see Frich et al. 2002). Each has advantages and disadvantages, some of which are unexpected (see Zhang et al. 2004, manuscript submitted to *J. Climate*, hereafter ZHZK04). In this study, we restrict ourselves to relatively simple indices that allow a direct comparison between mean and extreme events. The index of the most rare events studied here (annual extremes) can be used to estimate the return periods of very rare events that have potentially severe consequences (see Zwiers and Kharin 1998; Kharin and Zwiers 2000; KZ04; WEH). We have also calculated indices based on threshold exceedances for precipitation, and they show qualitatively similar changes to the indices used here (see G04).

A third issue is that although simulated changes in extremes share some broad features between models, results can be very model dependent (see, e.g., Houghton et al. 2001; Meehl et al. 2000; Allen and Ingram 2002). To address this issue, we study changes in extreme rainfall and minimum and maximum temperature from two coupled climate models. The models, which are introduced below, are unrelated (they have different ocean, atmosphere, and sea ice components). By using two models we can extend the perfect model approach to begin to evaluate the sensitivity of detection results to differences in the models' responses to anthropogenic forcing. Such ''imperfect'' model studies should help us identify indices of climate change that permit both early *and* robust detection of changes in climate extremes. Ideally, a large number of models need to be used to get a complete estimate of model uncertainty. While this is beyond the scope of the present paper, a project to collect daily data from a range of models at a central location will facilitate multimodel studies of extremes in the future (see Hegerl et al. 2003b).

The remainder of this paper is structured as follows: Section 2 outlines the model data and indices for climate extremes used in this study. Section 3 compares climate change patterns for the means and extremes of temperature and precipitation. It establishes where changes in extremes are significantly different from those in seasonal means and where changes are consistent between the models. Section 4 describes estimates of signal-to-noise ratios for mean-to-extreme rainfall and temperature, investigating the detectability of changes. Finally, we draw some conclusions and discuss which questions need further investigation in section 5.

2. Data and indices for extremes

We use temperature and rainfall data in this analysis, namely, the lowest and highest temperature of each model day, referred to as daily minimum and maximum temperature (T_{min}, T_{max}), and daily precipitation data (P) at every model grid point. Daily data are used from 2 three-member ensembles of climate change simulations using two different models (Table 1):

TABLE 1. CGCM2 and HadCM3 segments of daily data available for analysis. Ensembles of three simlulations were available for both models. Climatological comparisons and signal-to-noise ratio assessments are performed using the $1 \times CO_2$ and $2 \times CO_2$ segments. Segments in italics contribute only to the estimation natural internal variability; all others are used to estimate both internal climate variability and climate change fingerprints and are used in lieu of observations to determine the detectability of changes.

Segment	CGCM2	HadCM3
$1 \times CO_2$ (first natural variability segment)	N/A	*1959–79*
$1 \times CO_2$	1975–95	1975–95
$1 \times CO_2$ (second natural variability segment)	N/A	*1978–98*
$2 \times CO_2$	2040–60	2040–60
$3 \times CO_2$	*2080–2100*	N/A

- The second generation Coupled Global Climate Model (CGCM2) from the Canadian Centre for Climate Modelling and Analysis (CCCma; Flato and Boer 2001)—The atmospheric component of the model has a T32 horizontal resolution, which corresponds to a gridpoint resolution of approximately 3.75° latitude × 3.75° longitude. The oceanic component has double this resolution in each direction. We use data from simulations of twentieth- and twenty-first-century climate change due to greenhouse gas and direct aerosol forcing. Our analysis is based on three segments of 21-yr duration: segment 1975–95 (''$1 \times CO_2$''), which we use for defining climatological mean values for the present; segment 2040–60 (''$2 \times CO_2$''), which is approximately the time of CO_2 doubling relative to the base period; and segment 2080–2100 (''$3 \times CO_2$''). Changes in return values of annual extremes of temperature and precipitation simulated by this model have been analyzed in KZ04, and by its predecessor, CGCM1, in Kharin and Zwiers (2000). We also performed our analysis with output from CGCM1 and obtained quite similar results.
- The Third Hadley Centre Coupled Ocean–Atmosphere GCM (HadCM3; Gordon et al. 2000; Pope et al. 2000)—Its atmospheric component has the same resolution longitudinally as CGCM2, but a higher latitudinal resolution of 2.5° latitude. The oceanic component has a resolution of 1.25° in each direction. Here we use a three-member ensemble of climate change simulations (Johns et al. 2002), which includes greenhouse gas forcing, direct and indirect sulfate aerosol forcing, and tropospheric and stratospheric ozone forcing. We use data from 1959 to 1998, which includes 1975–95 segment ($1 \times CO_2$) and the same 2040–2060 segment ($2 \times CO_2$) as for CGCM2.

Climate change simulations with both models used greenhouse gas and direct aerosol forcing prescribed from the Intergovernmental Panel on Climate Change (IPCC) Special Report on Emission Series A2 (SRES

A2) emissions scenario. However, the CGCM2 simulations include direct aerosol forcing only, in addition to greenhouse gases, while the HadCM3 also includes the effects of tropospheric and stratospheric ozone and indirect aerosol forcing, leading to a total anthropogenic forcing of 3.43 W m^{-2} by 2050 relative to the average of 1880–1920 as compared to 4.29 W m^{-2} due to greenhouse gas forcing alone (note that the stratospheric ozone forcing is overestimated in some seasons; see Gillett and Thompson 2003). Therefore, the comparison between the climate change fingerprints of the two models will be somewhat affected by differences in forcing. This is similar to the situation when comparing observations with any model simulation because there would inevitably be differences between the two due to imprecisely specified forcing in the model. This comes about because the level of scientific understanding of anthropogenic nongreenhouse gas and indirect aerosol forcing remains limited (Houghton et al. 2001).

Both models have been used for the detection of climate change in the IPCC Third Assessment Report and thereafter (Mitchell et al. 2001; Tett et al. 2002; Jones et al. 2003; Stott et al. 2003; Zwiers and Zhang 2003). CGCM1 and HadCM2 (the predecessor of HadCM3) were also used in the United States National Assessment of the potential impacts of climate variability and change (National Assessment Synthesis Team 2001).

We have calculated a number of indices that sample the transition from climatic means to extremes from these data for each model. For minimum and maximum temperature, we extracted the boreal summer [June–August (JJA)] and winter [December–February (DJF)] season means and calculated the average of the 30, 10, 5, and 1 hottest and coldest days for each year. This yields 10 indices for each model grid point and year of the simulations for minimum and maximum temperature, respectively.

For precipitation, we calculated the annual mean (because the wet and dry seasons do not usually coincide with summer or winter within a hemisphere) and the average of precipitation on the 30, 10, 5, and 1 wettest days of the year. Additionally, we calculated the average accumulation per day on the 5 wettest consecutive days (from 365 overlapping 5-day periods) in a year. To keep the number of samples consistent for the different indices, we used 20 values from each 21-yr chunk from an individual ensemble member (each 21-yr chunk contains only 20 full DJF seasons). This yields 60 values from the three ensemble members for each 21-yr epoch. Additionally, we have processed the most extreme day in 5 yr, and the wettest pentad (wettest period among nonoverlapping 5-day accumulations in a year) for a comparison with observations. We also used a variation of the indices based on the 30th, 10th, and 5th most extreme days of the year (rather than the average of the 30, 10, 5, and 1 most extreme days) for comparison of tail shapes between models.

For several reasons, we have restricted our analysis primarily to changes in these indices over land. The reasons include the fact that there is a stronger impetus to study extremes that impact the inhabited part of the globe and that long in situ records (station data) are generally only available over land (daily data available online at www.ncdc.noaa.gov/gdcn.html; see Frich et al. 2002 for indices of extremes).

The indices have been calculated without removing the seasonal cycle. Therefore, the hottest temperatures will be generally drawn from the summer season, the coldest will be taken from the winter season, and the wettest days are expected to be drawn from each region's wet season(s). We consider it an advantage to study indices of absolute extremes because these kinds of extremes have potentially high societal impact. The disadvantage is that the present indices cannot be readily compared with observed indices that are based on extremes relative to climatological means, for example, the exceedance of a seasonally adjusted percentile (see Frich et al. 2002), which, however, may be problematic to use in a detection study (ZHZK04). We did compute the number of exceedances of the 99th and 99.7th percentile of rainfall on rainy days from both models, since this is similar to indices processed from observations (G04; see also Frich et al. 2002). Results are not shown in this paper, but patterns of change were similar to those from indices used here.

All index data have been transformed to the slightly coarser grid of the CGCM2 model prior to calculating correlations between climate change patterns and signal-to-noise ratios (see section 4). Ideally, this transformation should be applied to the daily data prior to calculating indices. However, given the well-organized structure that is apparent in the index fields, and the small difference between model grid size, we do not expect this to substantially affect results.

3. Simulated changes in temperature and precipitation extremes

a. Temperature

We have compared the climatological values (derived from the average of the $1 \times CO_2$ chunks) for mean and extreme temperature in both models. Figure 1 shows the climatologically warmest minimum temperature, that is, warmest night, of the year $\overline{\max_{1 \le t \le 365}\{T_{\min,t}\}}$, averaged over the $1 \times CO_2$ segments of the three ensemble simulations. The models are quite similar. In both models, the hot days and cold nights show peak values over the continents (not shown), while climatologically warm nights (Fig. 1) and cold days (not shown) have a more zonal pattern. The largest difference between the models seems to occur in the Tropics for the hottest day of the year, which is substantially hotter over Brazil and parts of Africa and southern North America in HadCM3 than in CGCM2 (with tropical land maximum temperatures

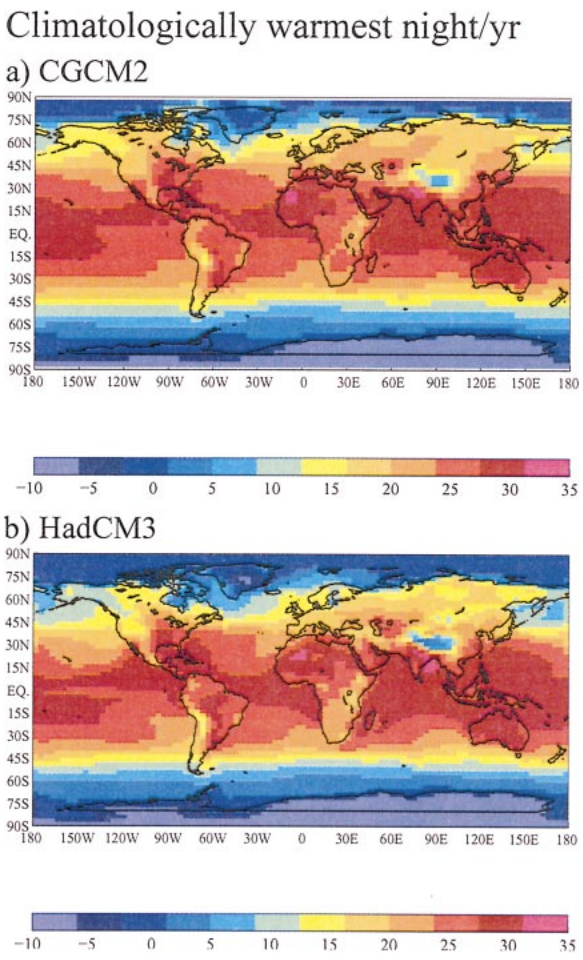

FIG. 1. The present-day mean temperature for the warmest night of the year (time averaged over the 1 × CO_2 segment for 1975–95 and across ensemble members) in (a) CGCM2 and (b) HadCM3. Units are °C.

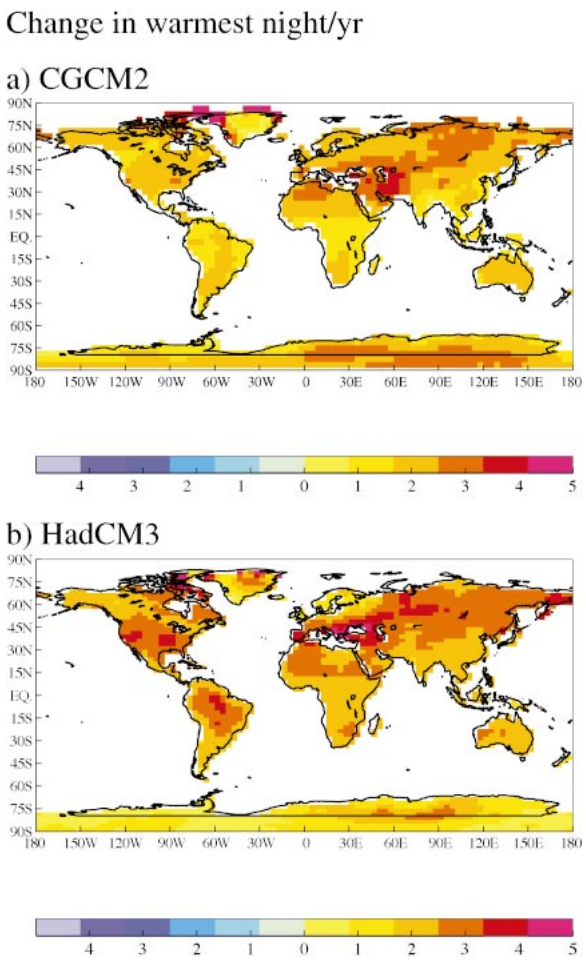

FIG. 2. Projected climate change for the warmest night of the year between the time of CO_2 doubling (segment 2040–60) and the 1 × CO_2 segment in (a) CGCM2 and (b) HadCM3. The climate change patterns are shown over land only. Units are °C.

averaging 40°C versus 33°C; not shown). Other differences are that the slightly higher resolution in HadCM3 results in somewhat more detail in mountainous regions. Overall, global land temperatures have lower cold extreme minimum temperatures $\min_{1\le t\le 365}\{T_{\min,t}\}$ (average 0.8°C versus 3.5°C) and cold extreme maximum temperatures $\min_{1\le t\le 365}\{T_{\max,t}\}$ (5.0°C versus 7.6°C) in HadCM3 than CGCM2. That is, the coldest night of each year and the coldest day of each year tend to be cooler on average in HadCM3. In contrast, globally averaged warm extreme minimum temperatures $\max_{1\le t\le 365}\{T_{\min,t}\}$, the warmest night of the year (20° and 21°C T_{min}; Fig. 1) and warm extreme maximum temperatures $\max_{1\le t\le 365}\{T_{\max,t}\}$, the warmest day of each per year (24° and 25°C T_{max}) are quite similar in the two models. A rigorous evaluation of the climatology in both models is beyond the scope of this paper [see Zwiers and Kharin (1998) for more discussion].

The patterns, or ''fingerprints,'' of climate change that we will use in our detection assessment are obtained by taking differences between averages of our temperature and precipitation indices during the 2 × CO_2 segment and the 1 × CO_2 segment. For each segment, averages are calculated over time (20 yr) and over ensemble members. Fingerprints for temperature are quite similar within a given model for the different indices, and therefore only the pattern of change for the warmest night of the year (warmest T_{min}) is shown in Fig. 2. Correlations between climate change patterns over land are given in Table 2. Correlations computed by including the spatial mean (which are more relevant for regression-based detection and attribution studies) are generally high. Results indicate that the spatial details for changes in warm means and extremes are in greater agreement than for cold means and extremes. All climate change patterns show overall warming, and changes are significant at the 10% level at nearly every grid point according to a nonparametric Mann–Whitney test (von Storch and Zwiers 1999). In both models, global land average changes are a bit larger for cold means

TABLE 2. Spatial correlations between the CGCM2 and HadCM3 climate change patterns for a range of indices of max and min temperatures [seasonal mean, average of 30, 10, 5, and 1 warmest (warm) and coldest (cold) extreme (ex) days of the year]. The seasonal mean gives correlations for boreal summer (JJA) for warm extremes and boreal winter (DJF) for cold extremes. In all cases, the spatial mean is included. Correlations with spatial means removed (in parentheses) are generally lower but are less relevant for regression-based detection methods used in this study.

	Seasonal mean	30ex	10ex	5ex	1ex
Warm T_{max}	0.90 (0.61)	0.92 (0.71)	0.92 (0.72)	0.92 (0.72)	0.92 (0.72)
Cold T_{max}	0.87 (0.49)	0.85 (0.41)	0.84 (0.41)	0.84 (0.41)	0.83 (0.41)
Warm T_{min}	0.93 (0.63)	0.94 (0.72)	0.94 (0.73)	0.94 (0.73)	0.94 (0.72)
Cold T_{min}	0.88 (0.50)	0.83 (0.43)	0.82 (0.44)	0.82 (0.43)	0.81 (0.43)

and extremes (2.4°–3.0°C) than for warm means and extremes (2.0°–2.6°C). Cold extremes exhibit strong increases over northern mid- and high latitudes while corresponding changes in the Southern Hemisphere (SH; not shown) are not as large. In contrast, warm extremes show overall strong warming over midlatitude landmasses (Fig. 2).

If the changes in the temperature extremes were due solely to a constant shift in the distribution of the daily values throughout the hot or cold season, then the extremes would change by the same amount as the corresponding warm or cold season means. However, this is not the case for a large part of the globe as shown in Figs. 3–5. The left-hand side in Figs. 3a,b shows the projected change of temperature on the warmest night of the year relative to the change in the warm season mean value of T_{min}. Note that the ''warm'' season is defined as JJA in the Northern Hemisphere (NH) and DJF in the SH, and thus change in the Tropics is difficult to interpret. Similarly, the right-hand side in Figs. 3a,b shows the projected change of temperature of the warmest day of the year relative to the change in the warm season mean value of T_{max}. Figure 4 displays corresponding changes in the coldest night of the year relative to the change in the cold season mean T_{min} and the coldest day of the year relative to change in the cold season mean T_{max}. Changes in both figures are displayed only where they are assessed to be significant (by a Mann–Whitney test) at the 10% level, which occurs at 62%–66% of model grid points for warm extremes and 39%–57% of grid points for cold extremes. Note that for warm extremes (Fig. 3), warm colors indicate larger extremes relative to the seasonal mean and hence a tendency toward a more extreme climate. For cold extremes (Fig. 4), blue colors indicate a tendency toward a more extreme climate. The patterns of change for cold extremes are more similar between the models ($r = 0.62$ for the coldest night of the year, and $r = 0.61$ for the coldest day of the year) than for warm extremes ($r = 0.43$ for the warmest night, and $r = 0.37$ for the warmest day). These relatively lower correlations for changes in extremes relative to means (as compared to the correlations for changes in extremes; see Table 2) are not surprising given that this is a quite subtle feature of temperature change.

This change in the shape of the distribution is particularly pronounced over parts of Europe. Figure 5 shows the change in the average location of the 30th, 10th, 5th, and 1st most extreme maximum temperature per year, and the most extreme event in 5 yr in a box containing large parts of Europe (37.5°–60°N, 0°–41.3°E). Results for T_{min} (not shown) are broadly similar. The tendency toward less severely cold winter days on average is consistent with an increase in NH annular modes, which leads to more westerly flow into Europe (Fyfe et al. 1999; Gillett et al. 2003). However, in HadCM3, the warm extremes become substantially more severe, as also seen in a regional climate model that is driven by HadCM3 [Schär et al. (2004), who also discuss reasons for this change, e.g., increased incidence of drought].

Warm nights also tend to change somewhat more for events farther out on the tail of the distribution for South America and eastern North America. In addition, HadCM3 simulates a change toward milder cold extremes in Siberia that is not supported by CGCM2 (Figs. 3 and 4). Overall, cold nights and days tend to become less extreme relative to the seasonal mean in parts of the mid- to high-latitude NH in both models. These latter changes are, at least partly, related to a change in the shape of the annual cycle during the cold season (not shown) and likely also to a loss of snow and ice cover (Kharin and Zwiers 2000). In both models, changes in the day-to-day climate variance are generally similar to the patterns of change shown in Figs. 3 and 4, which is consistent with results from individual regions (Fig. 5).

We conclude that in both models, changes in annual extremes differ from those in the seasonal means, and it should not be assumed that the shape or spread of the T_{min} and T_{max} distributions remains the same under global warming. The reasons for these changes in the shape of the temperature distribution are complex and beyond the scope of this paper. The changes are probably related to changes in other variables such as cloudiness, soil moisture, and the retreat of snow and ice cover (Zwiers and Kharin 1998; Kharin and Zwiers 2000; KZ04).

b. Precipitation

The models have similar climatological annual global mean precipitation (2.8 mm day^{-1} for CGCM2 versus

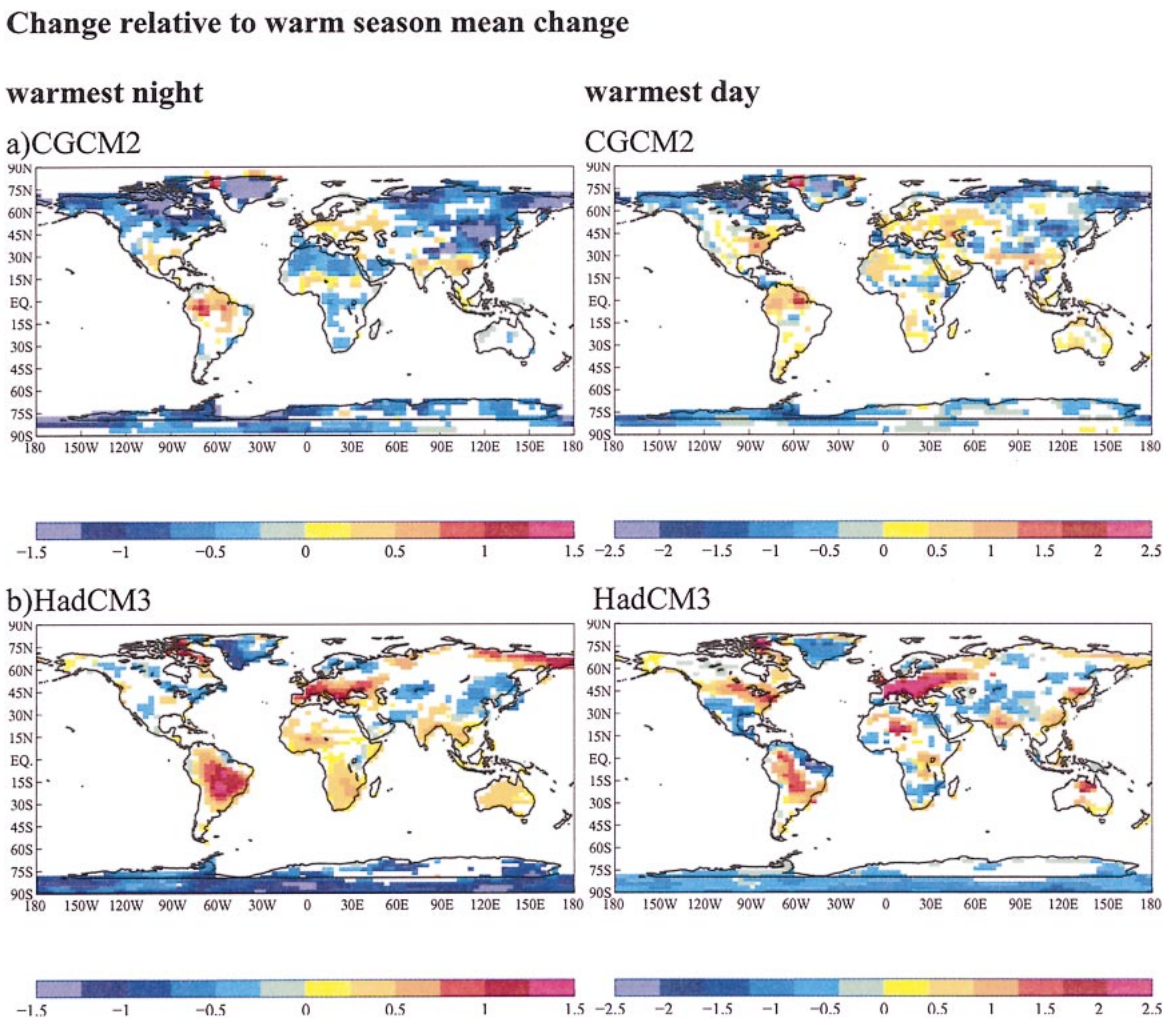

FIG. 3. (left) Difference between the projected change in the warmest night of the year (see Fig. 2) and the change in the respective warm season mean of T_{min} (JJA for NH; DJF for SH). The change relative to the seasonal mean is shown only over land, and only where changes are significant at the 10% level; (a) for CGCM2 and (b) for HadCM3. (right) As in left-hand side, except the difference between the warmest day of the year and the change in the warm season mean of T_{max}. Positive differences indicate that extremes change more strongly than the seasonal mean, yielding more severe warm extremes. Negative differences indicate a change toward less severe warm extremes. Units are °C.

2.9 mm day^{-1} for HadCM3), but the extremes are substantially more pronounced in CGCM2 than HadCM3 (a global mean of 56.7 versus 30.1 mm day^{-1} for the wettest day of the year). The spatial patterns of the climatological annual mean and extreme precipitation are arguably not as similar between the models of the temperature.

Figure 6 displays the spatial pattern of the average precipitation rate of the wettest pentad (5-day mean) of the year in the $1 \times CO_2$ segments from both models. Corresponding statistics from a blended observed data product (Xie et al. 2003) and the European Centre for Medium-Range Weather Forecasts (ECMWF) 15-yr Re-Analysis (ERA-15; Gibson et al. 1997) are also displayed. Note that the figure shows the strongest average rainfall rate among 73 adjacent pentads (5-day means) occurring over a year. This is slightly different from the index based on the wettest 5-day running mean per year that is used elsewhere in the paper. We display the wettest pentad here because 5-day running means are not available for the blended observational product. The global mean accumulation per day during the wettest pentad is higher in CGCM2 (20 mm day^{-1}) than in HadCM3 (14 mm day^{-1}), which is the same in the Xie et al. data and slightly less than the ERA value (16 mm day^{-1}). The spatial distribution of climatological extreme precipitation differs between models, although overall features are similar. The spatial pattern in HadCM3 shows a more pronounced intertropical convergence zone (ITCZ) and a smaller area of peak rainfall over the far western equatorial Pacific than CGCM2. The models also show marked differences in the location of wet areas over Africa, South America, and the Indian Ocean. Overall, both models simulate 5-day precipita-

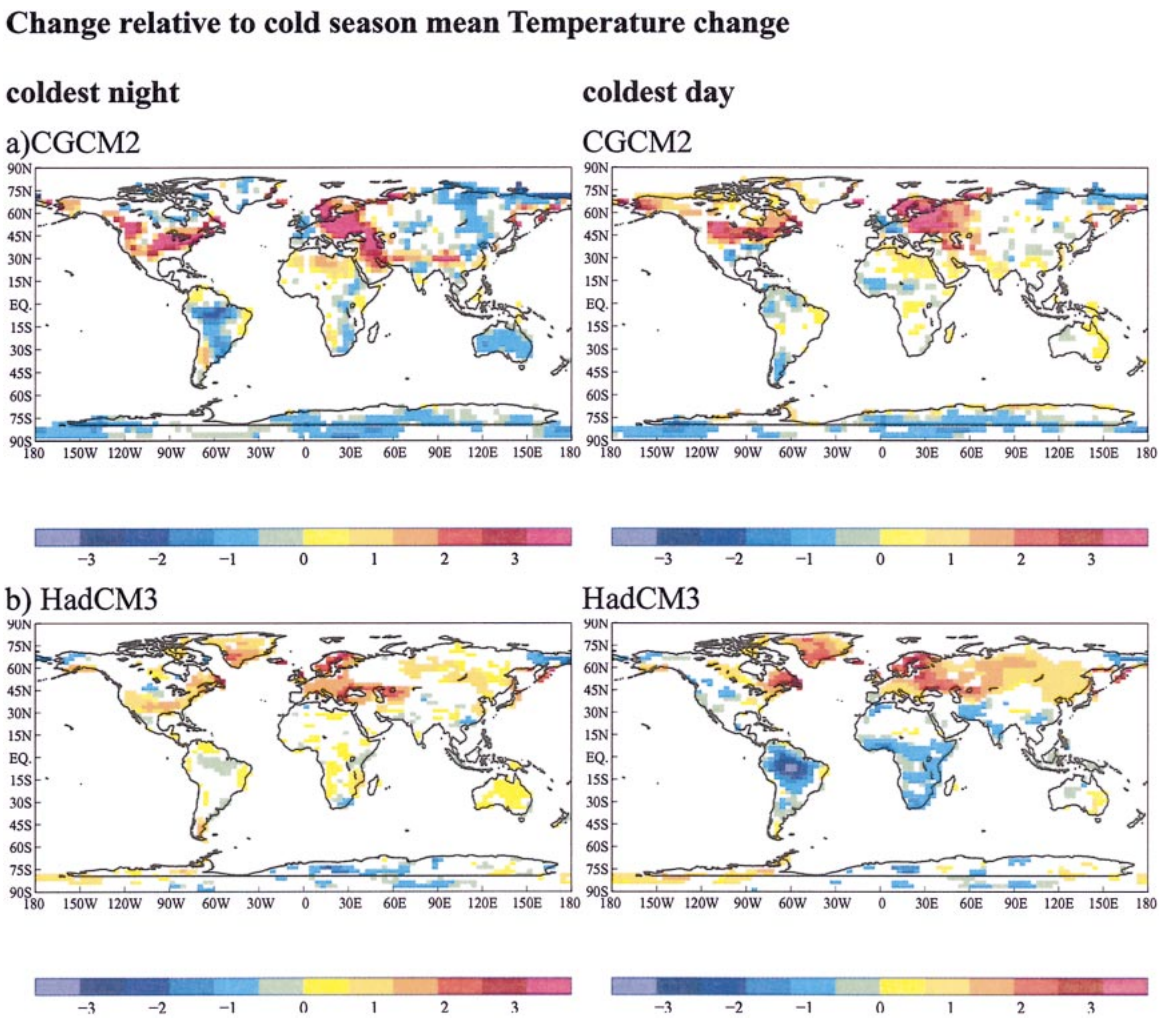

FIG. 4. (left) Difference between the projected change in the coldest night of the year (between the 1 × CO_2 and the 2 × CO_2 segments) and the change in cold season mean of T_{min} (DJF for NH; JJA for SH data). The change relative to the seasonal mean is shown only over land and only where changes are significant at the 10% level. (right) As in left-hand side, except the difference between the change in the coldest day of the year and the change in the cold season mean of T_{max}. Positive differences indicate that extremes change more strongly than the seasonal mean, yielding less severe cold extremes. Negative differences indicate a change toward more severe cold extremes. Units are °C.

tion events reasonably well, with HadCM3 appearing to agree a bit better with observational data over most regions of the globe.

Figure 7 shows the pattern of change in zonal mean, annual mean, and moderately extreme precipitation between the 1 × CO_2 and 2 × CO_2 segments in both models. The zonal mean changes are quite similar over the NH high latitudes, moderately similar over the NH and SH extratropics, and differ most strongly over the Tropics and parts of the subtropics. Note that while the changes become uniformly positive and stronger for the zonal average of moderately extreme compared to annual mean rainfall, the pattern of change for both is somewhat similar within a model. The similarity of the change pattern and the tendency toward stronger and more positive changes is also seen in indices of more extreme events within each model (not shown).

Figure 8 displays the percent change (change between global mean indices of mean and extreme rainfall, expressed as the percent of climatological global mean indices) in the global mean at the time of CO_2 doubling for annual mean precipitation and a range of indices of extremes that scan the tail of the distribution (the 30th wettest day of the year, the 10th wettest day of the year, the wettest day of the year, and the wettest day in 5 yr; changes in the average of 30, etc., wettest days are similar). We again see that simulated changes in extreme precipitation are larger than in mean precipitation, as has also been discussed in the literature (see, e.g., Meehl et al. 2000; Semenov and Bengtsson 2002). Allen and Ingram (2002) demonstrated a transition from moderate changes in mean precipitation to increases of around 20% for very extreme precipitation for HadCM3, which is close to changes following the Clausius–Clapeyron

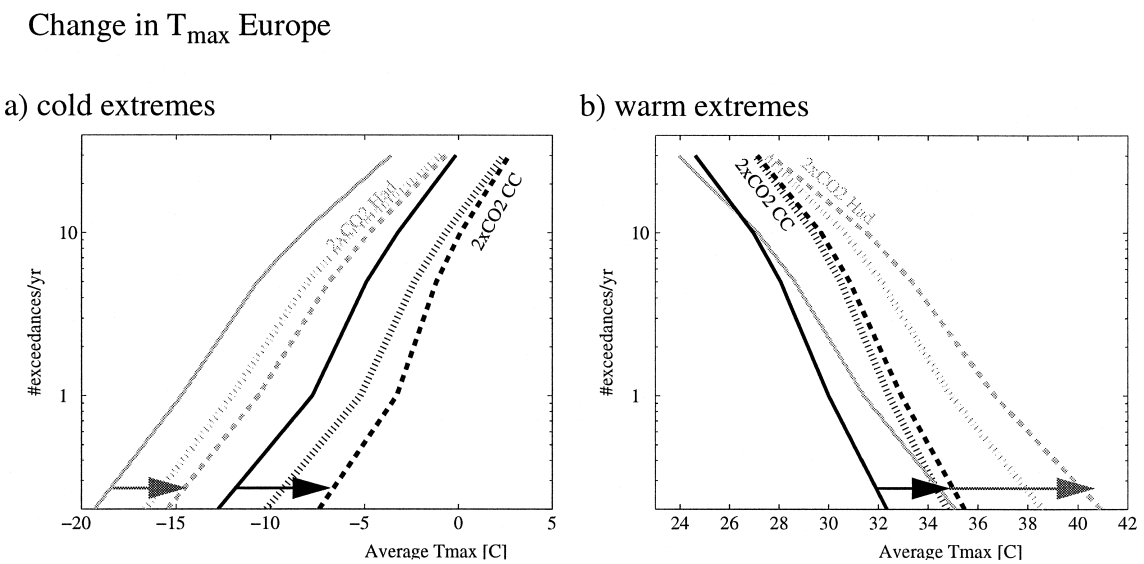

FIG. 5. 1 × CO_2 (solid) and 2 × CO_2 (dashed) distribution of (a) cold and (b) warm extreme max temperatures averaged over a box containing central Europe (37.5°–60°N, 0°–41.3°E). The horizontal axis gives the average temperature of extremes; the vertical axis gives the location of the 30th, 10th, 5th, and 1st most extreme days, and the most extreme day in 5 yr (logarithmic scale). Black (gray) lines show extremes for CGCM2 (HadCM3). The dotted lines show where the tail would be at the time of CO_2 doubling if changes were the same for extremes as for the seasonal mean. The arrows indicate the direction of the change. Units are °C.

relation. This increasingly strong change for the rarer precipitation extremes is qualitatively confirmed by changes in the indices for mean and extreme climate in the present paper for both models (Fig. 7), with peak global changes of 6.1% (CGCM2) and 9.5% (HadCM3) at the time of CO_2 doubling (and 13% at the time of tripling in CGCM2; not shown). Land-only changes (thin lines in Fig. 8) are somewhat reduced compared to global mean changes in CGCM2 (indicating that the heaviest increases in extreme rainfall occur over ocean); there is a weak opposite tendency in HadCM3.

While changes in the global means of the precipitation indices in the two models are comparable, the spatial patterns of change (fingerprints) are quite model dependent. This can be seen by comparing the left- and right-hand panels in the upper two rows of Fig. 9; they display the spatial patterns of anthropogenic climate change in annual mean and annual extreme precipitation. Table 3 lists spatial correlations between the climate change patterns for the various indices. Note that the changes displayed in Fig. 9 are expressed as a percent change at each grid point relative to the corresponding climatological 1 × CO_2 value. Changes are plotted only where they are significant at the 10% level according to a Mann–Whitney test.

The bottom row of Fig. 9 shows a spatially smoothed version of the mean pattern of precipitation change averaged across the two models. Smoothing is performed by applying a five-point spatial filter twice, where the filter assigns a weight of 1/3 to the central point and 1/6 to its four nearest neighbors. The spatial smoothing is used to focus attention on the larger-scale features of climate change. The two-model average is plotted only where the smoothed changes are consistent between the models (i.e., where a Mann–Whitney test does not find significant differences at the 10% level between the simulated changes). These panels therefore indicate the land areas that are relevant for detection between models.

Figure 9 shows that there is greater similarity between the spatial patterns of annual rainfall change in the high latitudes than in the mid-to-low latitudes and that the patterns of change are quite different over parts of the subtropics and Tropics. As a result, the correlation between the patterns is modest (Table 3). Some of the most pronounced changes in annual rainfall do not agree between the models. Note, for example, that precipitation decreases strongly over the Amazon basin and southwestern and northwestern Africa in HadCM3, while CGCM2 simulates more moderate changes in these regions. In contrast, CGCM2 shows a precipitation decrease over Southeast Asia and the Atlantic United States, which is not supported by HadCM3.

The difference between the climate change pattern of annual mean precipitation and the wettest day of the year is significant over most areas of the globe in both models (not shown) and further confirms the general widening of the tail of the precipitation distribution under climate change conditions that was demonstrated in Fig. 8. Correlations between annual mean and extreme climate change patterns are moderately high within a model (Table 3) but quite low between models. Correlations between models are largest for moderately extreme rainfall (peaking at the average of the 10 or 5 wettest days per year). Despite these low correlations, a larger fraction of model grid points shows consistent changes between the models for annual extreme (59%)

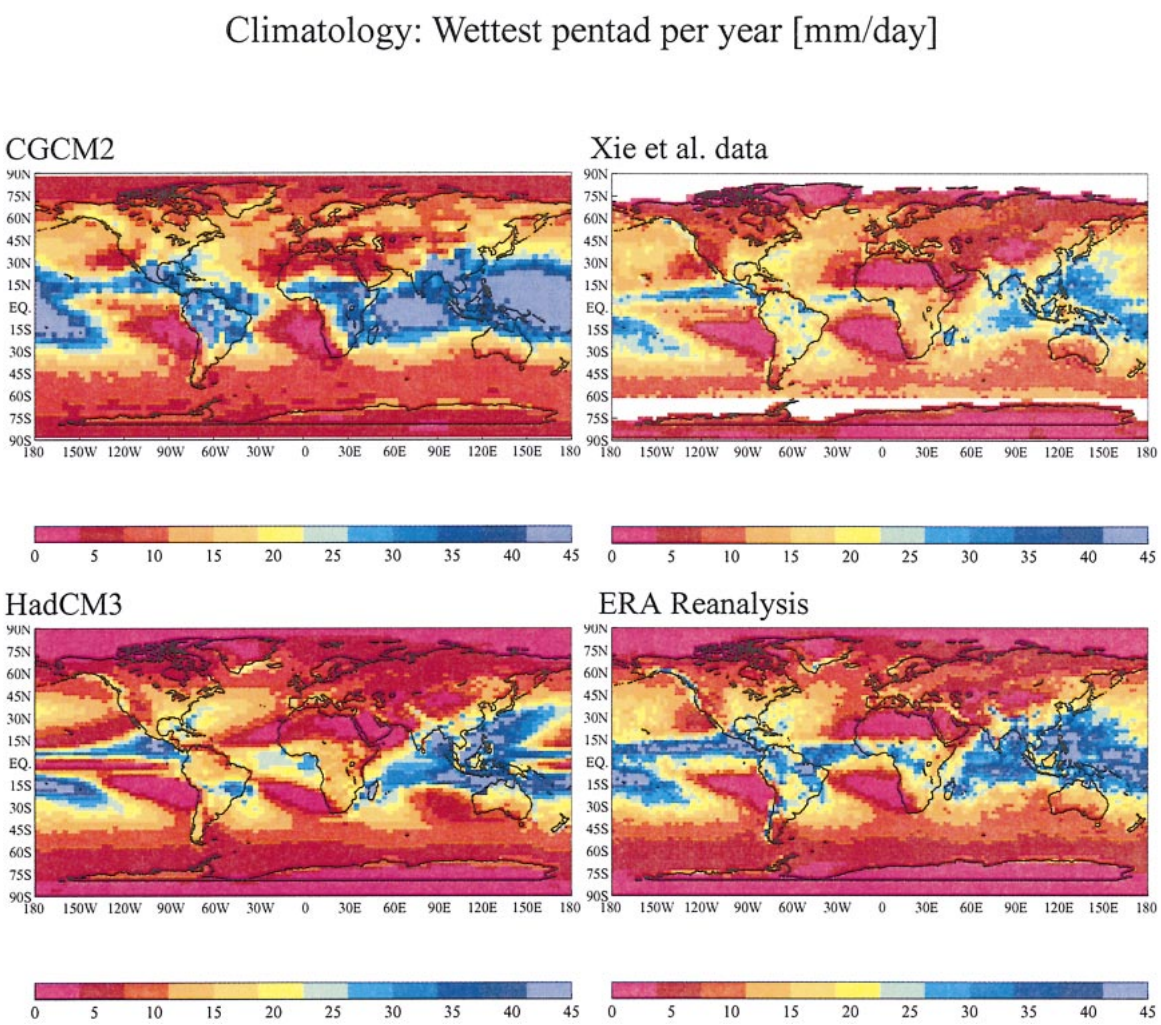

FIG. 6. Spatial pattern of the mean precipitation rate of the climatologically wettest pentad (nonoverlapping 5-day period) of the year. (left) Results from the climate models for the 1 × CO_2 segment. (right) Results from the Xie et al. (2003) blended dataset (top, 1979–98; blanks indicate missing data) and ERA-15 (bottom, 1979–93). Units are mm day^{-1}.

than for annual mean rainfall (41%); most of these changes indicate heavier extreme rainfall. In the next section, we will explore whether this indication for greater similarity between changes in extreme precipitation helps in detecting climate change between models.

4. Detectability of changes

a. Perfect and imperfect model studies

Climate change detection and attribution studies generally rely on sophisticated multiple regression methods (see Mitchell et al. 2001). These methods aim to estimate the amplitude of one or several model-simulated signals of climate change in observations (see, e.g., Hasselmann 1979, 1997; Allen and Tett 1999). We restrict ourselves to the detection of a single anthropogenic signal in this study. We use the spatial pattern of climate change between the 1 × CO_2 and 2 × CO_2 segments (each averaged over all 20 yr and at least two ensemble members) as the fingerprint representing the spatial pattern of the expected climate change signal.

The general regression equation for the detection of this climate change signal **g** (written as a vector over all grid points x) in observations **X** is given by

$$\mathbf{X} = a\mathbf{g} + \mathbf{u}. \quad (1)$$

If all forcings have been accounted for in the signal **g** then the residual vector **u** should contain only noise that reflects internally generated climate variability, as opposed to the forced climate response that we aim to detect.

A perfect model study assesses the detectability of a signal **g** in a ''data'' vector **X** of ''model observations'' that is obtained from a model simulation in order to estimate how detectable **g** might be in the real world at

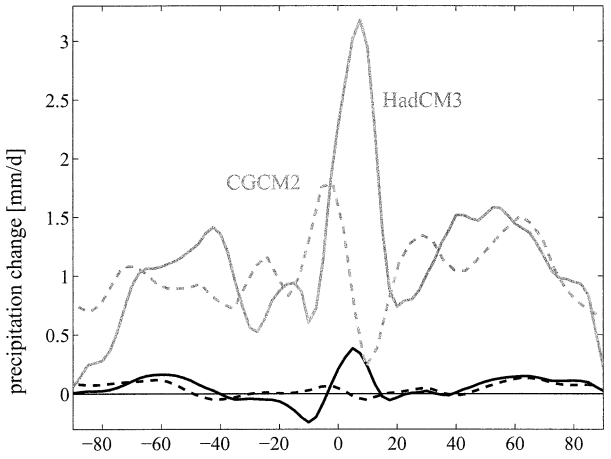

FIG. 7. Zonal mean precipitation change (land and ocean) for HadCM3 (solid) and CGCM2 (dashed). Annual mean changes are shown in black, and changes in the mean of the 10 wettest days of the year are shown in gray. Changes in other indices of extremes are qualitatively similar, with changes in extremes becoming overall positive and stronger for indices of increasingly rare events. Units are mm day^{-1}.

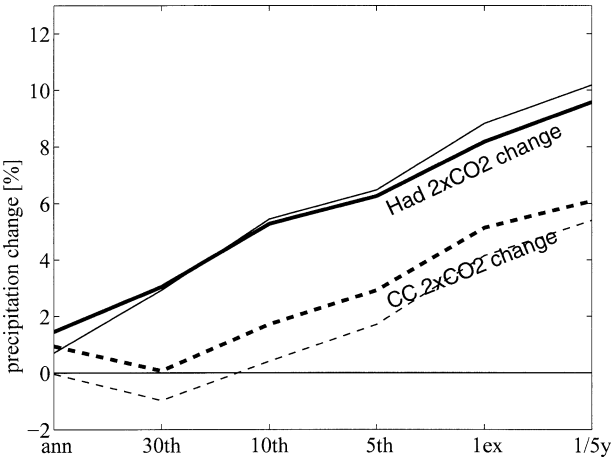

FIG. 8. Global mean precipitation change for HadCM3 (solid) and CGCM2 (dashed) between the $1 \times CO_2$ and $2 \times CO_2$ segments for the annual mean (ann), and the 30th, 10th, 5th, and 1st wettest day of the year (1ex) and the wettest day in 5 yrs (1/5yr). Values for land only are shown as thin lines. The changes are expressed as the global (or land only) mean of the precipitation change divided by the climatological global (or land only) mean precipitation for each index. Units are %.

some point in the future. We use model observations that are obtained by calculating the difference between the $1 \times CO_2$ and $2 \times CO_2$ segments of a single model simulation. This corresponds to the way in which a data vector **X** might be obtained in the real world because nature will also provide only a single realization of the climate change between two epochs. The fingerprint **g** is obtained from the two remaining simulations, which ensures that the noise in the fingerprint is statistically independent of the model observations. Unavoidably, the resulting fingerprint will be somewhat influenced by internal climate variability. However, the large-scale features of the fingerprints are robust across the three fingerprints that can be obtained from a given ensemble of three climate model simulations. We also apply the detection method to the spatial pattern of the trends ''observed'' during the $1 \times CO_2$ and $2 \times CO_2$ segments to determine if and when climate change should be detectable over shorter time spans. In this case, the model observations are the trend patterns in the individual simulations. However, the same fingerprints were used as in the main detectability study, since those fingerprints should be less influenced by noise than fingerprints derived from 20-yr trends.

Once a signal has been estimated by calculating a fingerprint, as discussed above, its presence or absence in the observations is evaluated by estimating the scaling factor a that appears in (1). The signal is detected if this estimate is found to be significantly greater than unity. The scaling factor is estimated from the observations by calculating

$$\hat{a} = (\mathbf{g}^T\mathbf{X})/(\mathbf{g}^T\mathbf{g}), \quad (2)$$

where the quantity in parentheses denotes a scalar product computed relative to a kernel matrix (discussed below). The best (least noise influenced) estimate $\hat{a}$ is obtained when the inverse noise covariance is used as the kernel in scalar product (2), thereby replacing **g** with the so-called ''optimal'' fingerprint (Hasselmann 1979, 1997; Allen and Tett 1999). However, a large sample of natural variability is needed to estimate the noise covariance. Obtaining such a large sample of daily data from control simulations was not feasible at the time of writing, and thus we apply only the standard regression approach that uses the identity matrix as the kernel in the scalar product. Use of the optimal fingerprint would increase the signal-to-noise ratio, although uncertainties in our knowledge of signal and noise, and data limitations, often limit the gain (Hegerl et al. 2000).

Prior to the analysis, all data are smoothed by twice applying the five-point spatial filter described above. The smoothing ensures that the spatial patterns are correlated over regions larger than $10^\circ \times 10^\circ$ and thus ensure that the analysis focuses on broad spatial scales. Small spatial scales do generally not show realistic variability in coupled climate models (see Allen and Tett 1999). Because precipitation extremes tend to vary on small spatial scales, for example, due to orography, we again express changes in percent of climatological mean and extreme precipitation. This makes changes more comparable between climatologically different regions. Spatial patterns of percent changes appear organized and show large-scale features (Fig. 9).

The amplitude estimate $\hat{a}$ [(2)] is subject to uncertainty due to climate noise, and therefore climate change is only considered to be detected if the estimated $\hat{a}$ is significantly larger than zero. The usual approach for estimating the uncertainty of $\hat{a}$ that is due to climate noise is to repeatedly apply the detection technique to

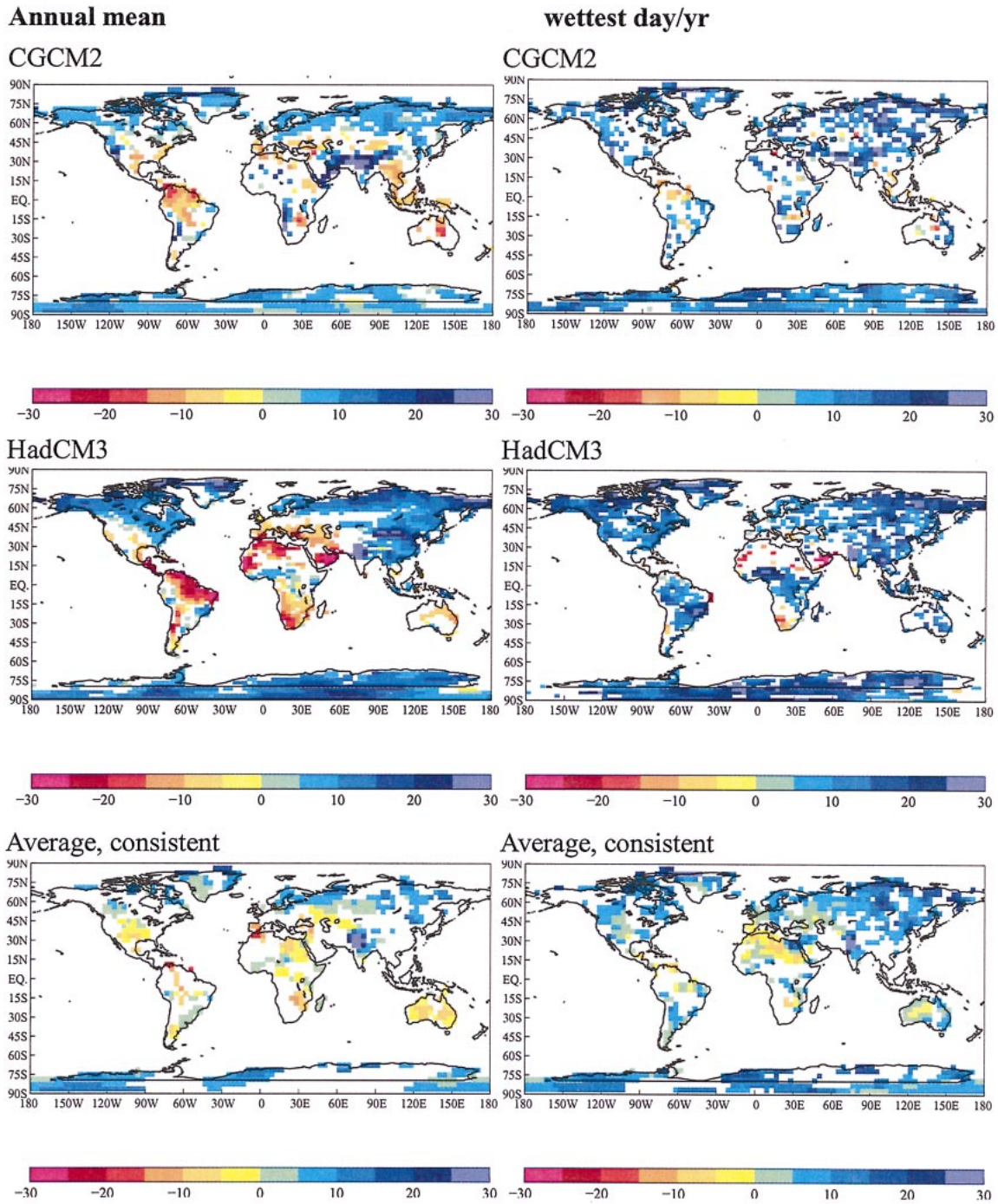

FIG. 9. Change in (left) annual mean precipitation and (right) the wettest day of the year in (top) CGCM2 and (middle) HadCM3 at the time of CO_2 doubling. Changes are expressed as percent of the present-day climatological value; the scale ranges from −30% to +30%. Changes are only plotted where they are significant at the 10% level. (bottom) The average of climate change patterns from both models where the large-scale (smoothed) changes are consistent between the models.

independent samples of internal variability from long control simulations. However, daily data from the CGCM2 and HadCM3 control simulations were not available at the time of writing, and thus we use the variability within ensembles to estimate internal climate variability.

The quantity of interest in this case is the internal variability of the difference between well-separated 20-

TABLE 3. Spatial correlations between precipitation change patterns for the different indices (annual mean; average of 30, 10, 5, and 1 wettest day of the year; and the precipitation rate for the 5 wettest consecutive days of the year). The first four rows give correlations between the annual mean and extreme climate change patterns for both models using all land grid points and only land grid points where changes are significant. The bottom two rows give correlations between climate change patterns for the same index in the two models. In all cases, the spatial mean is included.

		Ann	30 wettest	10 wettest	5 wettest	1 wettest	5-day rate
CGCM2	Ann vs extremes	1	0.96	0.88	0.83	0.74	0.75
	Significant grid points	1	0.98	0.95	0.92	0.84	0.90
HadCM3	Ann vs extremes	1	0.89	0.78	0.72	0.59	0.74
	Significant grid points	1	0.94	0.90	0.87	0.80	0.91
CGCM2 vs HadCM3	All land grid points	0.27	0.28	0.29	0.28	0.21	0.21
	Significant grid points	0.43	0.45	0.45	0.46	0.41	0.39

yr means. In the case of CGCM2, this variability was estimated as follows: First, for each segment ($1 \times CO_2$, $2 \times CO_2$, or $3 \times CO_2$), we computed a time average for each ensemble member and subtracted the ensemble mean time average to produce a sample of three anomalies. Variability amongst these is due to internal variability. We next computed the difference between anomalies for the $2 \times CO_2$ and $1 \times CO_2$ segments, both for anomalies within a given climate simulation and for anomalies in different simulations. The latter is justified because we do not expect statistical dependence either within or between ensemble members on time scales of the approximately 40 yr that separate the segments. This differencing produced a sample of nine anomaly differences. Similarly, nine anomaly differences were obtained from the $3 \times CO_2$ and $2 \times CO_2$ segments. The resulting sample of 18 differences was used to estimate the internal variability of the difference in time averages between segments. The variance estimate that was obtained was inflated by a factor of 3/2 to account for the dependence between anomalies that is caused by subtracting the ensemble means. For this same reason, the variance estimate was assumed to contain only 6 degrees of freedom (2 for each segment). A similar procedure was used to obtain an estimate of the HadCM3 variability, this time using differences between the $2 \times CO_2$ and 1959–79 segments, $2 \times CO_2$ and $1 \times CO_2$ segments, and the $2 \times CO_2$ and 1978–98 segments. The variance obtained from the resulting 27 anomaly differences was again scaled by a factor of 3/2. Despite being based on a larger number of differences, this variance estimate was again assumed to contain only 6 degrees of freedom because of the overlap, and proximity, in time of the 1959–79, $1 \times CO_2$, and 1978–98 segments.

We do not expect that changes in our indices of extremes will be Gaussian at the gridpoint level at individual points in time. However, the vector products used in (1) and (2) can be thought of as weighted averages taken over a large number of grid points. We assume that this averaging leads to a distribution of $\hat{a}$ that is close to Gaussian due to the central limit theorem. Unfortunately, we do not have enough samples of variability to rigorously test this assumption.

With this assumption, the 5%–95% uncertainty band for a is given by

$$(\hat{a} - t_{6,0.95}\hat{V}_{\hat{a}}^{1/2},\ \hat{a} + t_{6,0.95}\hat{V}_{\hat{a}}^{1/2}), \tag{3}$$

where $\hat{V}_{\hat{a}}$ is the variance of $\hat{a}$ [(2)], estimated by using samples of climate variability instead of model observations $\mathbf{X}$, and $t_{6,0.95} = 1.94$ is the 95th percentile of the Student's t distribution with 6 degrees of freedom. The null hypothesis that a is zero or negative can be rejected at the 5% significance level (i.e., detection of the signal at that significance level) when interval (3) contains only positive amplitude estimates.

We estimate the signal-to-noise ratio with

$$\hat{R} = \frac{|\overline{\hat{a}}|}{\hat{V}_{\hat{a}}^{1/2}}, \tag{4}$$

where $\overline{\hat{a}}$ is the average of the three estimates of a that are obtained by using each of three ensemble members as model observations in turn. If we assume that internal variations in the three ensemble members are independent of each other, then

$$E(\overline{\hat{a}}) = a, \qquad V(\overline{\hat{a}}) = V_{\hat{a}}/3, \quad \text{and}$$
$$E(\overline{\hat{a}}^2) = a^2 + V_{\hat{a}}/3, \tag{5}$$

where E denotes the expectation. Assuming that $\hat{a}$ is an unbiased estimator of a, the theoretical signal-to-noise ratio R is then given by

$$R^2 = \frac{a^2}{V_{\hat{a}}} = \frac{E(\overline{\hat{a}}^2) - V_{\hat{a}}/3}{V_{\hat{a}}} = \frac{E(\overline{\hat{a}}^2)}{V_{\hat{a}}} - \frac{1}{3}. \tag{6}$$

Equation (4) therefore provides a slightly positively biased estimate of the signal-to-noise ratio. However, (6) is impractical to use because it can produce negative estimates of R^2.

The uncertainty of estimate (4) that results from the limited sample that is available for estimating natural variability is difficult to determine using standard theoretical arguments because $\hat{R}^2$ has a noncentral distribution (we do not expect the mean of $\overline{\hat{a}}^2$ to be zero) and because $\overline{\hat{a}}$ and $\hat{V}_{\hat{a}}$ are not statistically independent. It is also not possible to assess this uncertainty with a bootstrapping (resampling) approach because the available ensemble sizes are so small. However, the resulting

signal-to-noise estimates proved to be robust to changing the sampling strategy (i.e., using only anomalies within individual climate simulations in the variance calculation). Also, results with the completely independent samples from CGCM1 were quite similar. Therefore, we believe that the resulting signal-to-noise ratios give some indication of which indices of extremes are more detectable within a model and between models.

b. Results from perfect and imperfect model studies

We now assess the detectability of the changes that occur between the $1 \times CO_2$ and $2 \times CO_2$ segments in a single ensemble member. This will help identify the indices of mean and extreme climate that show a response to anthropogenic forcing that is distinct from internal climate variability. Such changes are potentially better suited for the early detection of an anthropogenic influence on climatic extremes. We investigate the detectability of climate change in two ways. First, we use a fingerprint and model observations from the same model to explore detectability under ideal conditions (no model error in the fingerprint pattern). Second, we use fingerprints and model observations from different models to obtain an assessment of the signal-to-noise ratio that takes some account of model uncertainty.

For the perfect model assessment, we use the $1 \times CO_2$ to $2 \times CO_2$ change in one of the three available ensemble members as model observations and base the fingerprint on the mean change exhibited by the other two ensemble members. For the imperfect model assessment, we use the same ensemble member as model observations, while basing the fingerprint on two of the simulations from the other model (we use only two simulations to allow for similar contamination by internal climate variability within and between models; results are not very sensitive to this). Once model observations and a fingerprint pattern have been determined, the amplitude of the fingerprint in the model observations is determined by applying (2). This procedure is repeated three times using each ensemble member in turn in lieu of observations and using the remaining two ensemble members to derive the fingerprint.

Figure 10 displays the resulting scaling estimates (realizations of $\hat{a}$) and their associated 5%–95% uncertainty bands for the T_{min} fingerprints obtained from CGCM2. These estimates should be close to a value of 1 within a model because the model should be able to reproduce its own fingerprint with the correct amplitude. Between models, the scaling factor will be influenced by both differences in the magnitude and the pattern of simulated climate change. Figure 10 shows that both models simulate consistent amplitude estimates for boreal summer and that the CGCM2 data project with near-unit amplitude onto the HadCM3 fingerprint for the warmest night of the year. In both cases, the uncertainty range is very small compared to the best estimate, yielding very high signal-to-noise ratios.

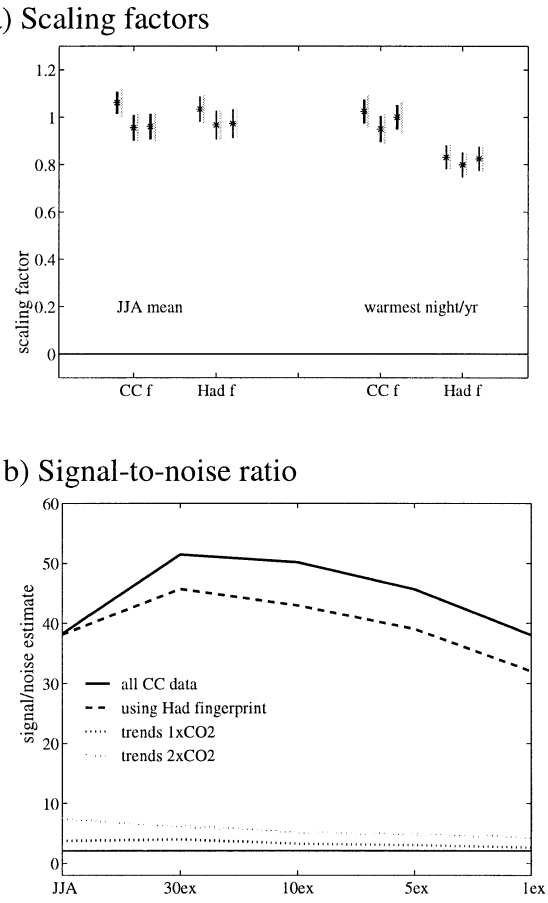

FIG. 10. Result of a model detection study for changes in seasonal mean and warm extremes of daily minimum temperature. (a) The estimated scaling factor a (asterisk) for (left) boreal summer temperature and (right) the warmest night of the year. The three realizations of $\hat{a}$ originate from the use of different members of the CGCM2 ensemble simulations to represent observed climate change. The bars labeled CC f indicate perfect model results (i.e., using CGCM2 fingerprints, thick bars), while those labeled Had f indicate imperfect model results (i.e., using HadCM3 fingerprints, thin bars). The vertical bars give the 5%–95% uncertainty range for each estimate of $\hat{a}$ (black bars using CGCM2 noise; gray bars using HadCM3 noise). (b) The resulting signal-to-noise ratios for the transition from seasonal mean (JJA) to extremes (average of the 30, 10, 5, and 1 warmest night of the year denoted by 30ex, 10ex, 5ex, and 1ex). The solid black line shows the results of the perfect model study; the dashed line shows results of the imperfect model study using fingerprints from HadCM3, and noise is based on CGCM2. The thin dotted lines give the signal-to-noise ratio for trends from over 20 yr at the present and the time of CO_2 doubling (using CGCM2 only). Changes that exceed the solid thin line should be detectable at the 5% (one sided) significance level.

The average amplitude estimate and the pooled variance of the amplitudes derived from samples of internal climate variability are then used to determine the estimate of the signal-to-noise ratio [(4)] that is displayed in Fig. 10b. As expected, the signal-to-noise ratios are large, since we in effect use the pattern of climate change that is projected to occur over 85 yr. The signal-to-noise ratios decrease only slightly when the HadCM3

fingerprint is used (dashed line), and they do not vary strongly for the transition from minimum temperature means to extremes. We also repeated the signal-to-noise analysis by replacing the segment mean indices with 20-yr within-segment trends. The results, shown as dotted lines in Fig. 10b, indicate that even the trends over the $1 \times CO_2$ segment should be detectable within the CGCM2 model. This is consistent with the marginal detection of annual mean temperature trends over that time period in observations (Hegerl et al. 1996). We have also calculated the signal-to-noise ratio for changes in warm minimum temperatures from HadCM3 data. Results are similar; except that the peak signal-to-noise ratio occurs for the summer average rather than for the 30 warmest nights of the year.

For both models, the results for seasonal mean and extreme daily cold minimum and cold and warm maximum temperatures (not shown) are qualitatively similar to the results shown in Fig. 10. Signal-to-noise ratios generally peak at seasonal means or means over the 30 warmest or coldest days, and they decrease only slightly for more extreme indices. This indicates that changes in temperature extremes that occur on average once to several times per year should be easily detectable well before the time of CO_2 doubling. In both models, signal-to-noise ratios for warm T_{max} and T_{min} means and extremes are generally higher than for cold means and extremes, although changes in cold extremes and seasons tend to be larger. This is consistent with the greater natural variability of surface air temperature in the cold season. In all cases, the signal-to-noise ratios are very similar for perfect and imperfect model studies. There is some suggestion that changes in temperature extremes might be detectable at the present time in 20-yr trends, particularly in warm means and extremes. However, this result tends to vary somewhat between the models and the indices.

Climate variability generally influences the scaling factors for precipitation more strongly than for temperature, yielding lower signal-to-noise ratios (Fig. 11). Note that when using the other model's fingerprint, the estimated scaling factors vary somewhat between cases because the HadCM3 fingerprints need to be scaled down to match CGCM2 data. Scaling factors are in better agreement for the wettest day of the year than for annual mean data. That is, the changes in each model project more strongly onto the other model's wettest day signal than onto the corresponding annual mean signal. This is consistent with the larger changes in extreme precipitation (Fig. 8) and the larger areas of consistent climate change between the models (Fig. 9).

The evolution of the signal-to-noise ratio as the indices make the transition to more extreme aspects of climate shows that anthropogenic change in precipitation can be detected best within a model (perfect model study) in averages of approximately the 10 or 5 wettest days. Note, however, that the difference between signal-to-noise ratios for the annual total and these moderately extreme events is small and probably not robust. In both models, detectability decreases somewhat as events become more extreme. Results also indicate that trends in precipitation over 20 yr are not detectable at present but that they should be detectable at the time of CO_2 doubling. In contrast, the signal-to-noise ratios obtained when using the other model's fingerprint (dashed lines) is lowest for annual mean changes and peaks relatively sharply for extremes. In the case of CGCM2 model observations, the peak signal-to-noise ratio when using HadCM3 fingerprints occurs at the average of the five wettest days. Similarly, the signal-to-noise ratio for the CGCM2 fingerprints in HadCM3 model observations occurs for the wettest day of each year. Using the other model's fingerprint results in signal-to-noise ratios that exceed those obtained in the perfect model study. This is due to a smaller projection of climate variability onto the other model's fingerprint (cf. smaller black dashed than gray dashed uncertainty bars in the left column of Fig. 11, and smaller gray solid than black solid bars).

We have analyzed the sensitivity of our results to using changes in absolute rather than relative units and to details of the analysis. In all cases, the main result of a more robust detectability for intermediately extreme precipitation has been confirmed.

These results suggest that early detection of an anthropogenic signal in moderately extreme precipitation may be less sensitive to the choice of model than the detection of such a signal in annual mean precipitation, where the pattern of change is inconsistent between the two models over a larger portion of grid points. It remains to be investigated whether this effect carries over when more models are used. However, an analysis with older CGCM1 data (with independent samples of natural variability) also showed this increased detectability for extreme rainfall between models. Moreover, the tendency toward overall increases in heavy precipitation (as opposed to a pattern of increase and decrease as in annual precipitation) has been also reported from climate change simulations with other models (cf. Meehl et al. 2000; Semenov and Bengtsson 2002; WEH). This suggests that our tentative conclusion that moderately extreme precipitation events may be more robustly detectable than annual mean precipitation may extend to other models.

5. Discussion and conclusions

We find evidence that seasonal mean temperature changes are significantly different from changes in temperature extremes over large portions of the globe (39%–66% of model grid points). A similar significant difference was found between very moderate extremes (change in the 30 most extreme days) and the most extreme day of the year. This shows that it cannot be assumed that seasonal mean values are sufficient to describe changes in extremes but that the tail of the distribution (and additionally, in high latitudes, the annual

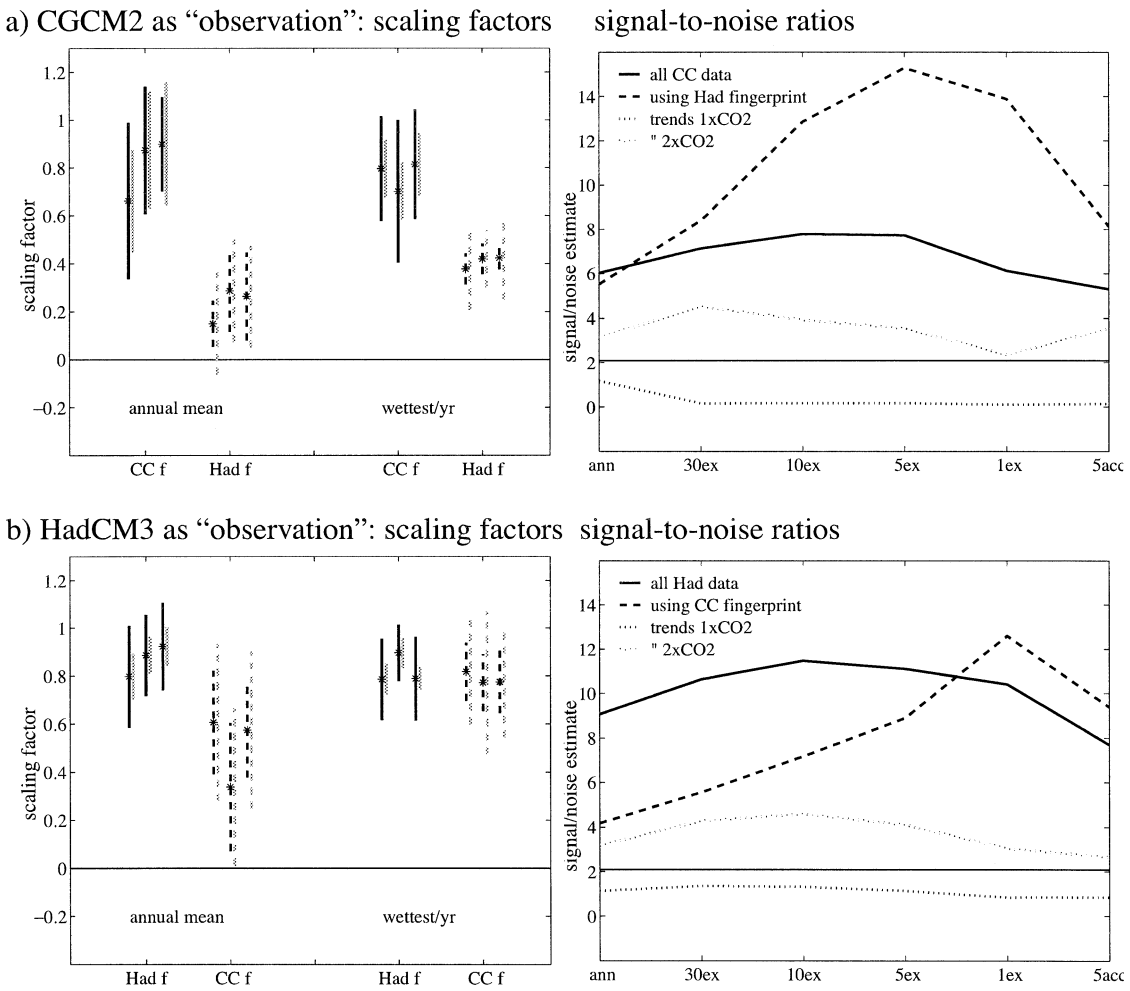

FIG. 11. As in Fig. 10, but for (left) the estimated scaling factors using model observations of precipitation changes from (a) CGCM2 and (b) HadCM3, from a perfect model study (solid uncertainy bars) and an imperfect model study (dashed bars). The left six pairs of scaling factor estimates are based on annual mean rainfall, and the right six pairs are based on changes in the wettest day of the year. The black bars are based on variability data from the same model as the model observations, and the gray bars are from the respective other model. The right column shows the signal-to-noise ratio for the transition from ann to 1ex, and the average precipitation rate during the 5 wettest days of the year (5acc). Note that changes in the wettest 5 or 1 days of the year become more detectable when the other model's fingerprint is used. The dotted lines show the signal-to-noise ratio for trends from over 20 yr at the present (black) and at the time of CO_2 doubling (gray; perfect model case only), indicating that it is not expected that short trends in precipitation can be detected at the present time but that they might become detectable in the future.

cycle) changes in both variables in a warmer world. Temperature extremes moderate in some regions of the globe and become more extreme in others relative to the change in seasonal means. Changes in temperature extremes appear robust and similar between models, although the changes in the tail of the distribution are model sensitive, particularly for warm extremes. Changes in moderately extreme temperature should be rather easily and robustly detectable.

For precipitation, the distribution generally becomes wider, with greater increases in extreme precipitation than in annual mean precipitation. Changes in annual mean precipitation are very model sensitive, with small correlations between model climate change patterns. There is greater consistency between the models for changes in annual extreme precipitation than mean precipitation for a larger fraction of model grid points (59% rather than 41%). Signal-to-noise ratios in perfect model studies decrease only slightly for changes in moderate-to-strong extreme precipitation (more so for events that occur once in 5 yr; not shown). However, in imperfect model analyses, when the other model's fingerprint is used, signal-to-noise ratios for changes in extremes are substantially greater (by a factor of 2–3) than those for changes in the mean precipitation, indicating that moderately extreme precipitation may be more suitable for an early detection of climate change given the uncertainty in fingerprints of climate change. This finding was also confirmed with data from CGCM1 (not shown). Although this result needs to be confirmed with data

from more models, we believe that it is qualitatively robust. The tendency toward overall increases in heavy precipitation (as opposed to a pattern of increase and decrease as in annual precipitation, which is more difficult to detect) has also been reported from climate change simulations with other models.

Nonetheless, our results, particularly those for precipitation, should be interpreted cautiously. The fingerprints used in this study were estimated by averaging only two to three simulations over relatively short 20-yr periods. This did not pose a serious problem in our study of the detectability of changes in temperature because the response to anthropogenic forcing in temperature at the time of CO_2 doubling is robust. This is reflected both by the estimates of signal-to-noise ratio (Fig. 11b), in the small uncertainty of the estimates in perfect model scaling factors (Fig. 11a), and the fact that these scaling factors have values very close to 1. On the other hand, perfect model scaling factors for the precipitation signals remain uncertain, and tend to be less than unity, at the time of CO_2 doubling. This indicates that the precipitation fingerprints are affected by internal variability despite the fact that they represent a mature climate change signal that has evolved over an extended period of time. This suggests that care needs to be taken to use fingerprints for the detection of precipitation changes that contain as small a contribution from internal variability as possible.

Acknowledgments. GCH was supported by NSF Grants ATM-0002206 and ATM-0296007; by NOAA Grant NA16GP2683 and NOAA's Office of Global programs; by DOE in conjunction with the Climate Change Data and Detection element; and by Duke University. PAS was funded by the United Kingdom Department for Environment, Food and Rural Affairs under Contract PECD 7/12/37. We wish to thank Tom Crowley and two anonymous reviewers for their suggestions, and Jesse Kenyon and Hank Seidel for their technical support.

REFERENCES

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

——, and W. J. Ingram, 2002: Constraints on future changes in climate and the hydrologic cycle. *Nature,* **419,** 224–232.

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

Chelliah, M., and C. F. Ropelewski, 2000: Reanalyses-based tropospheric temperature estimates: Uncertainties in the context of global climate change detection. *J. Climate,* **13,** 3187–3205.

Cubasch, U., J. Waszkewitz, G. C. Hegerl, and J. Perlwitz, 1995: Regional climate changes as simulated in time-slice experiments. *Climatic Change,* **31,** 273–304.

——, and Coauthors, 2001: Projections of future climate change. *Climate Change 2001: The Scientific Basis,* J. T. Houghton et al., Eds., Cambridge University Press, 525–582.

Easterling, D. R., J. L. Evans, P. Ya. Groisman, T. R. Karl, K. E. Kunkel, and P. Ambenje, 2000a: Observed variability and trends in extreme climate events: A brief review. *Bull. Amer. Meteor. Soc.,* **81,** 417–425.

——, G. A. Meehl, C. Parmesan, S. Changnon, T. R. Karl, and L. O. Mearns, 2000b: Climate extremes: Observations, modeling and impacts. *Science,* **289,** 2068–2074.

Flato, G. M., and G. J. Boer, 2001: Warming asymmetry in climate change simulations. *Geophys. Res. Lett.,* **28,** 195–198.

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

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

Gibson, J. K., P. Kållberg, S. Uppala, A. Hernandez, A. Nomura, and E. Serrano, 1997: ERA description. ECMWF Reanalysis Report Series 1, 72 pp.

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

——, H. Graf, and T. Osborn, 2003: Climate change and the NAO. *The North Atlantic Oscillation,* J. W. Hurrell et al., Eds., Amer. Geophys. Union, 193–210.

Gordon, C., C. Cooper, C. A. Senior, H. Banks, J. M. Gregory, T. C. Johns, J. F. B. Mitchell, and R. A. Wood, 2000: The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dyn.,* **16,** 147–168.

Groisman, P. Ya., and Coauthors, 1999: Changes in the probability of heavy precipitation: Important indicators of climatic change. *Climatic Change,* **42,** 243–283.

——, R. W. Knight, and T. R. Karl, 2001: Heavy precipitation and high streamflow in the contiguous United States: Trends in the twentieth century. *Bull. Amer. Meteor. Soc.,* **82,** 219–246.

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

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

Hegerl, G. C., H. von Storch, K. Hasselmann, B. D. Santer, U. Cubasch, 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-aerosol, and solar forced climate change. *Climate Dyn.,* **13,** 613–634.

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

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

——, G. Meehl, C. Covey, M. Latif, B. McAvaney, and R. Stouffern, cited 2003b: 20C3M: CMIP collecting data from 20th century coupled model simulations. *CLIVAR Exchanges,* **8** (26). [Available online at http://www.clivar.org/publications/exchanges/ex26/supplement/index.htm.]

Hennessy, K. J., J. M. Gregory, and J. F. B. Mitchell, 1997: Changes in daily precipitation under enhanced greenhouse conditions. *Climate Dyn.,* **13,** 667–680.

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.

Johns, T. C., and Coauthors, 2002: HCTN22: Anthropogenic climate change for 1860 to 2100 simulated with the HadCM3 model

under updated emissions scenarios. *Climate Dyn.,* **20,** doi: 10.1007/s00382-002-0296-y.

Jones, G. S., S. F. B. Tett, and P. A. Stott, 2003: Causes of atmospheric temperature change 1960–2000: A combined attribution analysis. *Geophys. Res. Lett.,* **30,** 1228, doi:10.1029/2002GL016377.

Jones, P. D., E. B. Horton, C. K. Folland, M. Hulme, D. E. Parker, and T. A. Basnet, 1999a: The use of indices to identify changes in climatic extremes. *Climatic Change,* **42,** 131–149.

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

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

——, ——, D. R. Easterling, and R. G. Quayle, 1995a: Trends in U.S. climate during the twentieth century. *Consequences,* **1,** 3–12.

——, ——, and N. P. Plummer, 1995b: Trends in high-frequency climate variability in the twentieth century. *Nature,* **377,** 217–220.

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

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

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

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

McCollum, J. R., and W. F. Krajewski, 1998: Uncertainty of monthly rainfall estimates from rain gauges in the Global Precipitation Climatology Project. *Water Resour. Res.,* **34,** 2647–2654.

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

Mitchell, J. F. B., D. J. Karoly, G. C. Hegerl, F. W. Zwiers, M. R. Allen, and J. Marengo, 2001: Detection of climate change and attribution of causes. *Climate Change 2001: The Scientific Basis,* J. T. Houghton et al., Eds., Cambridge University Press, 695–638.

National Assessment Synthesis Team, 2001: Climate change impacts on the United States: The potential consequences of climate variability and change. Report for the U.S. Global Change Research Program, Cambridge University Press, 620 pp.

Pope, V. D., M. L. Gallani, P. R. Rowntree, and R. A. Stratton, 2000: The impact of new physical parametrizations in the Hadley Centre climate model HadAM3. *Climate Dyn.,* **16,** 123–146.

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

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

Schär, C., P. L. Vidale, D. Lüthi, C. Frei, C. Häberli, M. A. Liniger, and C. Appenzeller, 2004: The role of increasing temperature variability in European summer heatwaves. *Nature,* **427,** doi: 10.1038/nature02300.

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

Stott, P. A., 2003: Attribution of regional-scale temperature changes to anthropogenic and natural causes. *Geophys. Res. Lett.,* **30,** 1728, doi:10.1029/2003GL017324.

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

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

——, and Coauthors, 2002: Estimation of natural and anthropogenic contributions to twentieth century temperature change. *J. Geophys. Res.,* **107,** 4306, doi:10.1029/2000JD000028.

Thorne, P. W., and Coauthors, 2002: Assessing the robustness of zonal mean climate change detection. *Geophys. Res. Lett.,* **29,** 1920, doi:10.1029/2002GL015717.

Trenberth, K. E., 1999: Conceptual framework for changes of extremes of the hydrological cycle with climate change. *Climatic Change,* **42,** 327–339.

——, A. Dai, R. M. Rasmussen, and D. B. Parsons, 2003: The changing character of precipitation. *Bull. Amer. Meteor. Soc.,* **84,** 1205–1217.

von Storch, H., and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research.* Cambridge University Press, 484 pp.

Wehner, M. F., 2004: Predicted twenty-first-century changes in seasonal extreme precipitation events in the Parallel Climate Model. *J. Climate,* in press.

Xie, P., J. E. Janowiak, P. A. Arkin, R. Alder, A. Gruber, R. Ferraro, G. J. Huffman, and S. Curtis, 2003: GPCP pentad precipitation analyses: An experimental dataset based on gauge observations and satellite estimates. *J. Climate,* **16,** 2197–2214.

Zwiers, F. W., and V. K. Kharin, 1998: Changes in the extremes of the climate simulated by CCC GCM2 under CO_2 doubling. *J. Climate,* **11,** 2200–2222.

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