JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 113, D05115, doi:10.1029/2006JD008091, 2008

Global changes in extreme daily temperature since 1950

S. J. Brown,[1] J. Caesar,[1] and C. A. T. Ferro[2]

Received 2 October 2006; revised 19 October 2007; accepted 7 December 2007; published 13 March 2008.

[1] Extreme value analysis of observed daily temperature anomalies from a new quasi-global data set indicates that extreme daily maximum and minimum temperatures (>98.5 or <1.5 percentile) have warmed for most regions since 1950. Changes in extreme anomalous daily temperatures are determined by fitting extreme value distributions with time-varying parameters. Changes in the distribution of anomaly exceedances above a high threshold are found to be statistically significant at the 10% level for most land areas when compared with a time-invariant distribution and with the unforced natural variability produced by a coupled climate model. The largest positive trends in the location parameter of the extreme distribution are found in Canada and Eurasia where daily maximum temperatures have typically warmed by 1 to 3°C since 1950. The total area exhibiting positive trends is significantly greater than can be attributed to unforced natural variability. For most regions, positive trend magnitudes are larger and cover a greater area for daily minimum temperatures than for maximum temperatures. The comparatively small areas of cooling are found to be consistent with unforced natural climate variability. The North Atlantic Oscillation (NAO) is found to have a significant influence on extreme winter daily temperatures for many areas, with a negative NAO of one standard deviation reducing expected extreme winter daily temperatures by ~2°C over Eurasia but increasing temperatures over northeastern North America.

Citation: Brown, S. J., J. Caesar, and C. A. T. Ferro (2008), Global changes in extreme daily temperature since 1950, *J. Geophys. Res.*, *113*, D05115, doi:10.1029/2006JD008091.

1. Introduction

[2] The impact of severe weather on ecosystems, economies and society has been witnessed throughout history with the 2003 European summer heatwave being a recent example. This heatwave caused between 22,000 and 35,000 heat-related deaths and approaching US$ 14 billion in agricultural losses [*Schär and Jendritzky*, 2004]. There is increasing evidence from climate models that at least some types of extreme events will become more frequent and more severe in the future [e.g., *Cubasch et al.*, 2001; *Kharin and Zwiers*, 2000; *Kharin and Zwiers*, 2005; *Meehl et al.*, 2005; *Tebaldi et al.*, 2006]. Such projections naturally lead to questions on whether the nature of extreme events has already changed and whether this has been due to anthropogenic activities.

[3] Because of the limited availability and homogeneity issues of observed daily data suitable for studying extreme events, fewer studies of observed changes in extremes of daily data have been reported in the literature than for other aspects of the climate system such as mean temperature. *Frich et al.* [2002] used indices of various temperature and precipitation extremes and found that the number of days where the daily minimum temperature is above the 90th percentile has increased for most land areas, with a commensurate reduction in the number of frost days and increases in growing season length. Heavy rainfall events were found to have become more frequent with wet spells producing significantly higher rainfall totals. *Kiktev et al.* [2003] generally confirmed these results with more robust statistical methods, but did not find the trend in maximum annual 5-d rainfall to be field significant. *Alexander et al.* [2006], using an updated indices data set with much greater coverage, found over 70% of the sampled land area has a significant decrease in the annual occurrence of cold nights and corresponding increase in warm nights. Most precipitation indices indicated a tendency toward wetter conditions although there is considerable spatial variability and mixed significance. *Caesar et al.* [2006] used rank statistics to estimate changes since 1946 for the 5th to 95th percentiles of daily maximum and minimum temperature (Tmax and Tmin, respectively), finding that the largest warming has occurred in Northern Hemisphere winter and spring. *Christidis et al.* [2005] found that anthropogenic emissions of greenhouse gases and aerosols have had a significant impact on the global pattern of annual maximum and minimum Tmin and annual minimum Tmax. A general characteristic of these studies is their consideration of only relatively moderate extremes on account of either the simple empirical techniques used (e.g., percentiles from rank sta-

[1]Met Office Hadley Centre for Climate Change, Exeter, UK.
[2]School of Engineering, Computing and Mathematics, University of Exeter, Exeter, UK.

tistics or maximum/minimum within a given period, usually 1 year) or that the available data has allowed (e.g., the indices of *Frich et al.* [2002] and *Alexander et al.* [2006]). If the area of interest is extreme events with multiyear return periods then other approaches can be beneficial. In this paper stationary and nonstationary extreme value distributions are fitted to observations of daily maximum and minimum temperatures to determine whether such extreme daily temperatures have changed since 1950.

2. Observed and Model Data

[4] Long-term gridded monthly temperature data sets with near global coverage have been available to climate researchers for many years [e.g., *New et al.*, 2000] but data on daily timescales has been limited to regional areas for data sets of sufficient length for climate studies [e.g., *Janowiak et al.*, 1999]. Here we use the new gridded daily temperature data set with quasi-global coverage of daily maximum and minimum temperatures of *Caesar et al.* [2006] but extended to cover the period January 1950 to December 2004. A network of ~2500 stations has been gridded onto a 2.5° latitude by 3.75° longitude grid using an angular-distance weighting algorithm following *New et al.* [2000]. To avoid biases, particularly in areas of high-elevation variability, anomalies are gridded. Daily anomalies are calculated separately for temperature minima and maxima as the difference between each daily temperature and a normal value. The normal value varies with the day of year and is computed as the average of those daily temperatures recorded between 1961 and 1990 that are within two calendar days of the day of interest. The anomalies are therefore the result of removing a mean annual cycle. Stations must have at least 20 years of data available within the reference period and at least 350 daily normal values out of 366 calendar days, otherwise they are excluded from further consideration. As will be seen from Figure 2 there is good coverage for North America, Eurasia, South Africa and Australia but poor coverage for parts of Africa, the Middle East, India and South America. There are also very few observations for Antarctica and Greenland so they are omitted from this study. The time-varying nature of the station coverage has the potential to introduce nonclimatic signals into this data set as is the case with most gridded observed data. However, limiting stations only to those which span the entire record would reduce the spatial coverage to only a few small regions. These potential limitations in the observed data need to be borne in mind when interpreting the results.

[5] An alternative approach to gridding in this way would be to determine the statistics of extreme temperatures of the individual stations and then grid these statistical measures. This may produce a better representation of extremes at a given point in space as this would avoid interpolation between individual events of different severity from different sites. However, we wish to analyse the extremes from our data set, which was constructed primarily to aid climate model validation, and detection and attribution studies. Climate models only produce temperatures averaged over entire grid boxes and so the construction of the gridded observations is designed to best match this. How future projections at the climate model gridbox scale are related to point locations is an important question but is outside the scope of this paper.

[6] The maximum and minimum tails of the daily maximum and minimum temperatures (xTmax, nTmax, xTmin and nTmin, respectively) are analysed. The focus of this paper is on extreme temperature anomalies that could occur on any day of the year (ANN), that is pooling data from all days of the year, to provide a general description of changes in extremes temperature anomalies. In regions where there is a strong seasonality in the variance and possibly higher moments of the raw temperatures ANN results may be dominated by extremes in a single season. To assess this, seasonal data are also analysed, the results of which are presented in Figures S1–S8.[1] The North Atlantic Oscillation (NAO) has a significant impact on winter temperatures in the Northern Hemisphere [*Hurrell and van Loon*, 1997]. Its influence on extreme temperatures is investigated in a separate section.

[7] Ideally, there would be sufficiently long observational records to determine whether or not any observed trends are outside what might be expected from natural variability. This is not possible here, so data from the HadCM3 atmosphere-ocean general circulation model [*Gordon et al.*, 2000] is used for determining the level of unforced internal natural climate variability (i.e., that which comes from the chaotic nature of the climate system rather than from naturally occurring climate forcings such as volcanoes). Daily temperatures from 1500 years are taken from a control climate simulation (CNTRL) after the model has reached near equilibrium and where the levels of CO_2 are kept constant at preindustrial levels.

3. Extreme Value Methodology

[8] The use of extreme value distributions is increasingly common in climate studies [e.g., *Zwiers and Kharin*, 1998; *Kharin and Zwiers*, 2000; *Wettstein and Mearns*, 2002; *Zhang et al.*, 2004; *Kharin and Zwiers*, 2005] although this has generally been restricted to Generalized Extreme Value (GEV) distributions applied to the most extreme value within a period, typically annual maxima or minima. Although more straightforward to apply, the use of annual extremes results in extensive data reduction and may not fully capture all extreme events. For example the two largest events in a record may occur in the same year. Extremal analysis can be adapted to include the top r values, which alleviates these problems to some degree [*Zhang et al.*, 2004]. An alternative, threshold approach is adopted below. Furthermore, few studies have investigated trends in characteristics of observed extremes, *Nogaj et al.* [2006] being a rare example. As climate change accelerates through this century [*Cubasch et al.*, 2001] an assumption of a stationary climate will become untenable if the evolution of extremes is to be correctly described. *Kharin and Zwiers* [2005] demonstrated the nonstationary nature of temperature and precipitation extremes that might occur if future emissions follow the IPCC SRES A2 scenario [*Nakienovic and Swart*, 2001]. The uncertainty in the future projection of extremes, however, seems to be dominated by the uncertainty in

[1] Auxiliary materials are available in the HTML. doi:10.1029/2006JD008091.

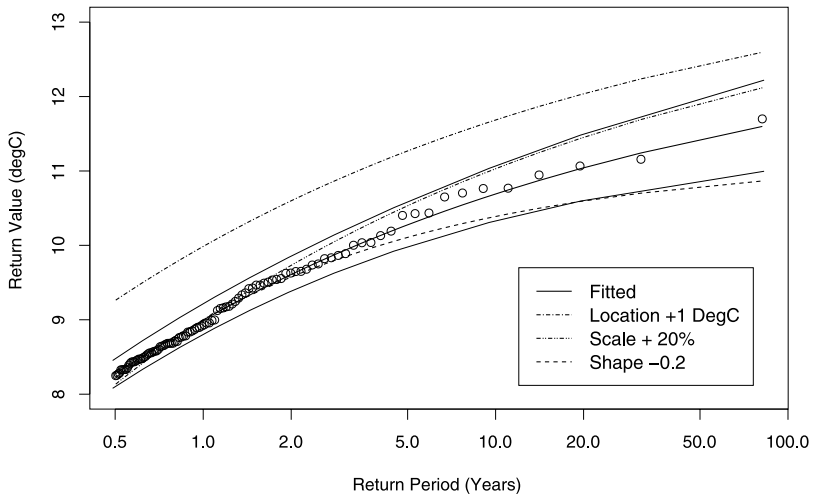

Figure 1. Return-level curve derived from fitting marked point process extreme value distribution to the top 1.5 % of daily Tmax data for the grid box containing London. Circles represent empirical return levels, the heavy solid line represents the fitted return values, and light lines represent the corresponding 5−95% uncertainty range. Nonsolid lines represent return-level curves where distribution parameters are adjusted according to the legend.

modelling the climate [*Kharin et al.*, 2007]. This paper asks if there is any evidence of nonstationarity in the amounts by which observed temperature anomalies from the annual cycle fall above high or below low thresholds.

[9] To include more of the extremal data in the analysis, as opposed to just annual maxima or minima, a peaks-over-threshold model is used to describe all exceedances above a high threshold, u. The exceedances are assumed to occur according to a Poisson process, and the excesses of those exceedances above the threshold are assumed to follow a Generalized Pareto distribution [*Katz et al.*, 2002]. The expected number of exceedances per year above any level x, conditional on x being greater than u, is written

$$\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}, \tag{1}$$

where μ, σ, and ξ are termed the location, scale and shape parameters, respectively [*Coles*, 2001; *Katz et al.*, 2005]. This formulation of the marked point process (MPP) model ensures that the scale parameter is invariant to the threshold u, and that the parameters are equivalent to those in the GEV distribution for annual maxima. The MPP parameters are estimated by maximum likelihood [*Coles*, 2001] in preference to other estimation procedures to enable both the inclusion of covariates for modelling trends in the parameters, and the imposition of constraints that ensure all observed exceedances are feasible under the estimated model. From such a model it is useful to determine the estimated exceedance value for a given probability, commonly referred to as the return level (z_m) experienced on average once every m years [*Coles*, 2001]:

$$z_m = \begin{cases} \mu-(\sigma/\xi)\left\{1-\left[-\ln\left(1-\frac{1}{m}\right)\right]^{-\xi}\right\} & \xi \neq 0 \\ \mu-\sigma\ln\left[-\ln\left(1-\frac{1}{m}\right)\right] & \xi = 0 \end{cases} \tag{2}$$

[10] Figure 1 gives an example of such a return-level curve for an MPP model fitted (as described later) to xTmax data for the gridbox containing London. Empirical return levels and probabilities of the actual data are included as points. To illustrate the role of different distribution parameters, the curves obtained when the location, scale and shape parameters are altered unilaterally are superimposed. The location parameter is analogous to the mean of a normal distribution, so increases in μ uniformly shift the distribution to higher values, increasing all extremes equally. Whereas σ and ξ determine the rate with which the magnitude of extremes alters with rarity.

[11] The MPP model described above assumes stationary extremes. Nonstationarity is modelled by allowing the parameters to depend linearly on time throughout the data record. Following *Kharin and Zwiers* [2005] and *Coles* [2001], this gives

$$\begin{aligned} \mu_t &= \alpha_0+\alpha_1 t \\ \sigma_t &= \exp(\beta_0+\beta_1 t), \\ \xi_t &= \gamma_0+\gamma_1 t \end{aligned} \tag{3}$$

where the covariate t is the date during the observational period scaled to range from 0 to 1. More complex time dependence was not investigated as linear forms are deemed an adequate first-order approximation to any nonstationarity and avoid finding nonlinearity that is just a product of natural internal climate variability. Furthermore the observational period is relatively short so there is a danger of over fitting if more complex time dependence is introduced. Applying this approach to future projections of climate may require more sophisticated covariates, such as global mean temperature, as future temperature changes are unlikely to be linear [*Cubasch et al.*, 2001]. The inclusion of NAO as a covariate is described later.

[12] The threshold, u, is also replaced by a threshold that changes linearly in time to ensure that exceedances are obtained throughout the record. This time-dependent thresh-

old is obtained separately for each grid point by fitting a linear regression through all temperature anomalies recorded at the grid point and then shifting the fitted regression curve uniformly until it is exceeded by a predetermined proportion of the data. Results presented below correspond to the top 1.5 % of the data. For the lower tails of the distributions the data are negated prior to this extremes selection. The sensitivity to threshold choice is determined by repeating the analysis with thresholds of 0.75% and 3%.

[13] The statistical analysis is simpler if exceedances are serially independent of one another. This is achieved approximately by retaining only the maximum (or minimum) temperatures within 10-d, nonoverlapping windows for each grid point. The sensitivity to the width of this declustering window is determined by repeating the analysis using widths of 5 and 30 d.

[14] Two approaches are adopted for assessing the significance of nonstationarity in the MPP parameters. The first determines whether or not the fit to the data achieved with one model is statistically significantly better than that achieved with a simpler model by means of a likelihood ratio test conducted at the 10% level. For example, the model with a time-dependent location parameter is deemed to be a significantly better description of the data than the same model with a time-invariant location parameter if the deviance between the models exceeds the upper 10% quantile of the chi-squared distribution with one degree of freedom [*Coles*, 2001].

[15] The location parameter is the only parameter for which significant nonstationarity is found. No time or NAO dependence for scale or shape, or combinations of the three parameters, is found to be a significant improvement from the stationary MPP model or the location-only dependent model other than for small regions which are not field significant. Although more complex changes in the distributions of extremes have been discussed in the literature [*Brabson et al.*, 2005; *Kharin and Zwiers*, 2005] it is not possible by the approach used here to find any such changes from the available observed data. Therefore, only four MPP models are considered further: a stationary model (STAT), the location parameter having a linear trend (L-TREND), the location parameter having NAO dependence (L-NAO) and the location parameter having a linear trend and NAO dependence (L-TREND-NAO).

[16] The second significance test follows the approach of *Karoly and Wu* [2005] to determine field significance. A number of grid boxes may exhibit warming trends purely due to unforced natural variability and these can have large spatial coherence. The test uses 1500 years of daily data from CNTRL divided into thirty 50-year periods from which time-dependent parameters are calculated in the same manner as the observed data. The fraction of grid boxes with significant trends is calculated for each of the thirty samples from which a 5–95% confidence interval is estimated for the area which may exhibit significant trends on account of unforced natural variability as simulated by HadCM3. However, unlike *Karoly and Wu* [2005] resampling of CNTRL to increase the number of 50-year samples and increase the precision of the confidence intervals is not practicable as the computational expense is prohibitive. The result of such a small sample of natural variability will mean that the significance tests for Table 2 are less robust. However, all but four of the areas that are deemed significantly different from natural variability are larger than the test threshold by at least 10% (in percentage area terms) suggesting a degree of robustness in the results.

[17] The accuracy with which each MPP model represents the distribution of the threshold excesses is determined by a Kolmogorov-Smirnov goodness-of-fit test. Specifically, suppose that the temperature X_t at time t exceeds the threshold u_t, and the MPP parameters at time t are (μ_t, σ_t, ξ_t). If the MPP is a good model for the data then the transformed exceedance

$$1 - \left[1 + \xi_t\left(\frac{X_t - u_t}{\beta_t}\right)\right]^{-1/\xi_t},$$

where $\beta_t = \sigma_t + \xi_t(u_t - \mu_t)$, will be approximately uniformly distributed on the interval (0, 1). Transforming all the exceedances in this way, we can test the model fit by comparing their empirical distribution to a uniform distribution. The Kolmogorov-Smirnov test provides a formal comparison. In practice, we must use estimates of the MPP parameters to transform the exceedances, and this affects the critical values of the test statistic, which are therefore determined by simulation as follows. Data are simulated from an MPP model with parameters equal to the parameter estimates. The model is then refitted to the simulated data, the exceedances transformed and the test statistic computed. This is repeated 10 000 times and the upper P% quantiles of the statistics define the P% critical values of the test.

4. Observed Changes in Extreme Temperatures

[18] The location, scale and shape parameters from L-TREND for xTmax, nTmax, xTmin and nTmin are plotted in Figure 2 for annual daily anomalies (ANN) and all land points with sufficient data for a retrieval. The location parameter exhibits a general progression to higher values with higher latitudes, with values typically ~4°C at the equator rising to ~16°C nearer the pole. Differences in location reflect differences in variability between regions rather than absolute extreme temperatures as the data are anomalies from the annual cycle. Daily temperature variability can be larger at higher latitudes and in continental interiors. This can be due to several mechanisms, such as air masses from quite different climatological regions influencing a region hence increasing variability. Also, at high latitudes strong inversions and their demise can allow large and rapid changes in temperature and hence high variability. For coastal regions, the sea can moderate temperature variability because of its high heat capacity. This, together with the prevailing winds, could be a contributing factor to the east-west progression of the location parameter seen for North America and for Europe into Russia. With North America this eastward influence appears to be contained by the Rocky Mountains. Both of the lower tails, nTmax and nTmin, show a greater range of location values with larger regions of high location values in the north of the Northern Hemisphere continents. The scale parameter has a similar spatial pattern to the location parameter which may be due

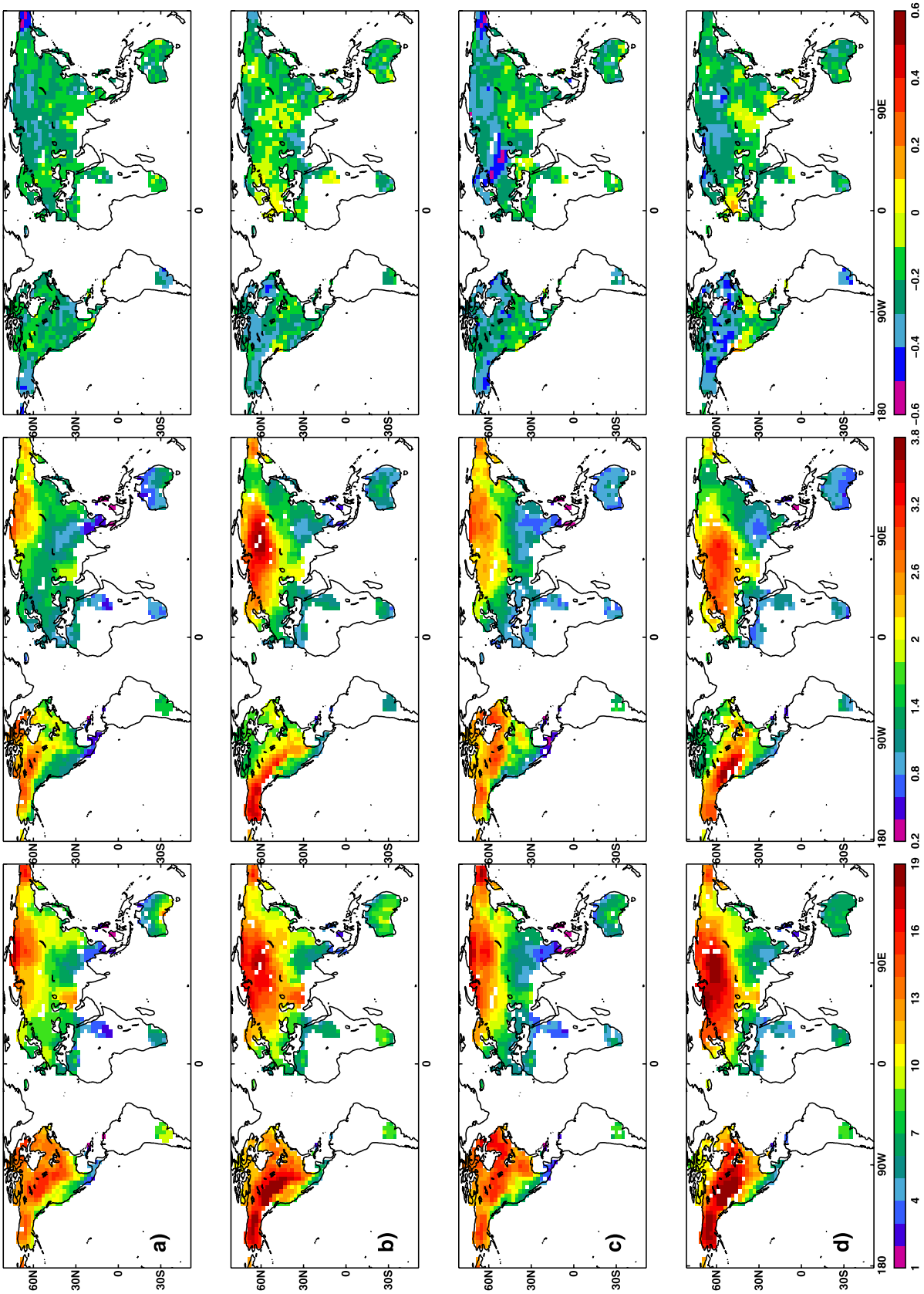

Figure 2. (left) Location parameter, (middle) scale parameter, and (right) shape parameter for marked point process extreme distribution (L-TREND) fitted to annual daily anomalies (ANN) (a) xTmax, (b) nTmax, (c) xTmin, and (d) nTmin. Points failing the goodness-of-fit test at the 1% level are masked as missing.

Table 1. Averaged Marked Point Process Distribution Parameters for the Subregions[a]

	xTmax				nTmax			
Region	Location	Location Trend	Scale	Shape	Location	Location Trend	Scale	Shape
Northern Africa	5.3	1.7	1.0	−0.2	6.4	0.5	1.1	−0.2
Southern Africa	5.6	1.8	0.8	−0.1	8.2	−0.5	1.2	−0.2
Central Asia	10.0	1.5	1.4	−0.2	12.6	1.3	2.0	−0.2
Australia	6.9	0.9	0.9	−0.2	7.7	0.3	1.0	−0.2
Europe	8.1	1.1	1.1	−0.2	9.9	0.9	1.7	−0.1
North America	11.6	0.3	1.6	−0.2	13.5	1.2	1.9	−0.2
Southern South America	9.1	1.2	1.4	−0.3	8.3	−0.6	1.1	−0.2
South Asia	6.9	0.7	1.1	−0.2	7.9	0.8	1.2	−0.2
Arctic	13.1	1.2	2.0	−0.2	14.2	1.9	2.2	−0.2
	xTmin				nTmin			
	Location	Location Trend	Scale	Shape	Location	Location Trend	Scale	Shape
Northern Africa	4.7	1.2	0.9	−0.1	5.3	0.3	0.9	−0.2
Southern Africa	5.2	1.2	0.8	−0.1	6.2	−0.5	0.9	−0.2
Central Asia	10.0	1.9	1.7	−0.2	12.8	2.5	1.9	−0.2
Australia	6.2	0.9	0.8	−0.2	6.6	0.2	0.8	−0.2
Europe	7.7	1.4	1.1	−0.2	11.1	1.6	1.9	−0.2
North America	11.8	0.5	1.7	−0.2	13.7	1.9	1.9	−0.2
Southern South America	8.2	2.1	1.3	−0.3	8.3	−1.2	1.0	−0.3
South Asia	5.0	1.1	0.9	−0.2	6.9	1.8	1.1	−0.1
Arctic	14.0	1.6	2.2	−0.3	14.0	2.5	1.8	−0.2

[a]Northern Africa [20W, 10S, 45E, 36N], southern Africa [20W, 40S, 45E, 10S], central Asia [45E, 36N, 180E, 65N], Australia [105E, 45S, 180E, 10S], Europe [20W, 36N, 45E, 65N], North America [165W, 15S, 30E, 65N], southern South America [115W, 55S, 30W, 10S], South Asia [45E, 10S, 160E, 36N], and Arctic [180W, 65N, 180E, 90N].

simply to statistical correlation between the location and scale parameter estimates. The smallest scale values ($\sim$0.4°C) are in the tropical maritime continent rising to $\sim$3°C in the highest latitudes.

[19] For the shape parameter there is less spatial variation with no strong, large-scale coherent patterns. The shape parameter for nTmin shows small areas of western Europe and western North America which are positive. Further, visual inspection of the data for these areas suggests the selected extremes form two populations: ''less'' extreme cold anomalies and ''more'' extreme cold anomalies. Such composite distributions are not catered for by the extreme distribution used here which assumes that all samples come from the same parent distribution and indicates that for these regions the thresholds may be too small. When the data are divided into seasons the occurrence of positive shape values is much reduced or even removed for these regions, further suggesting the positive shape values are a result of mixing extreme temperature anomaly events which have arisen from different weather types. Given this is found for only a small number of points it is not seen as a major issue for the adopted methodology. The area average of L-TREND parameters for selected regions are given in Table 1 for the four distribution tails. The shape parameter can be seen to be somewhat invariant at these spatial scales with most regions having values of −0.2. The scale parameter shows more variation although there is a degree of similarity between upper or lower tail values, as might be expected if the same weather is responsible for either hot events in Tmax and Tmin or cold events. The greatest spatial variation is seen with the location parameter.

[20] Corresponding seasonal results can be found in Figures S1–S8 but a short discussion of seasonality is provided here. We find the shape parameter exhibits little coherent variation through the seasons although the increased spatial noise, due to the reduced volume of data arising from separating into seasons, may be obscuring any seasonality. Location and scale both have considerable seasonality. In JJA the retrieved location parameters have a much reduced range of values and equator/pole gradient when compared with DJF, which has the largest range. Similar seasonality is seen with the scale parameter. This indicates that the largest extreme temperature anomalies occur during DJF and the smallest during JJA for the Northern Hemisphere.

[21] The retrieved location trend for L-TREND and ANN data is plotted in Figure 3 for xTmax, nTmax, xTmin and nTmin. This shows that most land areas have a significant positive trend in the location parameter and hence where there has been a significant positive trend in extreme tails of the temperature distributions over the observed period. Significance presented in Figure 3 is determined by the likelihood ratio test against STAT on a grid point basis at the 10% significance level. Similar patterns of significance are found when tested against the unforced natural variability from CNTRL. Greatest trend values and largest areas of significant positive trends are located in Canada, Alaska and Eurasia with values typically 1 to 3°C over the 55 years of the data. Smaller areas of significant negative trends are found in the United States, Mexico and eastern coast of Canada for xTmax and xTmin and in Uruguay, South Africa, Egypt and parts of Australia for nTmax and nTmin. The size and magnitude of these areas with negative trends are not found to be significantly different from those found in the CNTRL simulations. Both nTmin and xTmin have significant positive trends of greater magnitude and over a greater land area than the corresponding Tmax data.

[22] The trend in the location parameter seen over North America is similar to the pattern and magnitude of changes

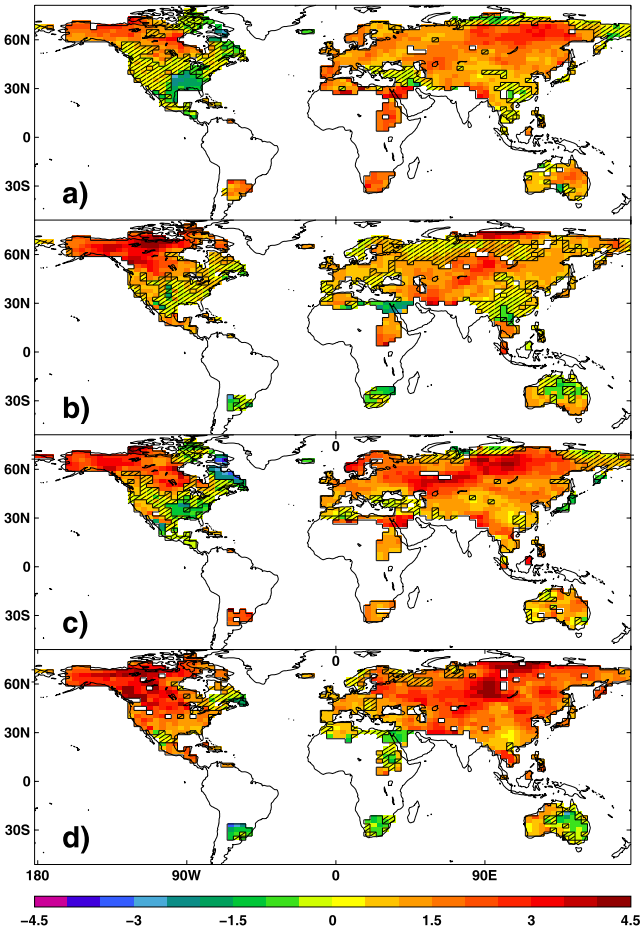

Figure 3. Trend in location parameter derived from time-dependent marked point process extreme distribution (L-TREND) fitted to (a) xTmax, (b) nTmax, (c) xTmin, and (d) nTmin. Values correspond to the change over the observation period 1950–2004 of all extreme temperatures more extreme than the data selection threshold; units are °C. Unshaded areas have significant trends rejecting the hypothesis of no trend at the 10% significance level derived from a likelihood ratio test with the corresponding stationary marked point process extreme distribution. Points failing the goodness-of-fit test at the 1% level are masked as missing.

found for the 90th and 10th percentiles of daily Tmin and Tmax by *Robeson* [2004]. The largest increases in Tmin and Tmax are similarly found to be in Northern Hemisphere winter and spring (not shown, see Figures S1–S8) and occurring in the western and northwestern parts of the continent. The strongest period showing extremes cooling for some regions is in SON as with *Robeson* [2004], however the inclusion of all percentiles in the cluster analysis of *Robeson* [2004] probably explains the different patterns of extremes cooling between the two studies. The percentile trend analysis of *Caesar et al.* [2006] using the same data set as used here but focussing on less extreme temperatures (5th to 95th percentile) and that of *Alexander et al.* [2006] which used daily extreme indices (occurrence of warm or cold days or nights above/below the 90th/10th percentiles) are in broad agreement with the results shown here. Both these studies find that the greatest warming of temperature extremes occurs over Eurasia and that this has predominately occurred in DJF and MAM as found here (not shown, see Figures S1–S8). *Nogaj et al.* [2006] apply an alternative nonstationary extreme value methodology to NCEP reanalysis data restricted to the North Atlantic region. For the land points in their domain there is considerable agreement with a decrease (increase) in warm JJA extremes for the southern (northern) east coast of North America and extremes increasing for most of Europe. For DJF we find more extensive indication that extreme cold events are getting warmer (see Figures S1–S8). Possible reasons for these differences could include the use of daily mean temperatures from the reanalysis and the limitations of the land surface schemes in the assimilation model resulting in discrepancies with the observations [*Kharin et al.*, 2005].

[23] The sensitivity of these results to the declustering window width was found to be small. Repeating the analysis using 5-d and 30-d windows produces very similar location trend patterns with the only difference being a larger number of retrievals failing to converge to a solution with the numerical solver for the 30-d window due to the smaller volume of data available for fitting. The magnitudes of the retrieved trends were found to be within one standard error of the original estimates for most grid boxes. Likewise sensitivity to threshold level was also found to be small. Thresholds of 0.75% and 3.0% produce very similar location trend patterns although substantially more retrievals fail to converge due to lack of data with the 0.75% threshold. Magnitudes of the trends from the different thresholds were also found to be within one standard error of the original estimates for most grid boxes.

[24] The average magnitude of the observed positive trends in the location parameter together with the area covered by positive trends, are presented for selected regions in Table 2. Assessment of field significance with respect to unforced natural climate variability is made through comparison with equivalent values from CNTRL as outlined earlier. Trends and areas not significantly different from that which might be expected from the model's unforced natural variability, at the 10% significance level, are italicized in Table 2. All regions have positive trends and nearly all their averages are greater in magnitude than can be expected from unforced natural variability for the four measures of extreme temperatures. Exceptions are nTmax and nTmin extremes for southern Africa and nTmax extremes for Europe. Considering the area with positive trends within these regions, 25 out of the 36 measures have areas which are greater than expected from unforced natural climate variability at the 10% significance level. The area of positive trends for all four temperature tails for Europe, central Asia, South Asia and the Arctic are significantly greater than expected from a stable climate. If the control climate of HadCM3 is an accurate measure of the unforced natural variability of daily temperature extremes the results of Table 2 suggest that the observed changes in extreme temperatures are being driven by external forcing of the climate system. The similarity between the significance mask derived from the likelihood ratio test and the CNTRL data (not shown) provides additional confidence that the detected changes in extreme temperatures are real. These

Table 2. Percentage Land Area Having Positive Trends and Corresponding Average Positive Trend Magnitude for Subregions Defined in Table 1[a]

	Northern Africa		Southern Africa		Central Asia		Australia		Europe	
	%		%		%		%		%	
Extreme	Area	Trend	Area	Trend	Area	Trend	Area	Trend	Area	Trend
xTmax	94	1.83	*100*	1.78	96	1.57	*89*	0.99	92	1.28
nTmax	*71*	1.28	*30*	*0.27*	95	1.39	*58*	0.86	97	*0.96*
xTmin	91	1.37	100	1.23	97	1.99	89	1.03	90	1.56
nTmin	*64*	0.76	5	*0.40*	100	2.48	*57*	0.87	100	1.64

	North America		Southern South America		South Asia		Arctic	
	%		%		%		%	
Extreme	Area	Trend	Area	Trend	Area	Trend	Area	Trend
xTmax	*64*	1.01	*94*	1.32	81	0.89	79	1.64
nTmax	84	1.50	0	-	76	1.19	96	1.95
xTmin	*62*	1.47	*100*	2.09	97	1.20	84	1.98
nTmin	94	2.07	0	-	99	1.79	100	2.49

[a]Trends and areas not significantly different from HadCM3's unforced natural variability, at the 10% significance level, are italicized.

results are consistent with the findings of *Christidis et al.* [2005] who detected the warming influence of increased greenhouse gases on the warmest and coldest night and coldest day of the year (here xTmin, nTmin and nTmax) although not for the warmest day (here xTmax).

[25] An important question, particularly for adaptation planning, is whether the distributions of regional daily temperature extremes will change at a different rate to those of daily mean temperatures. For the future, climate model simulations suggest that for some regions extreme changes could be considerably different to the changes in the mean [*Kharin and Zwiers*, 2005; *Clark et al.*, 2006]. To see whether this has already happened, Figure 4 shows the differences between the linear trend in the mean (derived from all daily anomalous temperature data) and the trend in the ANN location parameter. The statistical significance of these differences is determined at the 10% level via a bootstrap approach. The original data at each grid-point are resampled in blocks of 30 d to preserve serial correlation, and the mean and extreme trends are recalculated. The p-value is the proportion of resampled trend differences greater in magnitude than the trend difference in the original data. Because of computing limitations, the test was restricted to points with absolute trend differences greater than 0.5°C over the period and to 400 resamples.

[26] For most of the globe, changes in extremes are indistinguishable from change in the mean over the period 1950 to 2004. However, there are a substantial number of regions showing significantly differing trends. For example, in North America the cold tails have generally warmed more than the mean, particularly nTmin, and a similar but opposite and smaller response for the warm tails. The Southern Hemisphere regions, Uruguay, southern Africa and Australia show the opposite patterns. Eurasia and northern Africa have a more complicated picture with opposing patterns of positive and negative differences between xTmax and nTmax. For Tmin there are large parts of central Eurasia that have both Tmin tails warming more than the mean. Assessment of the causes for these differences is left for future work but potential candidates include local feedbacks which may enhance or reduce the effectiveness of the processes causing changes in extremes with respect to the mean, such as changes in soil moisture [*Clark et al.*, 2006].

5. Influence of NAO on Extreme Temperatures

[27] The North Atlantic Oscillation (NAO) and the Arctic Oscillation (AO), both patterns of large-scale, low-frequency natural atmospheric variability in the Northern Hemisphere, have been shown to have a significant influence on average winter surface temperatures [e.g., *Hurrell and van Loon*, 1997; *Thompson and Wallace*, 1998] with positive NAO indicating strong zonal flow over the North Atlantic bringing warmer, moister and stormier weather to Europe. However, their influence on extremes has been restricted to

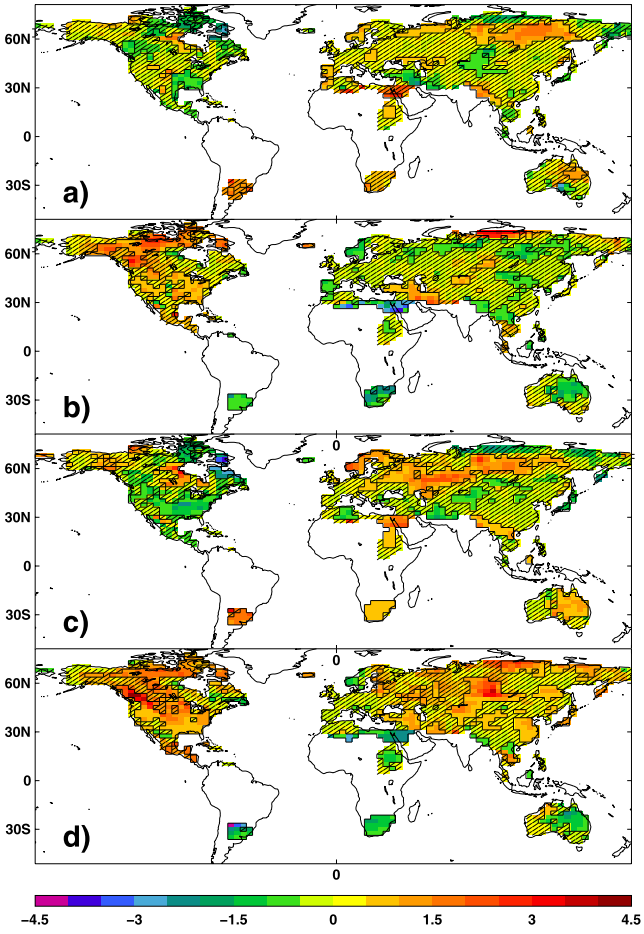

Figure 4. Difference between the location parameter trends of Figure 3 and trend in all daily temperatures: (a) xTmax - mean of Tmax, (b) nTmax - mean of Tmax, (c) xTmin - mean of Tmin, and (d) nTmin - mean of Tmin. Anomalies from all days of the year are used; units are °C. Significance is determined through bootstrapping (see section 3) with shaded areas failing to reject the null hypothesis of equal trends at the 10% level or where trends are within 0.5°C over the data period. Points failing the goodness-of-fit test at the 1% level are masked as missing.

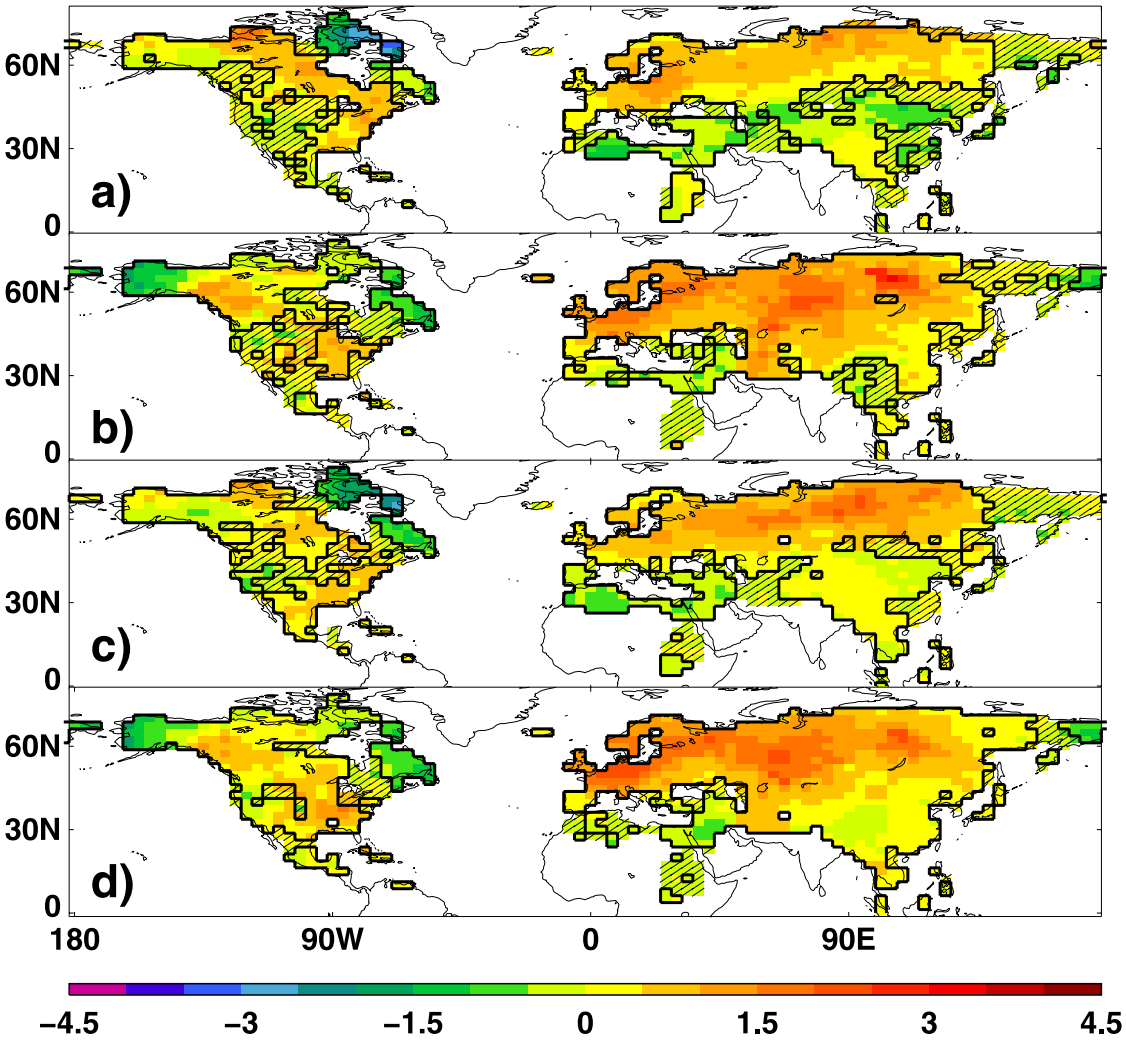

Figure 5. The effect the range of the North Atlantic Oscillation (NAO) has had on extreme DJF temperatures during the period 1950 to 2004 as derived from a marked point process model of extreme values of daily temperatures where the location parameter depends on both time and NAO (L-TREND-NAO) (a) xTmax, (b) nTmax,(c) xTmin, and (d) nTmin. Units are °C. Unshaded areas reject the hypothesis of no NAO effect and no trend at the 10% significance level derived from a likelihood ratio test. Points failing the goodness-of-fit test at the 1% level are masked as missing.

regional studies [e.g., *Wettstein and Mearns*, 2002; *Higgins et al.*, 2002; *Gong and Ho*, 2004]. Here, the influence of the NAO on the location parameter is determined through the introduction of an NAO index (defined as the standardized seasonal mean pressure difference between the Azores and Iceland) as a covariate in the estimation of the location parameter (L-NAO) and also where the location parameter depends on both NAO and time (L-TREND-NAO) to give location parameters that are, respectively:

$$\begin{aligned} \mu_t &= \alpha_0 + \alpha_2 nao \\ \mu_t &= \alpha_0 + \alpha_1 t + \alpha_2 nao \end{aligned} \tag{4}$$

These are tested for significance using the likelihood ratio test against STAT and L-TREND.

[28] For most regions there is little or no NAO influence on extreme daily temperatures found for ANN or for seasons other than DJF (not shown). For DJF, L-TREND-NAO is found to produce the largest number of grid boxes with significantly better fits with respect to STAT than L-TREND. The areas of significance for L-TREND and L-NAO with respect to STAT indicate regions for which both covariates have an influence, as for Europe and central Eurasia, in addition to regions where only one has a significant contribution, as with L-TREND in central and western North America (not shown). The introduction of an NAO covariate does not alter the retrieved temperature trends (α_1 in equations (2) and (3)) for most areas, the largest difference occurring for western Europe and northern Eurasia with coefficients for nTmax and nTmin being up to 0.4°C/decade lower if the NAO is included in the MPP model. Similarly the introduction of a trend covariate does not have a significant effect on the magnitude of the NAO coefficient (not shown).

[29] Figure 5 plots α_2 of L-TREND-NAO in equation (4), representing the effect NAO has had on extreme DJF

anomalous daily temperatures for all four distribution tails. The general patterns of influence correspond well with the patterns found by *Hurrell and van Loon* [1997] for mean winter temperatures. Positive NAO brings warmer temperature extremes for most of northern Eurasia and cooling, albeit weaker, for northern Africa, the Middle East and southern Eurasia. The cooling for northeastern North America is also reproduced. The impact of NAO on the rest of North America agrees less well with *Hurrell and van Loon* [1997] with a complicated pattern of warming and cooling being found here, and cooling with increased NAO found for Alaska, the opposite of that found by *Hurrell and van Loon* [1997]. These regional differences could be due to the use here of pressure data from the Azores rather than Lisbon in the formation of the NAO index.

[30] Unlike *Hurrell and van Loon* [1997] which looked at the influence of NAO on mean winter temperatures, here we have four extreme temperature measures. Figure 5 indicates that these measures do not always respond equally to NAO. The differences between the NAO coefficients for each pair of tails (not shown) are tested for statistical significance with a t-test at the 10% level. The test assumes independence between the extremes in the two tails, so may be liberal in case of dependence; however, any serial correlation in the data will make the test conservative. Notwithstanding these limitations, the NAO response of xTmax in Eurasia is seen to be significantly less positive in the north and more negative in the south. Alaska has a large negative response for nTmax and nTmin, as does western Europe, with significantly larger coefficients for the cold tails of the distributions than for the warm tails. These differences between the different temperature measures are potentially due to the different weather which causes extremes in each tail. The lower tail of nTmin is generally associated with cold winter nights and the weather causing such extremes (e.g., an arctic outbreak over Europe) will be different from that which causes extreme warm winter days for xTmax (e.g., a strong southerly flow bringing warm air from lower latitudes). The similarities in NAO influence between nTmin and nTmax suggest that similar weather types are causing extremes in both tails.

[31] These results are in qualitative agreement with *Wettstein and Mearns* [2002] who found high AO conditions increased (decreased) expected mean winter Tmin and Tmax for the region around New York (eastern Canada). *Higgins et al.* [2002] find a positive (negative) AO relationship for the number of warm (cold) days for central and eastern United States. These relationships correspond well with the NAO influence shown in Figure 5. Repeating our extreme analysis with an AO generally reproduces the patterns found for the NAO in North America. The extreme measures used by *Higgins et al.* [2002] of daily average temperature (rather than Tmax and Tmin) and the number of days exceeding a threshold or the separation of the effects of positive and negative phases of the AO could account for the differences in detail between their results and those presented here.

[32] The presence of an NAO effect might cast doubt on the earlier trend estimates for ANN (Figure 3). The comparison between the location trends of L-TREND-NAO and L-TREND (not shown) indicates that the effect of omitting the NAO dependence in L-TREND is very small for ANN (less than 0.04°C/decade) unlike that found earlier for DJF. Thus we conclude that the omission of NAO in the MPP parameter estimation for ANN has not unduly biased the inferred widespread changes in extreme daily temperatures since 1950 presented earlier.

6. Summary

[33] Changes in extreme daily temperatures from a new observed gridded daily maximum and minimum temperature data set are examined using both a stationary marked point process extreme value distribution and one which can have time and/or NAO dependence. The use of nonstationary extreme value distributions is found to be an effective method for discerning changes in extreme temperatures. The location parameter of the extreme distribution (conceptually analogous to the mean of a normal distribution) is the only parameter which is found to have discernable nonstationary characteristics from the observations. Changes in the location parameter represent a uniform shift for the entire extreme tail of the distribution it represents.

[34] Since 1950 there has been a significant positive trend in extreme daily temperature anomalies for both the upper and lower tails of the daily maximum and daily minimum temperature distributions. This is characterized by a trend in the location parameter of a marked point process extreme value distribution. This has occurred for most land areas for which there is suitable data, with extreme daily temperatures in the upper or lower 1.5% of the daily distribution increasing by 1 to 3°C over the observed period. The greatest warming is found for the cold tail of daily minimum temperatures with some regions in Russia and Canada warming by over 4°C. The cold tail of the daily maximum temperatures shows the least warming. Small areas of cooling are found, the largest being for eastern United States and Mexico for xTmax (approximately 1°C) and Australia for nTmin (approximately 0.5°C).

[35] The magnitude of the positive trends and the area they cover are outside the unforced natural climate variability (as simulated by the HadCM3 climate model) for most regions. Whether these increases are a result of anthropogenic emissions of green house gases is not assessed in this study. However, these results are in good agreement with those of *Christidis et al.* [2005] who detected an anthropogenic greenhouse gas influence in the warming of the annual extremes of xTmin, nTmin and nTmax at a global scale. They did not detect any such influence on xTmax but the similarity of the warming patterns between the four extreme measures suggests that the causes of the patterns are the same. It is possible that the noise characteristics of xTmax may be such that detection of an anthropogenic influence is not yet possible.

[36] Trends in extreme temperatures are not significantly different from trends in the mean for most of the land surface for which there is data. However, there are regions of significant differences, for example the cold (warm) tails of the daily distributions for most of North America appear to be warming faster (slower) than changes in the mean. The opposite is found for the Southern Hemisphere in the smaller regions for which there is data. Eurasia has a more complicated pattern but with large regions indicating that upper and lower tails of Tmin have warmed more than the

mean. This suggests that extremes may already be changing at a different rate than the mean in some areas and which may become clearer and more prominent in the future.

[37] The NAO is found to have had a significant influence on northern winter (DJF) extreme temperatures. A negative unit deviation of a standardized NAO index is found to cool the lower extreme tail of daily winter minimum temperatures by up to $\sim$2°C for western Europe and warm those for Newfoundland by $\sim$1.5°C. A smaller influence is seen for most of Eurasia. A warming trend in DJF is found for most regions even when the influence of NAO is taken into account.

[38] Nonstationary extreme value distributions are found to be a powerful and useful tool in the characterization of extremes in a changing climate. If the rate of climate change increases, as it is expected to in the future, such approaches will be needed if accurate information on the changing risk of extremes is to be provided. The forthcoming UK Climate Impact Programme (UKCIP) climate change scenarios will utilize such methods in their ambition to provide probabilistic estimates of future risk from extremes. It is envisaged that this approach to providing information on how extremes will change in the future will greatly assist adaptation activities and adaptation planning. The gridded extreme distribution parameters can be obtained from the Met Office Hadley Centre at http://www.hadobs.org.

[39] **Acknowledgments.** This work is funded by the UK Department for Environment, Food and Rural Affairs (contract PECD/7/12/37). CF was funded by the UK Natural Environment Research Council's National Centre for Atmospheric Science.

References

Alexander, L. V., et al. (2006), Global observed changes in daily climate extremes of temperature and precipitation, *J. Geophys. Res.*, *111*, D05109, doi:10.1029/2005JD006290.

Brabson, B. B., D. H. Lister, P. D. Jones, and J. P. Palutikof (2005), Soil moisture and predicted spells of extreme temperatures in Britain, *J. Geophys. Res.*, *110*, D05104, doi:10.1029/2004JD005156.

Caesar, J., L. Alexander, and R. Vose (2006), Large-scale changes in observed daily maximum and minimum temperatures: Creation and analysis of a new gridded data set, *J. Geophys. Res.*, *111*, D05101, doi:10.1029/2005JD006280.

Christidis, N., P. A. Stott, S. Brown, G. C. Hegerl, and J. Caesar (2005), Detection of changes in temperature extremes during the second half of the 20th century, *Geophys. Res. Lett.*, *32*, L20716, doi:10.1029/2005GL023885.

Clark, R. T., S. J. Brown, and J. M. Murphy (2006), Modeling Northern Hemisphere summer heat extreme changes and their uncertainties using a physics ensemble of climate sensitivity Experiments, *J. Clim.*, *19*, 4418–4435.

Coles, S. G. (2001), *An Introduction to Statistical Modeling of Extreme Values*, 225 pp., Springer, New York.

Cubasch, U., et al. (2001), Projections of future climate change, in *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, pp. 525–582, Cambridge Univ. Press, New York.

Frich, P., L. V. Alexander, P. Della-Martin, 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, *Clim. Res.*, *19*, 193–212.

Gong, D. Y., and C. H. Ho (2004), Intra-seasonal variability of wintertime temperature over East Asia, *Int. J. Climatol.*, *24*, 131–144.

Gordon, C., C. Cooper, C. A. Senior, H. T. 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, *Clim. Dyn.*, *16*, 147–168.

Higgins, R. W., A. Leetmaa, and V. E. Kousky (2002), Relationships between climate variability and winter temperature extremes in the United States, *J. Clim.*, *15*, 1555–1572.

Hurrell, J. W., and H. van Loon (1997), Decadal variations in climate associated with the North Atlantic Oscillation, *Clim. Change*, *36*, 301–326.

Janowiak, J. E., G. D. Bell, and M. Chelliah (1999), A gridded data base of daily temperature maxima and minima for the conterminous United States: 1948–1993, *NCEP/CPC Atlas 6*, Clim. Predict. Cent., Natl. Cent. for Environ. Predict., Camp Springs, Md.

Karoly, D. J., and Q. G. Wu (2005), Detection of regional surface temperature trends, *J. Clim.*, *18*(21), 4337–4343.

Katz, R. W., M. B. Parlange, and P. Naveau (2002), Statistics of extremes in hydrology, *Adv. Water Resour.*, *25*, 1287–1304.

Katz, R. W., G. S. Brush, and M. B. Parlange (2005), Statistics of extremes: Modeling ecological disturbances, *Ecology*, *86*, 1124–1134.

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. Clim.*, *13*, 3760–3788.

Kharin, V. V., and F. W. Zwiers (2005), Estimating extremes in transient climate change simulations, *J. Clim.*, *18*, 1156–1173.

Kharin, V. V., F. W. Zwiers, and X. B. Zhang (2005), Intercomparison of near-surface temperature and precipitation extremes in AMIP-2 simulations, reanalyses, and observations, *J. Clim.*, *18*, 5201–5223.

Kharin, V. V., F. W. Zwiers, X. B. Zhang, and G. C. Hegerl (2007), Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations, *J. Clim*, *20*(8), 1419–1444.

Kiktev, D., D. M. H. Sexton, L. Alexander, and C. K. Folland (2003), Comparison of modeled and observed trends in indices of daily climate extremes, *J. Clim.*, *16*, 560–3571.

Meehl, G. A., J. M. Arblaster, and C. Tebaldi (2005), Understanding future patterns of increased precipitation intensity in a climate model simulations, *Geophys. Res. Lett.*, *32*, L18719, doi:10.1029/2005GL023680.

Nakienovic, N., and R. Swart (Eds.) (2001), *IPCC Special Report on Emission Scenarios*, 599 pp., Cambridge Univ. Press, New York.

New, M. G., M. Hulme, and P. D. Jones (2000), Representing twentieth century space-time climate variability. part II: Development of 1901–96 monthly grids of terrestrial surface climate, *J. Clim.*, *13*, 2217–2238.

Nogaj, M., P. Yiou, S. Parey, F. Malek, and P. Naveau (2006), Amplitude and frequency of temperature extremes over the North Atlantic region, *Geophys. Res. Lett.*, *33*, L10801, doi:10.1029/2005GL024251.

Robeson, S. M. (2004), Trends in time-varying percentiles of daily minimum and maximum temperature over North America, *Geophys. Res. Lett.*, *31*, L04203, doi:10.1029/2003GL019019.

Schär, C., and G. Jendritzky (2004), The European heatwave of 2003: Was it merely a rare meteorological event or a first glimpse of climate change to come? Probably both, is the answer, and the anthropogenic contribution can be quantified, *Nature*, *432*, 559–560, doi:10.1038/432559a.

Tebaldi, C., K. Hayhoe, J. M. Arblaster, and G. A. Meehl (2006), Going to the extremes: An intercomparison model-simulated historical and future changes in extreme events, *Clim. Change*, *79*(3–4), 185–211.

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

Wettstein, J. J., and L. O. Mearns (2002), The influence of the North Atlantic–Arctic Oscillation on mean, variance, and extremes of temperature in the northeastern United States and Canada, *J. Clim.*, *15*, 3586–3600.

Zhang, X. B., F. W. Zwiers, and G. L. Li (2004), Monte Carlo experiments on the detection of trends in extreme values, *J. Clim.*, *17*, 1945–1952.

Zwiers, F. W., and V. V. Kharin (1998), Changes in the extremes of the climate simulated by CC GCM2 under CO_2 doubling, *J. Clim.*, *11*, 2200–2222.

S. J. Brown and J. Caesar, Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK. (simon.brown@metoffice.gov.uk)

C. A. T. Ferro, School of Engineering, Computing and Mathematics, Harrison Building, University of Exeter, North Park Road, Exeter EX4 4QF, UK.