

Methods for Exploring Spatial and Temporal Variability of Extreme Events in Climate Data

C. A. S. COELHO

Centro de Previsão de Tempo e Estudos Climáticos, Instituto Nacional de Pesquisas Espaciais, Cachoeira Paulista, São Paulo, Brazil

C. A. T. FERRO AND D. B. STEPHENSON

School of Engineering, Computing and Mathematics, University of Exeter, Exeter, United Kingdom

D. J. STEINSKOG

Nansen Environmental and Remote Sensing Center, Bjerknes Centre of Climate Research, Bergen, Norway

(Manuscript received 17 November 2006, in final form 18 September 2007)

ABSTRACT

This study presents various statistical methods for exploring and summarizing spatial extremal properties in large gridpoint datasets. Extremal properties are inferred from the subset of gridpoint values that exceed sufficiently high, time-varying thresholds. A simple approach is presented for how to choose the thresholds so as to avoid sampling biases from nonstationary differential trends within the annual cycle. The excesses are summarized by estimating parameters of a flexible generalized Pareto model that can account for spatial and temporal variation in the excess distributions. The effect of potentially explanatory factors (e.g., ENSO) on the distribution of extremes can be easily investigated using this model. Smooth spatially pooled estimates are obtained by fitting the model over neighboring grid points while accounting for possible spatial variation across these points. Extreme value theory methods are also presented for how to investigate the temporal clustering and spatial dependency (teleconnections) of extremes. The methods are illustrated using Northern Hemisphere monthly mean gridded temperatures for June–August (JJA) summers from 1870 to 2005.

1. Introduction

Several different descriptive statistical methods have been found to be useful in climate science for exploring and summarizing the bulk statistical properties of variables at many different spatial locations (e.g., gridpoint datasets). For example, spatial dependencies, *teleconnections*, are often investigated using correlation maps obtained by calculating the correlation between the variable at one grid point and the variable at all other grid points. By ignoring spatial location and treating the set of gridpoint variables as a high-dimensional vector, it is possible to use multivariate techniques to summa-

rize gridpoint datasets. For example, principal component (empirical orthogonal function) analysis is widely used as a data-reduction technique to isolate leading spatial patterns (the principal component loading weights) that account for maximal variance (Hannachi et al. 2007).

Such approaches are appropriate for summarizing the bulk statistical properties of the whole distribution such as the variance and covariance structure, but are not ideally suited for exploring extremal properties in the tails of the distribution. Because of the potential impact of events having large values of meteorological variables (e.g., heat waves caused by large surface temperatures), there is a growing need for statistical methods suitable for exploring extremal properties in gridpoint datasets. Such methods can help address such important questions as follows:

- How does the probability distribution of extreme intensity vary spatially?

Corresponding author address: C. A. S. Coelho, Centro de Previsão de Tempo e Estudos Climáticos, Instituto Nacional de Pesquisas Espaciais, Rodovia Presidente Dutra, Km 40, SP-RJ, 12630-000 Cachoeira Paulista, São Paulo, Brazil.
E-mail: caio@cptec.inpe.br

- How does the probability distribution of extreme intensity depend on time-varying factors such as natural modes of variability [e.g., ENSO and the North Atlantic Oscillation (NAO)] and long-term climate change?
- How do extremes cluster in time at different locations?
- How are extremes at one location related to extremes at another location?

This study presents various extreme value theory (EVT) methods that we believe could be useful for addressing such questions when using large gridpoint datasets. Our approach uses values above large time-varying thresholds to infer properties about extremal behavior in the tail of the distribution. Rather than use sample statistics (e.g., extreme indices statistics) or models fitted separately at each grid point, we present a flexible EVT model that can include explanatory variables to account for spatial and temporal variation in the tail of the distribution. Smooth estimates are obtained by fitting the model over neighboring grid points while accounting for possible spatial variation across these points (spatial pooling). A simple approach is presented for how to choose the thresholds so as to avoid biases from seasonal and long-term nonstationarity in the gridpoint time series.

The paper is structured as follows. Section 2 presents a brief summary of the main concepts in EVT and reviews its use in climate science. Section 3 describes the Northern Hemisphere monthly mean temperature dataset used in this study that was motivated by the need to better understand large-scale heat waves such as the European heat wave of 2003. Section 4 discusses various methods for how to select large values and then shows how to choose a threshold that can avoid sampling biases caused by seasonal and long-term nonstationarity in the time series. Section 5 presents a spatial EVT model and demonstrates how it can be used to summarize the extremal properties of the temperature data. Section 6 shows how time-varying explanatory variables can be easily included into the spatial model to account for how extreme temperatures may depend on factors such as ENSO. Section 7 investigates temporal clustering and teleconnections of extremes. Section 8 concludes the paper with a summary of the main findings and suggestions for future research.

2. Extreme value theory and its use in climate science

EVT is the branch of probability theory and statistical science that deals with modeling and inference for extreme values (Coles 2001). EVT methods generally

use a subset of large values from the data sample to *infer* the extreme behavior of the underlying process that generated the data. The inference relies on models for the tails of distributions derived from asymptotic limit theorems and certain regularity assumptions about the distributional tails. Empirical checks and physical arguments should be used to justify the application of such models (e.g., Ferro 2007). Large values are selected in various ways, for example, block maxima (e.g., annual maxima), the r -largest values (e.g., the five largest events in the year), or peaks over threshold (e.g., exceedances above a predefined threshold). The large values are used to infer behavior about the tail of the distribution rather than provide an absolute binary definition of what is an extreme event. Extremeness is a relative concept not an absolute dichotomy.

Only a few published studies have analyzed extreme weather in gridded datasets using EVT. The most common practice based on EVT consists in individually fitting the generalized extreme value (GEV) distribution to samples of annual maximum daily surface temperature, precipitation, and near-surface wind speed at each grid point of reanalysis data, model simulations, and also for station data (e.g., Zwiers and Kharin 1998; Kharin and Zwiers 2000, 2005; van den Brink et al. 2004; Fowler et al. 2005; Kharin et al. 2005). These studies analyzed maps of changes over time in the estimated GEV parameters and also maps of return values (e.g., a 20-yr return value is exceeded once every 20 yr on average). Naveau et al. (2005) fitted the GP distribution to daily temperature and precipitation climate model simulations and examined changes in 30-yr return values over the Euro–Atlantic sector induced by changes in the intensity of the thermohaline circulation.

Other statistical methods are also used to investigate climate and weather extremes in the climate literature. Shabbar and Bonsal (2004) used maximum covariance analysis (also known as singular value decomposition) to investigate the relationship between low-frequency climate variability [e.g., ENSO, the Arctic Oscillation (AO), and the quasi-biennial oscillation (QBO)] and the occurrence of hot and cold extreme daily temperatures in Canada. Ferro et al. (2005) discussed methods for relating temperature and precipitation extremes to the center of the distribution. Beniston et al. (2007) presented diagnostic methods to determine how heat waves, heavy precipitation, droughts, wind storms, and storm surges change between 1961–90 and 2071–2100 in the European Union–Prediction of Regional Scenarios and Uncertainties for Defining Climate Change Risks and Effects (EU-PRUDENCE) project regional model simulations. Other examples can be found in the special

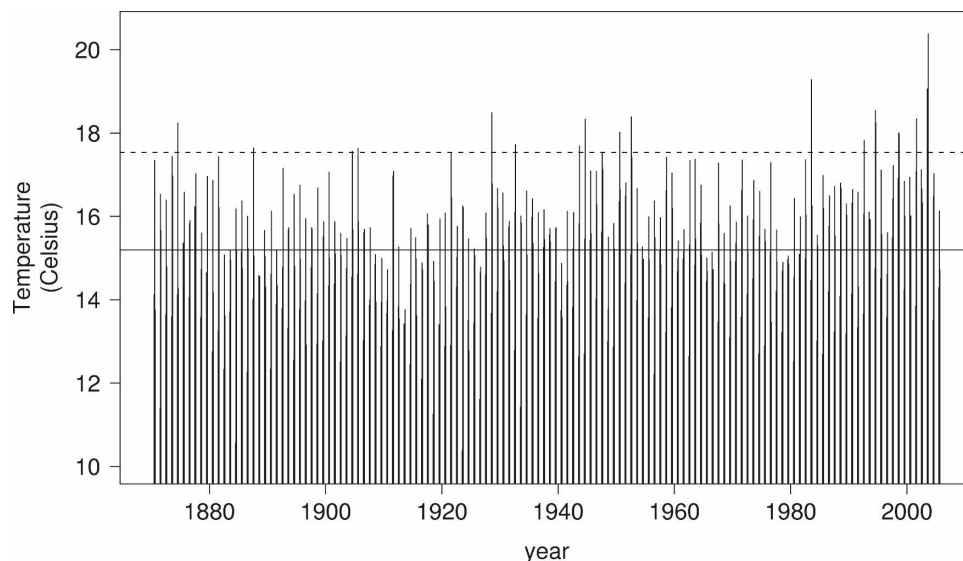


FIG. 1. Summer (June–August) monthly mean temperatures T ($^{\circ}\text{C}$) from 1870 to 2005 at a grid point in central Europe (47.5°N , 12.5°E). Each of the $n_s = 408$ vertical bar is the monthly mean for a particular summer month. The horizontal solid line is the long-term (1870–2005) summer monthly mean temperature of 15.2°C . The horizontal dashed line is the 95th quantile of summer monthly mean temperatures of 17.5°C .

issue of *Global and Planetary Change* (2004, Vol. 44) devoted to extreme climate events.

3. Heat waves in the Northern Hemisphere

Over 20 000 people are believed to have lost their lives during the summer of 2003 because of the persistently hot conditions over Europe (Beniston and Diaz 2004; Beniston 2004; Meehl and Tebaldi 2004; Schär et al. 2004). The event has been attributed to anthropogenic climate change (Stott et al. 2004) and such events are expected to become more frequent and more intense in the future because of climate change (Beniston et al. 2006). There is therefore a pressing need to better understand extreme behavior in surface temperature.

Such heat wave events are the result of persistent blocking conditions and so leave a clear signature in monthly mean temperatures. Therefore, we will illustrate our methods here by applying them to long historical series of observed monthly mean temperatures. This study uses monthly mean gridded surface temperature anomaly data from the Hadley Centre Climatic Research Unit Temperature version 2 (HadCRUT2v; Jones and Moberg 2003; Rayner et al. 2003). (These data are available online at <http://www.cru.uea.ac.uk/cru/data/tem2/>.) The dataset contains combined land and marine monthly mean analysis of surface temperature anomalies from January 1870 to December 2005 in a regular $5^{\circ} \times 5^{\circ}$ global grid. This is one of the best datasets with long time coverage (136 yr) available for climate research. Such a long time series is appropriate

for the investigation of extremes because it contains a large number of episodes, which allows a proper investigation of the tails of temperature distributions. Note, however, that not all grid points have full-time coverage. Mainly Europe, North America, and the North Atlantic region have data covering most of the period during 1870–2005. For the analysis presented in the following sections only grid points with fewer than 50% of values missing (i.e., with at least 68 yr of available data) are used. It is worth noting that gridpoint values have different sampling uncertainties depending upon the number of meteorological data stations available, and sampling uncertainties can vary over time. Anomalies in the original dataset are expressed with respect to the 1961–90 period. To obtain the time series of actual temperatures at each grid point before performing the analysis, the climatological monthly means for the period 1961–90—also provided by the Climatic Research Unit (CRU)—are added to the anomaly temperature series.

Figure 1 shows summer (June–August) monthly mean temperatures (T) at a grid point in central Europe (47.5°N , 12.5°E), which was chosen arbitrarily for illustrative purposes. Each of the $n_s = 408$ vertical bars indicates the monthly mean for a particular summer month. August 2003 and August 1983 stand out as the first and the second hottest observed monthly mean values.

Figure 2 shows the first three sample moments of the distribution of summer temperatures. The median $T_{0.5}$

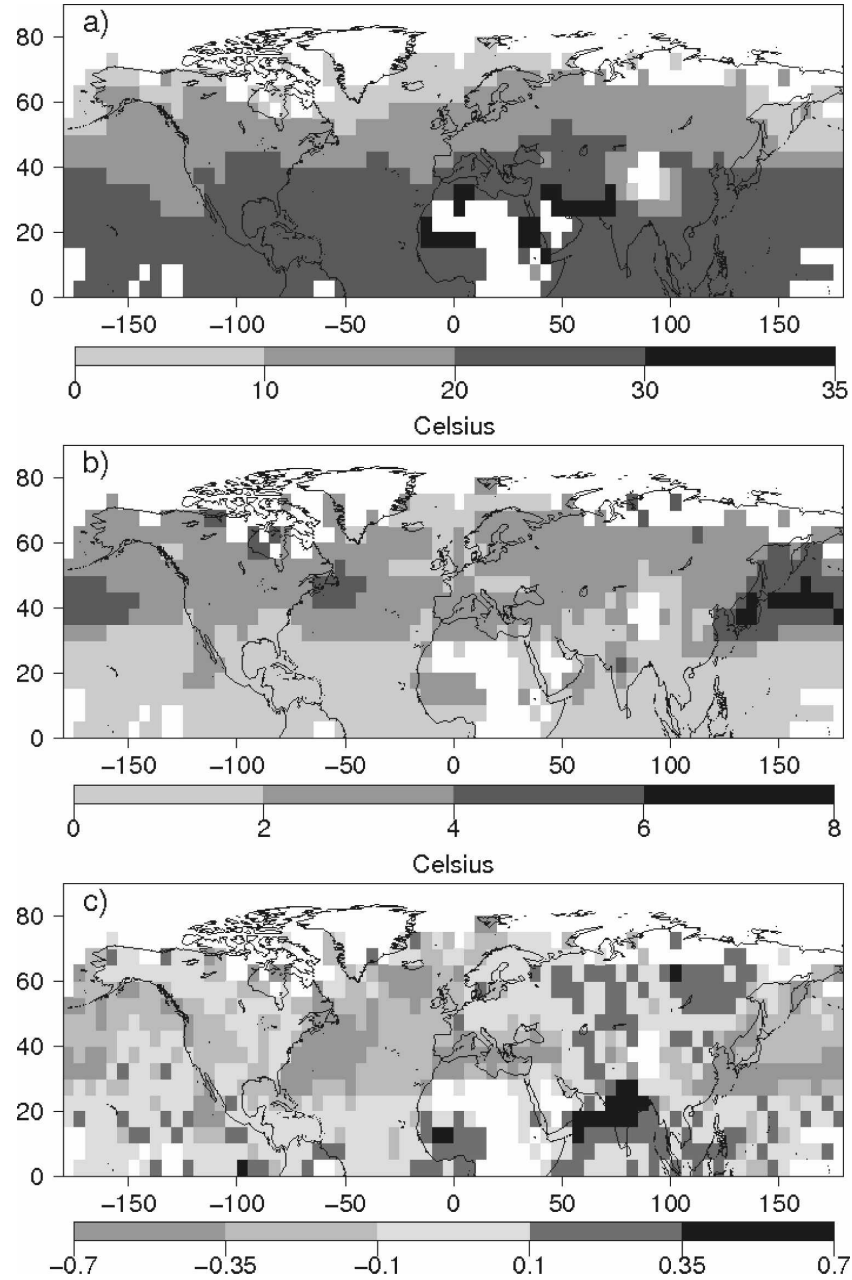


FIG. 2. (a) Median ($T_{0.5}$), (b) interquartile range, and (c) Yule-Kendall skewness statistics (γ), of monthly mean summer temperatures estimated over the 1870–2005 period. White shading signifies missing values.

(Fig. 2a) is a measure of location. The interquartile range $T_{0.75} - T_{0.25}$ (Fig. 2b) is a measure of variability, where $T_{0.75}$ and $T_{0.25}$ are the upper and lower quartiles, respectively. The Yule-Kendall skewness γ (Fig. 2c) is a measure of asymmetry of the distribution and is defined as

$$\gamma = \frac{T_{0.25} - 2T_{0.5} + T_{0.75}}{T_{0.75} - T_{0.25}}. \quad (1)$$

Summer temperatures in the tropics are higher (Fig. 2a) and less variable (Fig. 2b) than in the extratropics. The larger temperature variability in extratropical regions compared to tropical regions is due to mixing processes such as baroclinic instability. Regions of maximum temperature variability over the western Atlantic and Pacific Oceans (Fig. 2b) coincide with the genesis regions of the storm tracks.

Figure 2c shows that large parts of both Asia and the

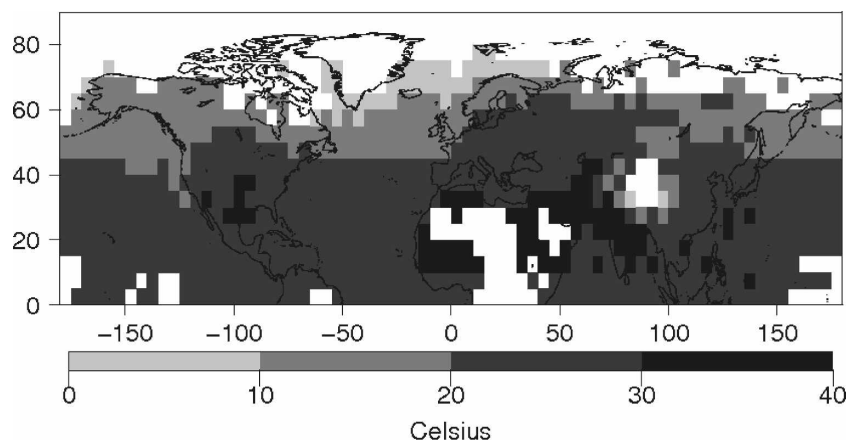


FIG. 3. Maxima of monthly mean summer temperatures over the 1870–2005 period.

tropics have a positively skewed distribution (i.e., distribution with longer tail toward higher temperatures) indicating higher frequency of very high temperatures in these regions when compared to other regions. West of North America, southern Europe, and the Atlantic sector have a negatively skewed distribution (i.e., distribution with longer cold tail) indicating higher frequency of very low temperatures in these regions when compared to other regions. The most common process generally responsible for high temperatures in extratropical regions is atmospheric blocking (Rex 1950). The connection between atmospheric blocking and high temperatures will be further discussed in sections 5b and 5d.

4. How to select the large values: Time-varying thresholds

There are various EVT approaches for selecting a subset of large values from the data sample to *infer* the extreme behavior of the underlying process that generated the data (Coles 2001).

One of the simplest yet least robust ways is to select the largest maximum value from the sample. For example, Fig. 3 shows the maximum values of summer monthly mean temperatures. The spatial pattern broadly resembles the mean temperature pattern (Fig. 2a) with higher temperatures in tropical regions and lower temperatures in extratropical regions. The largest values are observed in arid tropical regions in North Africa, the Middle East, India, and central North America where the lack of precipitation results in surface dryness and consequently high temperatures. The maximum value is based on only one value from the sample and so is not a generally reliable summary statistic. It is nonresistant to outliers (e.g., a single erroneously high temperature measurement).

Alternatively, one could define blocks of data (e.g., a year or a decade) and select the largest (or the r largest) values of each block, and then estimate parameters of the GEV distribution based on the selected sample values (Coles 2001). However, one should be careful about how to choose blocks. For monthly mean temperatures, the block maxima approach using, for example, annual blocks is not appropriate because the blocks are not large enough (only 12 values are available for each year) for the asymptotic properties that give rise to the GEV distribution to provide a close approximation. For monthly mean temperatures a larger block (e.g., a few decades) would be required, and would therefore substantially reduce the sample size for the estimation of GEV distribution parameters. The annual block maxima approach, however, could be appropriate for daily temperatures that have a larger block size of 365 values per year (Zwiers and Kharin 1998; Kharin and Zwiers 2000, 2005).

Rather than use only the largest values in blocks, the peaks-over-threshold approach considers all sample values that exceed a suitably high predefined threshold u (see Coles 2001, chapter 4). For example, all the exceedances $T > u_{0.95}$ above the 95th quantile $u_{0.95} = 17.5^\circ\text{C}$ (dashed line) of summer temperatures are shown in Fig. 1. This quantile illustrates extremes that recur on average once every 20 yr. The probability distribution of the *excesses* $T - u$ can then be modeled using smooth tail distributions such as the generalized Pareto (GP) distribution (see section 5). The theoretical tail distribution allows one to smoothly interpolate and extrapolate return probabilities for events having amplitudes above the threshold. The choice of threshold is often a compromise between having large enough values to represent events in the tail of the distribution, yet also having a sufficient number of exceedance

events to obtain reliable fits. Therefore, exceedances above the threshold do not always have to be far out in the tail of the distribution and so should not be uncritically referred to as *simple extreme events* or *extreme indices* as is often the case in the climate literature (e.g., Houghton et al. 2001). In other words, moderately rare events can sometimes be useful to make inference about extreme behavior. Exceedance above a threshold does not define an *extreme event*—it defines a value that can be used to infer the properties of the whole continuum of events having large values.

For nonstationary processes, it is necessary to consider time-varying thresholds in order to avoid sampling biases. For example, the use of a fixed (constant) threshold for monthly temperatures with a warming trend will oversample warmer periods (e.g., summer and the latter part of the record) compared to other periods in the record. Furthermore, warm-period exceedances above a fixed threshold may no longer be even in the tail of the distribution, which will lead to systematic biases when estimating tail distributions. It is therefore necessary to consider using time-varying thresholds that are able to account for seasonal and long-term variations when investigating extreme behavior in climatic datasets. The possibility of using time-varying thresholds makes the peak-over-threshold approach more readily able to accommodate nonstationarities than the block maxima approach. We recognize that although the block maxima approach cannot represent nonstationarity within blocks, it is perfectly straightforward to represent nonstationarity between blocks. It is important to note, however, that for some practical applications (e.g., human health or ecosystems) it may be important to consider exceedances above a fixed threshold. For these applications, trends in the occurrence of extreme events so defined are an important part of the signal.

Note that the use of a time-varying threshold should not be confused with whether or not the distribution of the excesses varies in time. To account for time variation in the excess distribution one also needs to allow the GP distribution parameters to depend on covariates such as calendar month, year, etc. (Davison and Smith 1990). However, it is quite possible to have excesses above a time-varying threshold that have constant-in-time probability distribution. The threshold may also be allowed to vary with covariates other than time, which may yield a more accurate and meaningful description of the data in specific applications. Such a generalization was uninformative for our data.

This study will investigate the distribution of excesses above a time-varying threshold designed to ensure the following: 1) an approximately constant exceedance

frequency, 2) that analysis is not biased toward the warmer climate of the end of the twentieth century, and 3) that excesses are yielded relative to contemporary climate, and are therefore designed to reflect effects of similar physical processes at all times. The time-varying threshold $u_{y,m} = L_{y,m} + \varepsilon$ in year y and calendar month m is considered to be the sum of a long-term trend $L_{y,m}$ and a positive constant ε , that is adjusted so as to have $\alpha\%$ of the observed values above $u_{y,m}$. The long-term trend $L_{y,m}$ models the effect of global warming and decadal variability and is distinct for each month m to account for differential trends within the annual cycle. Here $L_{y,m}$ is estimated with a local polynomial fit with sliding window of 20 yr. The constant ε is obtained empirically by lifting $L_{y,m}$ up until $\alpha\%$ of the observed values are above $u_{y,m}$. Appendix A describes the procedure for obtaining ε .

The time-varying threshold is illustrated in Fig. 4, which shows the observed monthly mean temperatures $T_{y,m}$ (black dots) for June (Fig. 4a), July (Fig. 4b), and August (Fig. 4c) for the grid point in central Europe (47.5°N, 12.5°E) from 1870 to 2005. The solid lines in Figs. 4a–c are the long-term trends $L_{y,m}$. The dashed lines are the thresholds $u_{y,m} = L_{y,m} + \varepsilon$ obtained by adding the constant ε to $L_{y,m}$ (solid lines) so as to have $\alpha = 5\%$ of the observed values above $u_{y,m}$. The threshold $u_{y,m}$ with $\alpha = 5\%$ is hereafter referred to as the 95% threshold. A constant fairly stable frequency of exceedances (i.e., events $T_{y,m} > u_{y,m}$ when black dots are above the dashed lines) is noted throughout the whole 1870–2005 period in Figs. 4a–c. The largest values are observed during June (Fig. 4a) and August (Fig. 4c) 2003. The summer of 2003 also stands out in Fig. 4d, which shows summer excesses $T_{y,m} - u_{y,m}$ above the 95% threshold $u_{y,m}$ (vertical bars above the horizontal line) for the same grid point. The 95% threshold has been used for all results to be presented in sections 5 and 6. Section 5c provides additional justification on why the use of the 95% threshold is appropriate, representing an acceptable balance between large enough values and a sufficient number of exceedances to obtain reliable GP distribution fits.

When a constant threshold u is chosen for each grid point it is possible to plot a single map of the threshold. However, when a time-varying threshold such as $u_{y,m}$ is used, multiple maps of the threshold need to be examined. To illustrate the typical threshold value at each grid point, Fig. 5 shows the long-term time mean of the 95% threshold $u_{y,m}$ for the summer months over the period from 1870 to 2005. It has a broadly similar pattern to Figs. 2a and 3 with larger values observed in arid regions of North Africa and part of the Middle East.

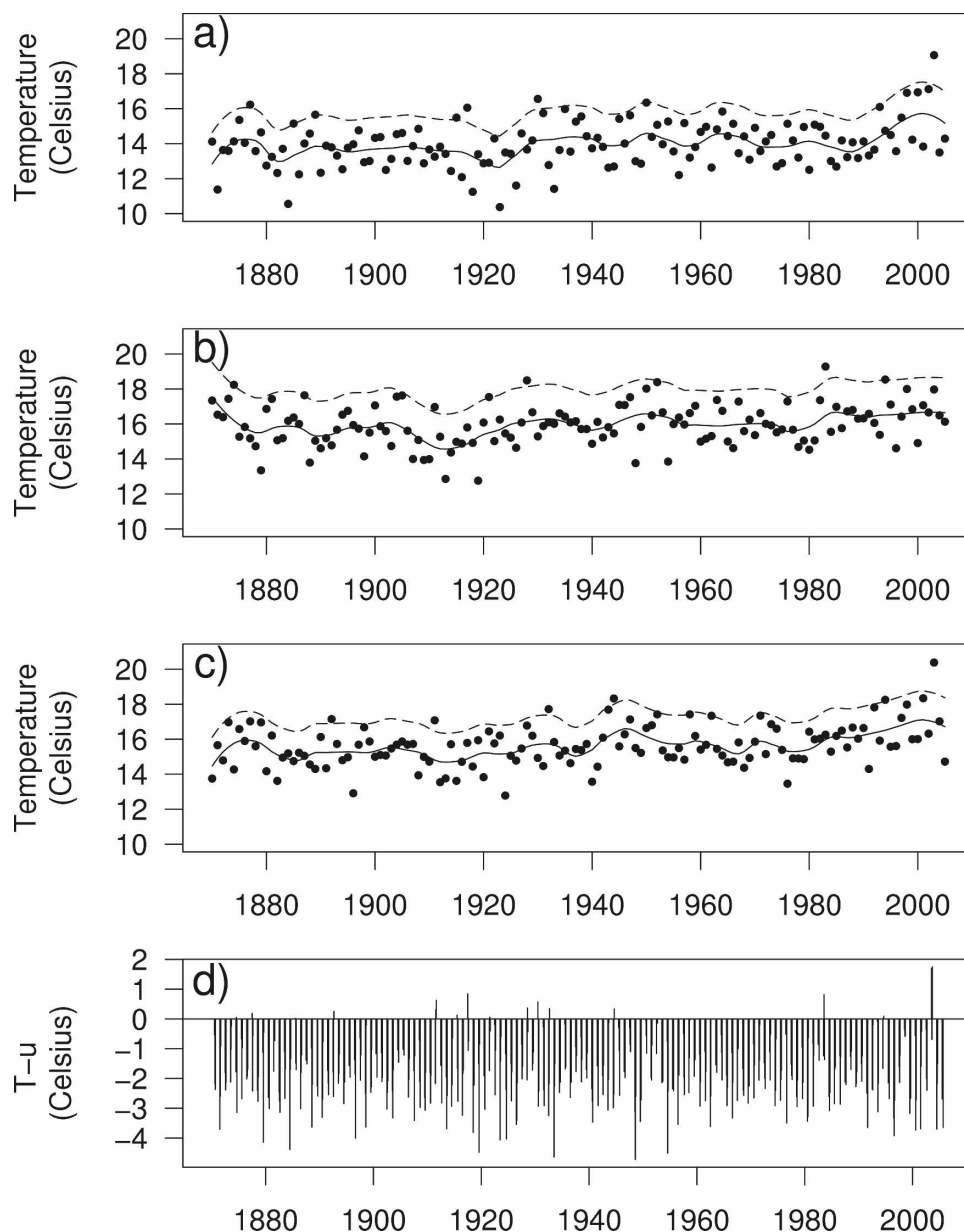


FIG. 4. Observed 1870–2005 monthly mean temperatures $T_{y,m}$ (black dots) for the grid point in central Europe (47.5°N , 12.5°E) during (a) June, (b) July, and (c) August. The solid curves are the long-term trends ($L_{y,m}$) that represent decadal variability. The dashed lines are the time-varying threshold $u_{y,m}$ given by $u_{y,m} = L_{y,m} + \varepsilon$, estimated for each summer month individually, where ε is the increment necessary to have $\alpha = 5\%$ of the observed values above $u_{y,m}$. (d) Summer month differences between $T_{y,m}$ and $u_{y,m}$ (95% time-varying threshold) for the grid point in central Europe (47.5°N , 12.5°E). Vertical bars above the horizontal solid line (zero line) are excesses $T_{y,m} - u_{y,m}$ above the threshold $u_{y,m}$.

Excesses $T_{y,m} - u_{y,m}$ for the exceedance events where $T_{y,m} > u_{y,m}$ can be used to provide distribution summaries of events having large values. Figure 6 shows the sample time median (Fig. 6a) and interquartile range (Fig. 6b, i.e., measure of variability) of the excesses. Large median and interquartile range of excesses are observed in colder and more variable extratropical

regions than in tropical regions, except the tropical Pacific, which also has large values due to the manifestation of ENSO. There are much higher values of median and interquartile range of excesses over extratropical land areas than over most oceanic and tropical regions. This contrast indicates that monthly temperature excesses are on average larger and have

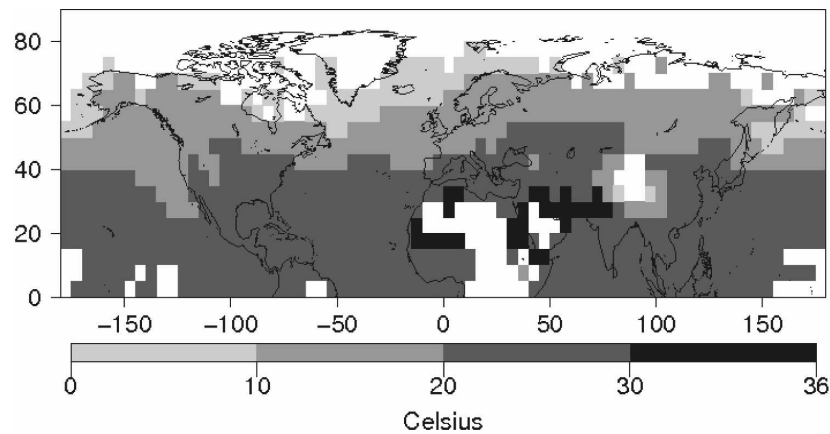


FIG. 5. Time mean of the 95% threshold $u_{y,m}$ for the summer months over the period 1870–2005.

larger variability over extratropical continental areas than over most oceanic and tropical regions. One possible reason for the larger excesses over extratropical land regions could be the much smaller heat capacity of land compared to the oceans (Peixoto and Oort 1993).

5. Spatial GP distribution model for high exceedances

a. Peaks-over-threshold spatial pooling

This section illustrates how peaks over threshold can be used to investigate the upper tail of temperature

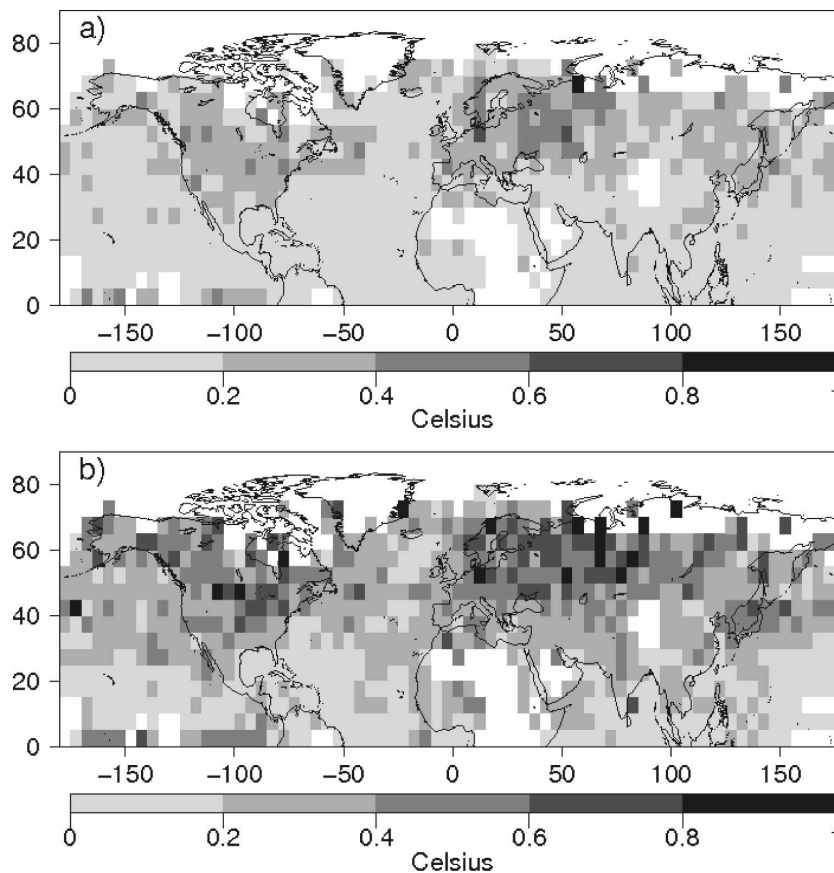


FIG. 6. (a) Median and (b) interquartile range of excesses above the 95% time-varying threshold $u_{y,m}$.

distributions. In the asymptotic limit for sufficiently large thresholds, the distribution of excesses $Z = T_{y,m} - u_{y,m}$ conditional on $T_{y,m} > u_{y,m}$ can be shown to approximate the GP distribution function

$$\Pr(Z \leq z) = H(z) = 1 - \left(1 + \frac{\xi z}{\sigma}\right)^{-(1/\xi)}, \quad (2)$$

which is defined for $z > 0$ and $1 + \xi z/\sigma > 0$, where $\sigma > 0$ is the scale parameter and ξ is the shape parameter of the distribution. The mean, median, variance, and interquartile range of Z are, respectively,

$$E(Z) = \frac{\sigma}{1 - \xi}, \quad (3)$$

$$\text{Med}(Z) = \frac{\sigma(2^\xi - 1)}{\xi}, \quad (4)$$

$$\text{Var}(Z) = \frac{\sigma^2}{(1 - \xi)^2(1 - 2\xi)}, \quad (5)$$

$$\text{IQR}(Z) = \frac{4^\xi \sigma(1 - 3^{-\xi})}{\xi}. \quad (6)$$

For the examples presented in this study the GP distribution parameters are estimated using maximum-likelihood methods (Coles 2001, section 2.6.3). To preserve the smooth nature of temperature fields a spatial pooling approach is applied for the estimation of these parameters (see spatial model in appendix B). The scale and shape parameters, σ_{os} and ξ_{os} , at a grid point, s , of interest are estimated using excesses from both grid point s and grid points in a neighborhood of s . This is achieved by modeling the scale parameter as a bilinear function of the distance between grid points, and constraining the shape parameter to be constant over the neighborhood.

b. Spatial patterns of GP distribution parameters

This section presents maps of GP distribution parameters estimated with the spatial model of appendix B. Because of the logarithmic link function in the scale parameter Eqs. (B1) and (B2) the quantity $e^{\sigma_{os}}$ needs to be examined. Figure 7a shows estimates of $e^{\sigma_{os}}$ for summer monthly mean temperature excesses $T_{y,m} - u_{y,m}$. The scale parameter provides information about the variability (or volatility) of the excesses. Regions with larger values of $e^{\sigma_{os}}$ have higher variability of large temperature values. In accordance with the interquartile range (i.e., variance) of excesses shown in Fig. 6b, higher variability of large temperature values (i.e., large $e^{\sigma_{os}}$) is found over the tropical Pacific and extratropical continental areas when compared to other oceanic and tropical regions. Note also that the scale parameter pat-

tern (Fig. 7a) is similar to the median of excesses (Fig. 6a). This similarity is noted because the median of excesses in (4) is proportional to the scale parameter σ . The maximum variability of large temperature values is observed over extratropical continental Europe and Asia and coincides with the region where atmospheric blocking is typically observed during the summer (Black et al. 2004). We hypothesize that atmospheric blocking is a potential process likely to increase temperature values over Europe and Asia. Shorter persistence (less than 10 days) of anticyclonic (high pressure) conditions in association with warm air advection from North Africa can contribute to increasing the variability of large temperature values over Europe (Nakamura et al. 2005).

Figure 7b shows estimates of the shape parameter ξ_{os} for summer monthly mean temperature excesses $T_{y,m} - u_{y,m}$. The shape parameter tells us about the form (or fatness) of the tail of the distribution of excesses. The tail of the distribution of excesses in regions with a smaller shape parameter is thinner than in regions with a larger shape parameter. Shape parameter values below zero indicate that the distribution has an upper bound (Coles 2001). Shape parameter values above or equal to zero indicate that the distribution is unbounded (i.e., it has an infinite upper tail). Figure 7b shows that most regions have negative shape parameter and hence have an upper-bound excess value equal to $-e^{\sigma_{os}}/\xi_{os}$ from the GP distribution fit. Continental regions (e.g., North America and Asia) have smaller shape parameters than oceanic regions. Figure 7c shows the upper bound of excesses. Regions where no bound is estimated are shaded in black. Larger upper bounds of excesses (between 3° and 6°C) are found in extratropical continental areas (e.g., central North America, most of Europe, and northwest Asia), indicating that excesses over 3°C can be observed in these regions.

To illustrate typical uncertainties on the estimated parameters and in the upper bounds of excesses we report in Table 1 confidence intervals obtained with a bootstrap resampling procedure (Davison and Hinkley 1997) for three grid points, one over North America (47.5°N, 92.5°W), one over central Europe (47.5°N, 12.5°E), and the other over equatorial Africa (12.5°N, 7.5°W). For the first two grid points the sign of the shape parameter is negative and the 90% confidence intervals also lie below zero, supporting the existence of upper bounds for the temperature distributions at these two grid points. For the third grid point the shape parameter is positive suggesting the nonexistence of an upper bound. However, the 90% confidence interval for this grid point ranges from negative to positive values, indicating uncertainty regarding the existence or

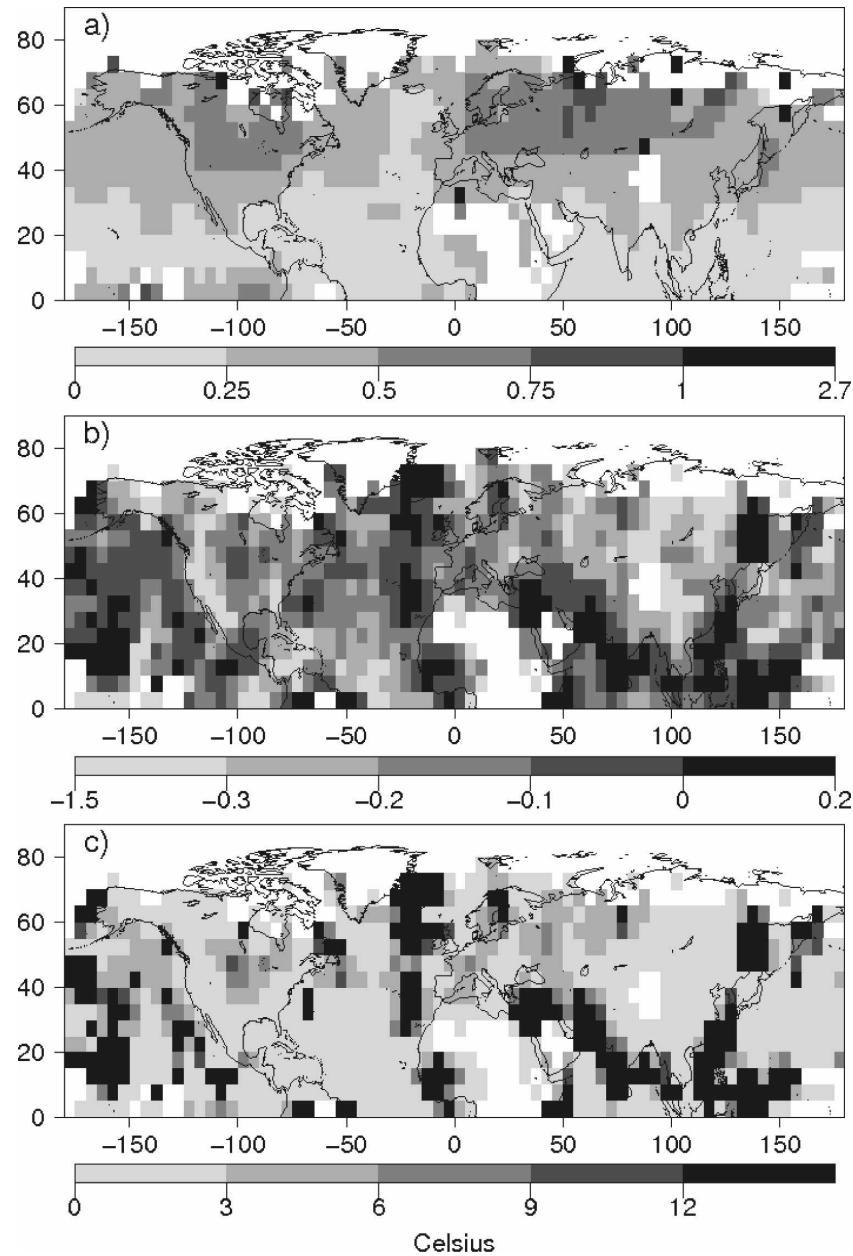


FIG. 7. Summer monthly mean temperature GP (a) scale parameter $e^{\sigma_{os}}$, (b) shape parameter ξ_{os} , and (c) upper bound $-e^{\sigma_{os}}/\xi_{os}$ of excess $T_{y,m} - u_{y,m}$ above the 95% time-varying threshold $u_{y,m}$. Grid points with $\xi_{os} \geq 0$ have an estimated infinite upper bound and are shaded in black.

TABLE 1. Estimates of scale ($e^{\sigma_{os}}$), shape (ξ_{os}), and upper bound ($-e^{\sigma_{os}}/\xi_{os}$) of excesses above the 95% time-varying threshold $-e^{\sigma_{os}}/\xi_{os}$ for three grid points: one over North America (47.5°N, 92.5°W), one over central Europe (47.5°N, 12.5°E), and the other over equatorial Africa (12.5°N, 7.5°W). Values in parentheses are the limits of the 90% confidence intervals estimated using a bootstrap resampling procedure.

	North America	Central Europe	Equatorial Africa
$e^{\sigma_{os}}$	0.57 (0.49, 0.70)	0.55 (0.47, 0.67)	0.24 (0.19, 0.30)
ξ_{os}	-0.14 (-0.31, -0.06)	-0.13 (-0.32, -0.03)	0.06 (-0.17, 0.20)
$-e^{\sigma_{os}}/\xi_{os}$ (°C)	4.2 (2.1, 9.0)	4.2 (2.0, 16.1)	∞ (1.36, ∞)

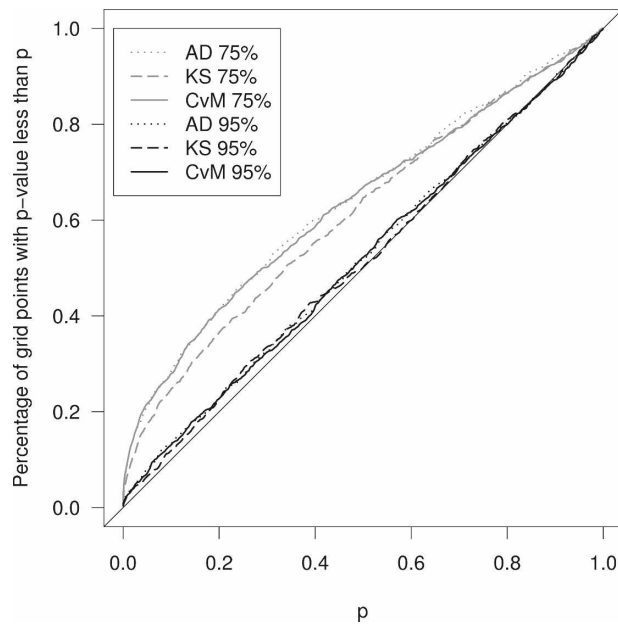


FIG. 8. Percentage of grid points with AD (dotted lines), KS (dashed lines), and CvM (solid lines) test statistics probability value (p value) less than or equal to p (a probability value between 0 and 1) for two choices of time-varying thresholds: 75% (gray curves) and 95% (black curves) of the summer values falling below the threshold.

nonexistence of an upper bound. This uncertainty is also reflected in the 90% confidence interval of the upper bound that ranges from 1.36°C to infinity. Figures presented in Table 1 suggest that the existence of an upper bound is a more conclusive result than the nonexistence of an upper bound. Grid points in Fig. 7c with an upper bound larger than 12°C should therefore be interpreted with caution.

c. Goodness of fit and threshold empirical check

When dealing with parametric distributions such as the GP it is always good practice to examine how well they fit the data. The goodness of fit can be examined using the Anderson–Darling (AD), Kolmogorov–Smirnov (KS), and Cramér–von Mises (CvM) test statistics (Choulakian and Stephens 2001). For this particular application these are all tests of the null hypothesis that the true distribution function of temperature excesses is a GP distribution. It is also advisable to examine if the chosen threshold is high enough so that the assumption of excesses above a sufficiently large threshold approximating a GP distribution is respected. One should critically examine whether the 95% threshold choice is high enough to satisfy this assumption.

Figure 8 shows the percentage of grid points with

AD, KS, and CvM p values less than or equal to p for each p between 0 and 1 for two choices of time-varying thresholds: 95% and 75%. The p -values are computed by bootstrap resampling as in Kharin and Zwiers (2000). If the true distribution function is GP, then the expected percentage of grid points with the p -value less than or equal to p should equal p with all points falling on the diagonal line. All curves for the 95% threshold are close to the diagonal line, indicating that the quality of the fit is good. For thresholds lower than 95% the curves fall on the left-hand side and far from the diagonal line as illustrated in Fig. 8 for the 75% threshold. This indicates that the GP distribution does not fit well to the data at the 75% threshold, and the assumption of a sufficiently high threshold is not valid for thresholds lower than 95%.

d. Excesses and return periods for August 2003

Figure 9a shows excesses during August 2003, the hottest ever recorded monthly mean temperature in central Europe (Figs. 1 and 4c). Excesses of nearly 2°C are observed in central Europe. This event has been linked to the occurrence of atmospheric blocking in central Europe (Beniston and Diaz 2004; Black et al. 2004). The persistence of anticyclonic (high pressure) conditions over Europe during the summer of 2003 resulted in cloudiness reduction, increased surface sensible heat fluxes into the atmosphere, and reduced surface latent heat fluxes (Black et al. 2004; Zaitchik et al. 2006). The lack of precipitation observed in many parts of western and central Europe during this event reduced soil moisture, surface evaporation, and evapotranspiration (Beniston and Diaz 2004). Such a reduction in moisture availability combined with the increase in sensible heat fluxes from the hot land surface contributed to increased temperatures locally.

Figure 9b shows the GP distribution return period estimates $[1-H(z)]^{-1}$ for the August 2003 excesses in Fig. 9a with scale and shape parameter estimates of Figs. 7a and 7b, respectively. The return period is the frequency with which one would expect, on average, a given value (e.g., an excess z of 2°C) to recur. Some grid points over Europe have a return period between 5 and 10 yr, others between 10 and 50 yr and some between 50 and 100 yr. For example, the return period for the grid point in central Europe (47.5°N , 12.5°E) is 61.2 yr with a 90% confidence interval of 39.1, 860.2 yr estimated using a bootstrap resampling procedure (Davison and Hinkley 1997). The immediate left neighbor grid point (47.5°N , 7.5°E) has a return period of 90.6 yr with a 90% confidence interval of 52.0, 3429.9 yr. The immediate right neighbor grid point (47.5°N , 17.5°E) has a return period of 6.5 yr with a 90% confidence interval

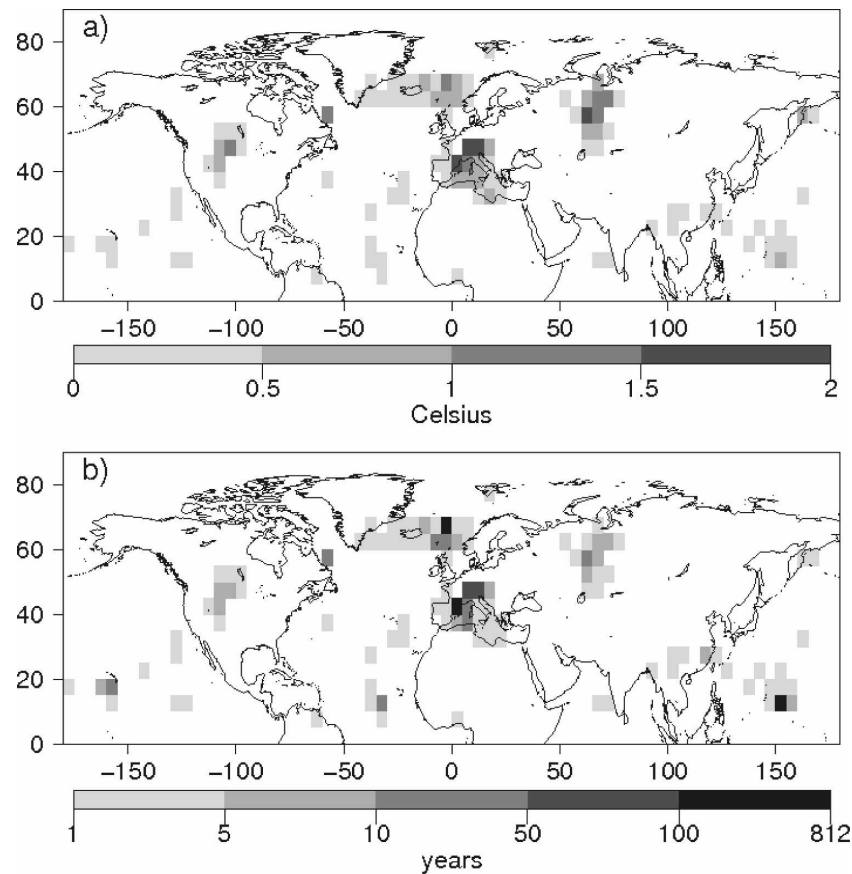


FIG. 9. (a) Excesses $T_{y,m} - u_{y,m}$ above the 95% threshold $u_{y,m}$ during August 2003 (i.e., $T_{y,m} - u_{y,m}$ when $T_{y,m} > u_{y,m}$). (b) Return period estimates for the August 2003 excesses using the GP distribution with the scale and shape parameters estimates of Figs. 7a and 7b, respectively.

of 5.2, 8.7 yr. These return periods are based on the probability of observing a large excess given that the 95% threshold is exceeded. If we multiply our return period estimates by 20 (i.e., 1 over the probability of exceeding the 95% threshold) we can then compare our return period estimates with the return period of 46 000 yr over Switzerland obtained by Schär et al. (2004) using a normal (Gaussian) distribution fitted to the mean June–August temperature during 1990–2002. Our return period estimates are much smaller than the value obtained by Schär et al. (2004).

6. Dependence of extremes on time-varying factors

The dependence of extreme temperatures on factors such as time and ENSO can be easily examined by modeling the shape and scale parameters of the GP distribution as functions of these factors. One can think of this as a regression model for extremes in which

the distribution of the response variable (e.g., the exceedance above a high threshold) is a parametric function of the explanatory variables. The idea of making extreme value distribution parameters dependent upon explanatory variables has previously been used by Coles (2001), Nogaj et al. (2006), Brown et al. (2008), among others. Here we further develop this idea using a model that attempts to preserve the spatial smoothness of the temperature fields (see appendix B for additional information about this spatial model).

For example, if one is interested in how the variability of summer temperature excesses is related to ENSO, the following model could be used for scale and shape:

$$\log \sigma = \sigma_0 + \sigma_1 x, \quad (7)$$

$$\xi = \xi_0, \quad (8)$$

where x is an ENSO index such as the Southern Oscillation index (SOI). Note that in (7) the logarithm of σ

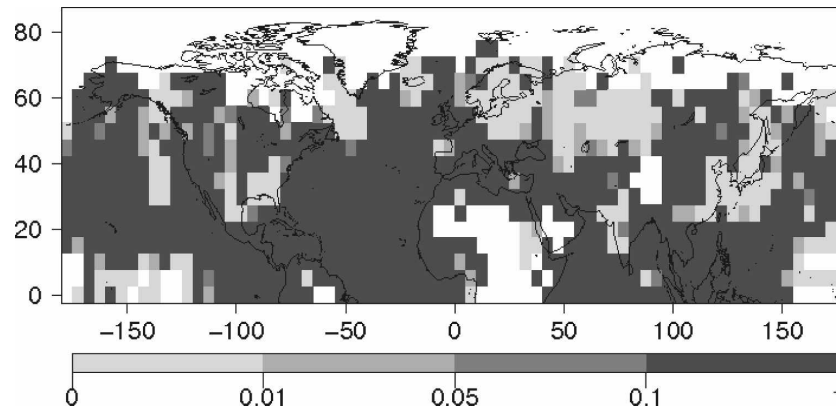


FIG. 10. The p values for the likelihood ratio test for model $\log \sigma_r = \sigma_{0r} + \sigma_{1r} x$ and $\xi_r = \xi_{os}$ against a simpler nested model given by $\log \sigma_r = \sigma_{0r}$ and $\xi_r = \xi_{os}$. The need for the extra parameter σ_{1r} is tested by the null hypothesis $H_0: \sigma_{1r} = 0$ against the alternative hypothesis $H_1: \sigma_{1r} \neq 0$.

is used instead of σ to ensure that $\sigma = e^{(\sigma_0 + \sigma_1 x)}$ is positive for all choices of parameter values σ_0 and σ_1 . Appendix B describes how a spatial pooling approach that uses data from neighboring grid points r in addition to data of the grid point of interest s to estimate the GP distribution parameters σ_0 , σ_1 , and ξ_0 is applied using maximum-likelihood methods (Coles 2001). This approach models σ_0 and σ_1 as bilinear functions of the distance between the neighboring grid points r and the grid point of interest s .

The appropriateness of the nonstationary spatial model given by (B1) can be tested by performing a likelihood ratio test (Coles 2001, section 2.6.6). The model of (B1) is tested against a simpler nested model given by $\log \sigma_r = \sigma_{0r}$ and $\xi_r = \xi_{0r}$. The need for the extra parameter σ_{1r} is tested by the null hypothesis $H_0: \sigma_{1r} = 0$ against the alternative hypothesis $H_1: \sigma_{1r} \neq 0$. For example, the model with a SOI-dependent scale parameter is deemed to be a significantly better description of the data than the same model with a SOI-invariant scale parameter if the deviance between the models exceeds the upper 5% quantile of the chi-squared distribution with one degree of freedom (Coles 2001, section 2.6.6).

Figure 10 shows the map of p -values for the hypothesis test above, where x is the SOI obtained from the Climate Prediction Center (more information available online at <http://www.cpc.noaa.gov/data/indices/>). The null hypothesis cannot be rejected at the 5% significance level over regions where p -values are greater than 0.05. This indicates that over oceanic regions such as the Atlantic, Indian Ocean, and subtropical Pacific the simple model with constant shape and scale parameters is enough to fit the excesses above the 95% time-varying threshold. There is therefore no advantage to

using the SOI factor to model large temperature values over these regions. On the other hand, Fig. 10 shows that over the tropical Pacific and extratropical continental regions in southeast North America, east Europe, Scandinavia, and northwest Asia p -values are less than 0.05. Over these regions the null hypothesis can be rejected at the 5% significance level in favor of the alternative hypothesis. These results suggest that SOI is a statistically significant factor for modulating high temperature variability over the tropical Pacific and these extratropical continental regions. In other words, the variability of large temperature values over these extratropical continental regions is affected by ENSO atmospheric teleconnections.

Because of the logarithmic link function in (B2) the parameters σ_{os} and σ_{1s} are not on the same scale as the response variable. The parameters can be expressed in terms of change in the response due to a unit change in any of the explanatory variables. For example, a unit change in x scales σ_r by $e^{\sigma_{1r}}$. Figure 11 shows maps of $e^{\sigma_{os}}$ (Fig. 11a), $e^{\sigma_{1s}}$ (Fig. 11b), and ξ_{os} (Fig. 11c) estimated using summer temperature excesses during the period 1882–2005, which is the period when the SOI was available. As one could expect, the map of $e^{\sigma_{os}}$ in Fig. 11a is similar to the map of $e^{\sigma_{os}}$ in Fig. 7a and the map of ξ_{os} in Fig. 11c is similar to the map of ξ_{os} in Fig. 7b, where the shape and scale parameters have been estimated by setting $x = 0$ in (B1) and (B2). Figure 11b shows that over most of North America, some regions in northern Europe, and Scandinavia $e^{\sigma_{1s}}$ is larger than 1, indicating that σ_r increases for larger values of the SOI (i.e., larger variance of excesses during La Niña conditions). Figure 12a illustrates this effect using a grid point over southeast North America (37.5°N, 87.5°W). Figure 12a shows a scatterplot of excesses for this grid

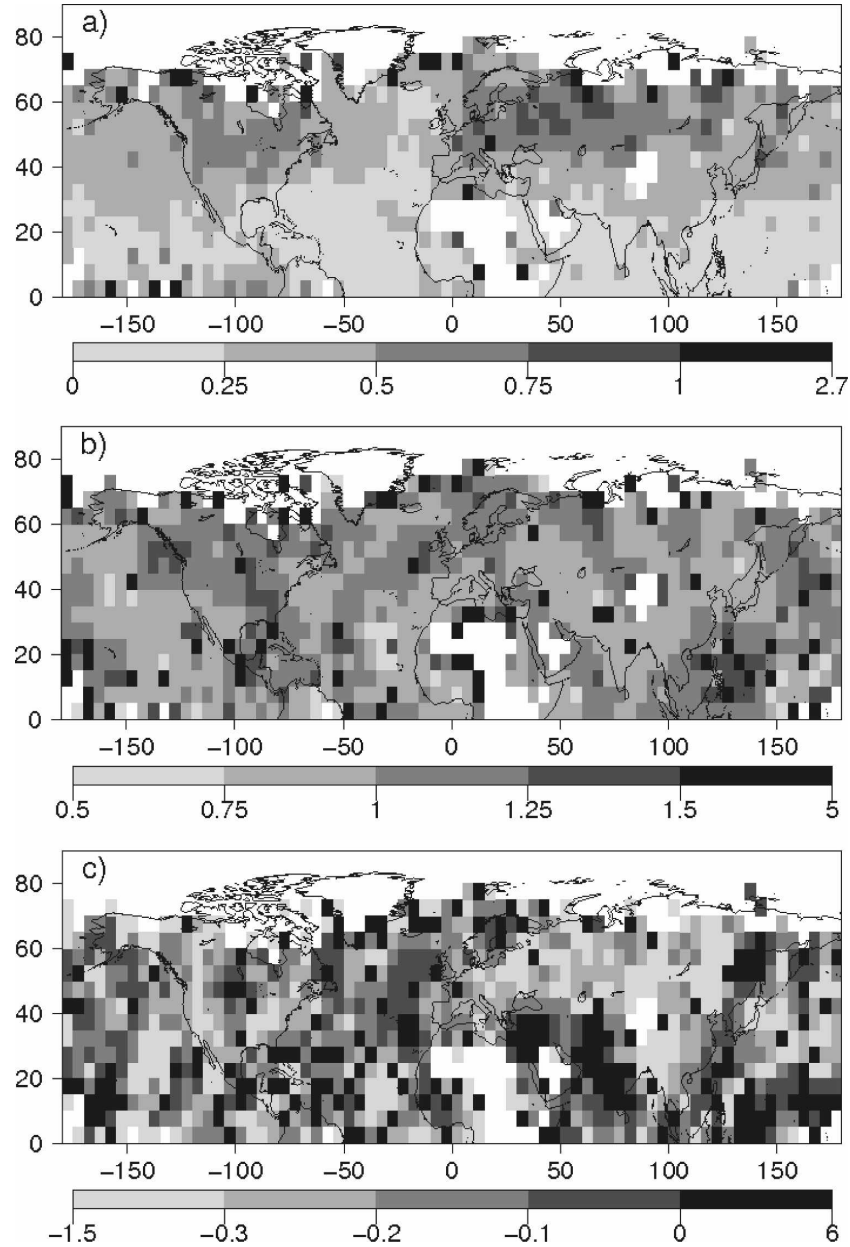


FIG. 11. Here (a) $e^{\sigma_{os}}$, (b) $e^{\sigma_{ls}}$, and (c) ξ_{os} are estimated using summer temperature excesses $T_{y,m} - u_{y,m}$ above the 95% time-varying threshold during the period 1882–2005.

point and the SOI with the median (solid line) and the upper and lower quartiles (dashed lines) of the GP distribution with the scale parameter $\sigma_r = e^{(\sigma_{or} + \sigma_{ls}x)}$ and shape parameter $\xi_r = \xi_{os}$ superimposed. The increased variability of excesses can be noted for larger SOI, even though high variability of large temperature values is also noted during neutral conditions (i.e., when SOI is close to zero).

Figure 11b also shows that for several grid points in the central Pacific $e^{\sigma_{ls}}$ is smaller than 1 indicating that

σ_r increases for smaller values of the SOI (i.e., larger variance of excesses during El Niño conditions). Figure 12b illustrates this effect using a grid point over the central Pacific (7.5°N, 152.5°W). Figure 12c shows a similar scatterplot for a grid point in the eastern Pacific (2.5°N, 92.5°W), where the effect of ENSO on local temperatures is most pronounced. The effect of larger variance of excesses during El Niño conditions is not observed, however: the GP distribution fit is nearly horizontal. This result suggests that for the 95% thresh-

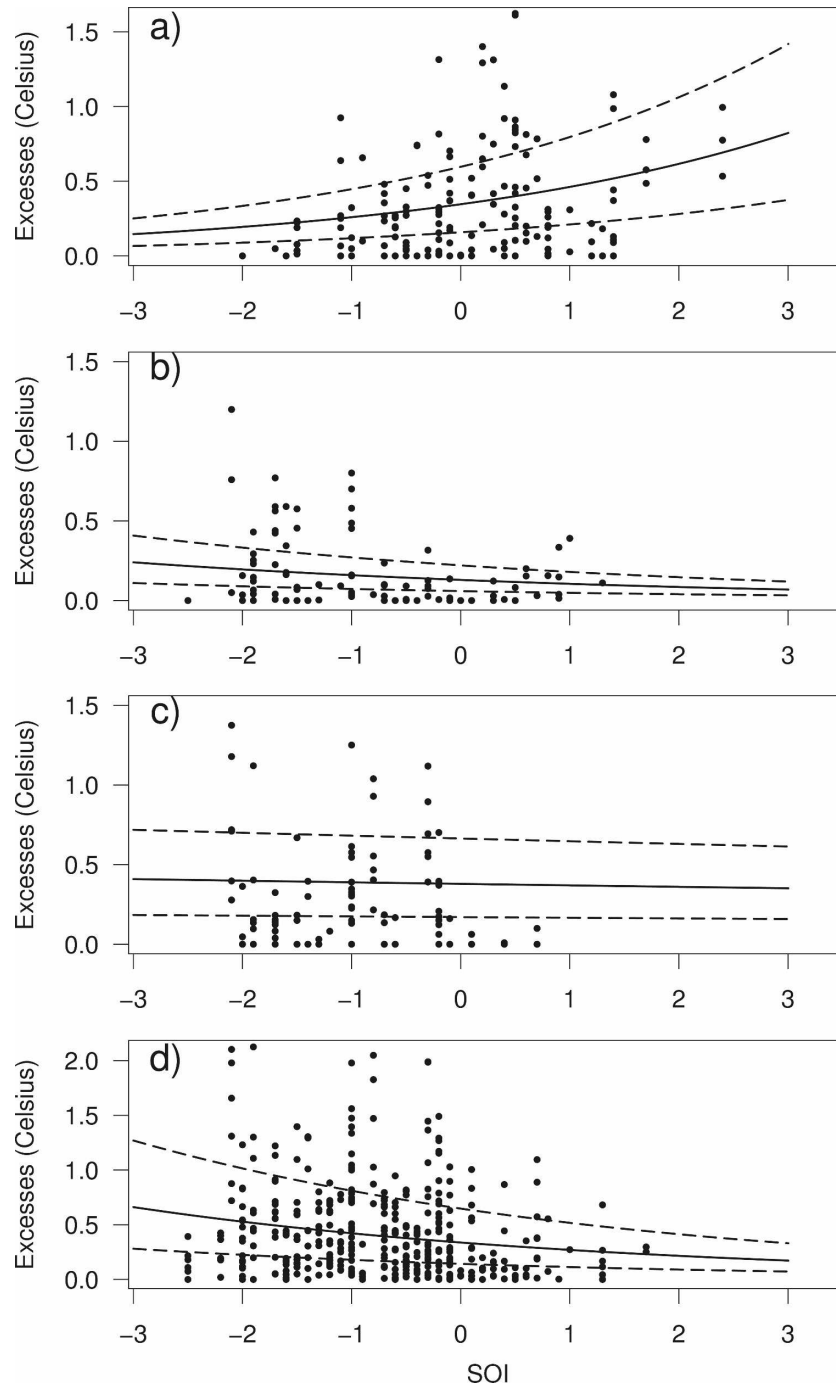


FIG. 12. Scatterplot of excesses $T_{y,m} - u_{y,m}$ and SOI for three grid points: (a) southeast North America (37.5°N , 87.5°W) for excesses above the 95% threshold, (b) central Pacific (7.5°N , 152.5°W) for excesses above the 95% threshold, (c) eastern Pacific (2.5°N , 92.5°W) for excesses above the 95% threshold, and (d) eastern Pacific (2.5°N , 92.5°W) for excesses above the 75% threshold. Median (solid line), upper, and lower quartiles (dashed lines) of the GP distribution with the scale parameter $\sigma_r = e^{(\sigma_{or} + \sigma_{1,r})}$ and the shape parameter $\xi_r = \xi_{os}$.

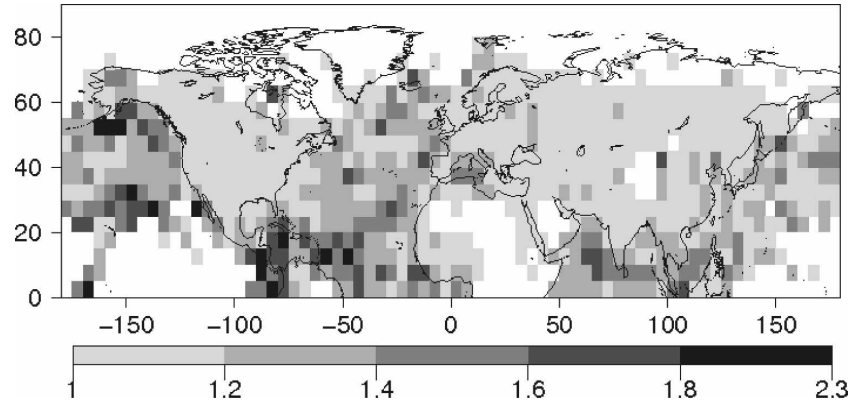


FIG. 13. Average number of summer exceedances \bar{n}_e obtained using the time-varying 95% threshold $u_{y,m}$ over the period 1870–2005.

old it was not possible to statistically detect the differential effect of both phases of ENSO (i.e., El Niño and La Niña) in the eastern Pacific because of the lack of excesses during La Niña conditions. Figure 10 supports this result indicating that in the eastern Pacific SOI is not a statistically significant factor for modulating large temperature variability over this region. Figure 12d shows that if the threshold is reduced to 75% it is then possible to visually detect the differential effect of ENSO on large temperature variability in the eastern Pacific (2.5°N, 92.5°W).

7. Temporal clustering and teleconnections of extremes

a. Temporal clustering analysis

The annual frequency of large values (e.g., the number of large temperature values such as exceedances observed during each summer) is a proxy for clustering of extremes. The average number of summer exceedances $\bar{n}_e = (1/N)\sum_{i=1}^N e$ that occur in years for which there is at least one exceedance provides a measure of the average cluster size. The binary variable $e = 1$ if an exceedance is observed and $e = 0$ if an exceedance is not observed; and N is the total number of summers with at least one observed exceedance. By examining maps of \bar{n}_e it is possible to identify regions where large values are more clustered in time (i.e., regions where there is more serial dependence).

Figure 13 shows \bar{n}_e computed for summer exceedances above the 95% threshold over the period 1870–2005. A clear contrast between continental and oceanic regions is noted. Large temperature values are more clustered over the Atlantic, east Pacific, and Indian Oceans (e.g., averages between 1.4 and 1.8 events per year) than over North America, Europe, and Asia (e.g., average between 1 and 1.2 events per year). This indi-

cates that extreme temperatures are more clustered over the oceans than over land mostly because of the longer memory of the oceans when compared to the continents. Typical values of \bar{n}_e with associated 90% confidence intervals (in parentheses) obtained with a bootstrap resampling procedure are 1.42 (1.00, 1.81) for a grid point over the tropical Atlantic (17.5°N, 47.5°W) and 1.20 (1.00, 1.38) for a grid point over central Europe (47.5°N, 12.5°E).

b. Spatial association of extremes: The χ statistics

Association of large values between different locations (i.e., teleconnection at extreme levels) can be studied using an asymptotic dependence measure χ (Buishand 1984; Coles et al. 1999) as follows. Suppose we are interested in investigating how large monthly mean temperature values at central Europe T_E are related to large monthly mean temperature values at another location T_O . If T_E and T_O have a common distribution function F , it is possible to define

$$\chi = \lim_{u \rightarrow u_+} \Pr[T_O > u | T_E > u], \quad (9)$$

where u_+ is the upper end point of F , so that χ is a limiting measure of the tendency for T_O to be large conditional on T_E being large (Coles et al. 1999). In other words, the probability of temperature at the other location to be high given that the temperature at central Europe is high. If $\chi = 0$ then T_E and T_O are “asymptotically independent.”

However, two different environmental variables could well have uncommon or even unknown distribution functions. Nevertheless, the true distribution of these variables can be estimated using their empirical distributions and one way of obtaining identical distributions is to transform them both to uniform distributions (i.e., ranging from 0 to 1). This can be done by

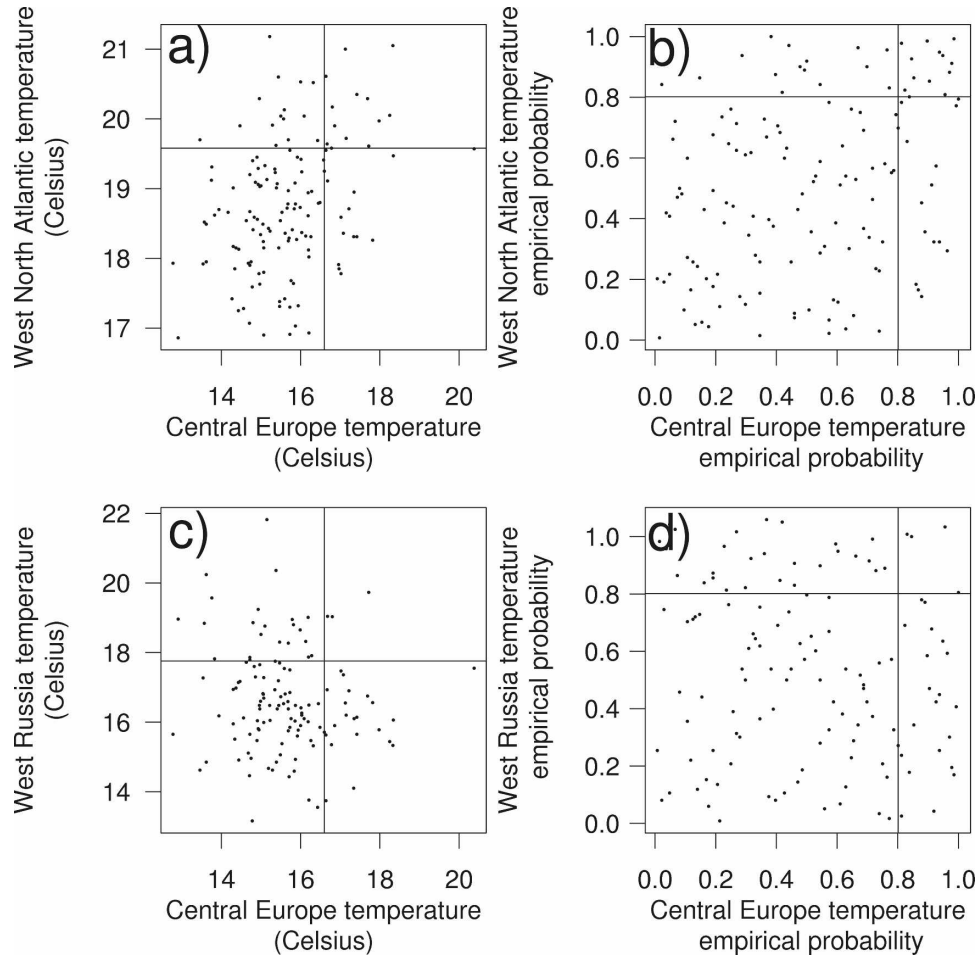


FIG. 14. (a) Scatterplot of August monthly mean temperatures T_E in a grid point in central Europe (47.5°N , 12.5°E) and August monthly mean temperatures T_O in a grid point in the west North Atlantic (42.5°N , 67.5°W). The vertical and horizontal lines are the 80th quantile of the August monthly mean temperatures in each grid point, respectively. (b) Scatterplot of transformed values of T_E and T_O [i.e., $F_{T_E}(T_E)$ and $F_{T_O}(T_O)$]. The vertical and horizontal lines indicate $u = 0.8$. (c), (d) As in (a), (b) but for the grid point in central Europe (47.5°N , 12.5°E) and a grid point in west Russia (57.5°N , 42.5°E).

ranking each set of observations T_E and T_O separately, and dividing each rank by the total number N' of observations in each set.

If F_{T_O} and F_{T_E} are the distribution functions of T_E and T_O , respectively, (9) can be rewritten as

$$\chi = \lim_{u \rightarrow 1} \Pr[F_{T_O}(T_O) > u | F_{T_E}(T_E) > u]. \quad (10)$$

It is possible to show that $\chi = \lim_{u \rightarrow 1} \chi(u)$, where

$$\chi(u) = 2 - \frac{\log \Pr[F_{T_E}(T_E) < u, F_{T_O}(T_O) < u]}{\log \Pr[F_{T_E}(T_E) < u]} \quad (11)$$

defined for thresholds u on the range $0 < u < 1$ (Coles et al. 1999). Therefore, by making the uniform transformations $\text{rank}(T_E)/N'$ and $\text{rank}(T_O)/N'$ to obtain $F_{T_E}(T_E)$ and $F_{T_O}(T_O)$ one can compute $\chi(u)$, where

$\text{rank}(\cdot)$ is the rank of the data. For large thresholds (i.e., $u \rightarrow 1$) the measure $\chi(u)$, which ranges from $2 - \log(2u - 1)/\log(u)$ to 1, provides a simple measure of extremal dependence between T_E and T_O . The sign of $\chi(u)$ determines whether the variables are positively or negatively associated at the quantile level u , with larger values indicating stronger dependence.

Figure 14a shows the scatterplot of August monthly mean temperatures T_E in a grid point in central Europe (47.5°N , 12.5°E) and August monthly mean temperatures T_O in a grid point in the west North Atlantic (42.5°N , 67.5°W). The scatterplot shows that temperatures at the two grid points are positively associated. Indications of extreme dependence are noticeable in that large values often occur simultaneously at the two grid points. The extreme dependence measure $\chi(u)$ is

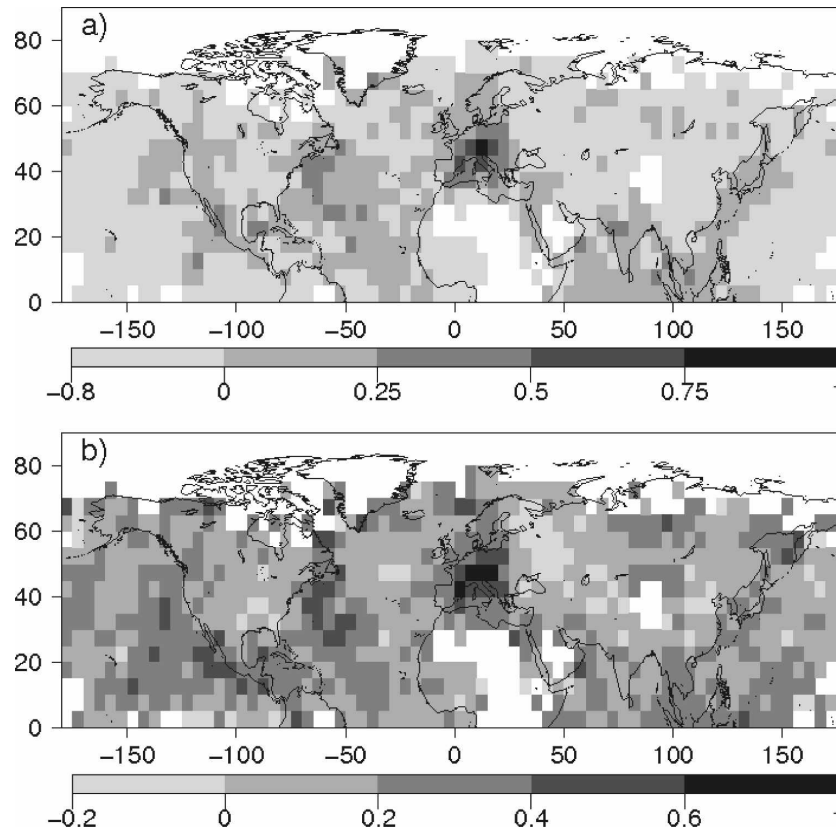


FIG. 15. Here (a) χ and (b) $\bar{\chi}$ for the August monthly mean temperatures for the grid point in central Europe (47.5°N , 12.5°E) with $u = 0.8$.

defined for any value of u between 1 and 0 and here we have used the 80th quantile (i.e., $u = 0.8$). The extreme dependence measure χ is obtained using the points in Fig. 14a that are located on the right-hand side of the 80th quantile threshold (vertical line) of August temperatures in central Europe. The χ statistic is given by the ratio between the number of points on the top-right-hand corner of the scatterplot (i.e., those points that are located above both 80th threshold lines) and the total number of points to the right-hand side of the vertical line. The χ statistics can also be computed as described above but instead using the transformed values of T_E and T_O [i.e., $F_{T_E}(T_E)$ and $F_{T_O}(T_O)$] as shown in Fig. 14b with $u = 0.8$ (vertical and horizontal lines). In practice χ is computed using (11). The value of χ for the data shown in Fig. 14a is 0.29 with normal (Gaussian) 95% confidence interval obtained with the delta method (Coles et al. 1999) of -0.22 , 0.80 . Figures 14c,d show similar scatterplots for the grid point in central Europe (47.5°N , 12.5°E) and a grid point in west Russia (57.5°N , 42.5°E). No sign of extreme dependence is noticeable between these two grid points. The value of χ for the data shown in Fig. 14c is -0.12 with 95% confidence interval of -0.73 , 0.49 .

Figure 15a shows a map of χ with $u = 0.8$ for August monthly mean temperatures for the grid point in central Europe (47.5°N , 12.5°E). As expected, grid points close to the central Europe grid point (47.5°N , 12.5°E) have large values of χ , indicating strong dependence. The west North Atlantic also shows some dependence, as previously illustrated in Figs. 14a,b. The χ statistics ranges from 0.25 and 0.5 in the west North Atlantic regions. Following the interpretation proposed by Coles et al. (1999) that was also used by Svensson and Jones (2002), a value of χ ranging from 0.25 and 0.5 means that if the temperature in central Europe exceeds the 80th quantile, then there is a 25%–50% risk that temperature in the west North Atlantic regions will also exceed the 80th quantile. This dependence is likely to be linked to the manifestation of large-scale planetary Rossby waves with ridges over central Europe and the west North Atlantic and a trough in between. This dependence could also be related to sea surface temperature conditions in the North Atlantic, but further investigation is required to better understand the mechanisms behind such a teleconnection. North America and west Russia show very weak dependence, as also previously illustrated in Figs. 14c,d.

c. Spatial association of extremes: The $\bar{\chi}$ statistics

The χ statistic provides a measure of extremal dependence for asymptotically dependent variables. Asymptotically independent variables (for which $\chi = 0$) can still exhibit dependence at subasymptotic levels, so another statistic is needed to measure the strength of this dependence. Following Coles et al. (1999), such a statistic is given by

$$\bar{\chi} = \lim_{u \rightarrow u_+} \bar{\chi}(u), \quad (12)$$

where

$$\bar{\chi}(u) = \frac{2 \log \Pr[F_{T_E}(T_E) > u]}{\log \Pr[F_{T_E}(T_E) > u, F_{T_O}(T_{T_O}) > u]} - 1 \quad (13)$$

is defined for thresholds in the range $0 < u < 1$. The $\bar{\chi}$ statistic ranges from -1 to 1 . For asymptotically dependent variables $\bar{\chi} = 1$. For completely independent variables $\bar{\chi} = 0$. As χ provides a summary measure of the strength of dependence for asymptotically dependent variables, $\bar{\chi}$ provides a corresponding measure for asymptotically independent variables. In other words, when $\chi = 0$ (or close to 0) then $\bar{\chi}$ is a more appropriate measure of the strength of extremal dependence. As the correlation coefficient is the standard measure of overall association between two variables, $\bar{\chi}$ is a measure of association between large values.

Figure 15b shows a map of $\bar{\chi}$ for August monthly mean temperatures for the grid point in central Europe (47.5°N , 12.5°E). As noticed in Fig. 15a, large temperature values in central Europe are strongly associated with large temperature values in neighboring grid points. Central Europe's large temperature values are also confirmed to be associated with large temperature values in the west North Atlantic (Figs. 14a,b). The value of $\bar{\chi}$ for the data shown in Fig. 14b is 0.39 with normal (Gaussian) 95% confidence interval obtained with the delta method (Coles et al. 1999) of 0.09, 0.69. The one-point correlation map between all August temperature values at the grid point in central Europe and all other grid points in the Northern Hemisphere also show positive association between central Europe and the west North Atlantic (not shown), which is also noticed in the scatterplot of Fig. 14a. Figure 15b still shows a regionalized negative association between temperatures in central Europe and west Russia, which is also noticeable in Fig. 14c, but could not be identified by examining Fig. 15a alone. The values of $\bar{\chi}$ for central Europe and west Russia shown in Fig. 14d is -0.10 with a 95% confidence interval of -0.36 , 0.16 .

8. Conclusions

This study has presented some novel methods for inferring extremal behavior in gridded datasets. A spatial GP distribution model is used to provide a structured and rigorous approach for summarizing the data. By spatial pooling over neighboring grid points, the model is able to provide smoother more robust spatial estimates than fitting different EVT models at each grid point. Care is taken to use a time-varying threshold that avoids sampling biases due to nonstationarity caused by the annual cycle and long-term trends. The model can include explanatory variables to account for temporal changes in the tail distribution such as the effect of natural modes of variability or long-term climate trends. Some approaches are also presented for how to explore temporal clustering of extreme events and teleconnections between extreme events at different spatial locations.

The methods have been illustrated using Northern Hemisphere gridded monthly mean surface temperatures for June–August months from 1870 to 2005. The application of the GP distribution model has revealed that the variance of the temperature excesses at high values is

- larger in extratropical continental regions than in oceanic and tropical regions;
- over tropical and most oceanic regions is mainly driven by local processes rather than by ENSO atmospheric teleconnections;
- over the tropical Pacific, southeast North America, east Europe, Scandinavia, and northwest Asia is affected by ENSO atmospheric teleconnections;
- large during La Niña conditions over southeast North America;
- large during El Niño conditions over the tropical Pacific.

The GP distribution model still helped to quantify the magnitude of the European heat wave observed in August 2003. For example, the return period for excesses above the 95% threshold for August 2003 for a grid point in central Europe was estimated to be of the order of 60 yr. An upper bound for these excesses was identified and estimated to be of the order of 4°C . In other words, our analysis suggests the existence of a finite value that temperatures can reach in central Europe.

Temporal clustering analysis has revealed that temperature excesses are more clustered over the Atlantic and east Pacific Oceans than over continental regions. Teleconnection analysis between extreme events has indicated that central Europe extreme temperatures

during August are related to extreme temperatures in the northwestern Atlantic. This finding suggests that large-scale atmospheric circulation patterns producing extreme temperatures over the North Atlantic sector are likely to simultaneously modulate extreme temperatures over central Europe.

The methods presented here could be further developed and improved. For example, quantile regression (Koenker 2005) could be used to define the threshold for obtaining the excesses. Such an alternative approach would avoid the estimation of the threshold by shifting the mean variability of the observed time series. The methods have been written in the R statistical language (see online at <http://www.r-project.org>) as part of the RCLIM initiative (R software for climate analysis). The software is freely available from the lead author.

Acknowledgments. CASC and DJS were supported as part of work package 4.3 (Understanding Extreme Weather and Climate Events) of the European Union-funded ENSEMBLES project (GOCE-CT-2003-505539; more information available online at <http://www.ensembles-eu.org>). CASC was also partially supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) Processes 2005/05210-7 and 2006/02497-6. CATF was funded by the U.K. Natural Environment Research Council's National Centre for Atmospheric Science. Comments by Adri Buishand, Geert Jan van Oldenborgh, and Deb Panja on an earlier version helped to improve the quality of this manuscript. We also acknowledge the help of Christoph Frei on the development of the RCLIM map plotting functions. Critical comments by three anonymous reviewers helped to substantially improve the quality of this manuscript.

APPENDIX A

Procedure for Obtaining ε

The constant ε that ensures $\alpha\%$ of the observed values above the time-varying threshold $u_{y,m}$ is computed as follows:

- Sort the quantity $T_{y,m} - L_{y,m}$ in ascending order.
- Store in vector (D_1, \dots, D_K) the K positive, nonrepeated, sorted values of $T_{y,m} - L_{y,m}$.
- Add D_1 to $L_{y,m}$ and check the percentage of $T_{y,m}$ values above $L_{y,m} + D_1$.
- Add D_2 to $L_{y,m} + D_1$ and check the percentage of $T_{y,m}$ values above $L_{y,m} + D_1 + D_2$.
- Keep repeating the procedure above until $\alpha\%$ of $T_{y,m}$ values are above $L_{y,m} + D_1 + D_2 + \dots + D_J$,

where $J < K$ is the j th increment to $L_{y,m}$ that ensures that $\alpha\%$ of $T_{y,m}$ values is above $L_{y,m} + D_1 + D_2 + \dots + D_J$.

- The constant ε is then computed as $\varepsilon = \sum_{j=1}^J D_j$.

APPENDIX B

Spatial Model for Generalized Pareto Distribution

The GP temperature parameters could well be estimated using data of each grid point separately. Such an approach produces noisy spatial patterns because it does not take advantage of the smooth spatial structure of the temperature field. To preserve the spatial nature of gridded datasets one can use multivariate techniques in which parameters are estimated by pooling data from several neighbor grid points and spatial dependence is modeled as follows.

Let σ_r and ξ_r be the scale and shape parameters for grid point r , respectively. Suppose we are interested in grid point s and neighborhood $N(s)$, for example, the eight immediately neighboring grid points of a central grid point of interest.

Consider the following model that allows a log-linear relationship between the scale parameter and a covariate x (e.g., SOI), but in which the shape parameter is invariant to x :

$$\left. \begin{aligned} \log \sigma_r &= \sigma_{or} + \sigma_{1r} x \\ \xi_r &= \xi_{or} \end{aligned} \right\} \text{ for all } r \in N(s). \quad (\text{B1})$$

We allow the log-linear parameters σ_{or} and σ_{1r} to vary over $N(s)$ but impose spatial smoothness through a bilinear structure:

$$\sigma_{or} = \sigma_{os} + \theta_{\sigma o}(\lambda_r - \lambda_s) + \phi_{\sigma o}(\delta_r - \delta_s),$$

$$\sigma_{1r} = \sigma_{1s} + \theta_{\sigma 1}(\lambda_r - \lambda_s) + \phi_{\sigma 1}(\delta_r - \delta_s),$$

where λ_r, δ_r are the latitude and longitude of grid point r . We impose greater spatial smoothness on the shape parameter by forcing ξ_{or} to be constant over $N(s)$: $\xi_{or} = \xi_{os}$ for all r in $N(s)$. Similar smoothing methods have been used by Buishand (1991), Zwiers and Kharin (1998), and Kharin and Zwiers (2000).

Our model contains seven parameters ($\sigma_{os}, \sigma_{1s}, \xi_{os}, \theta_{\sigma o}, \theta_{\sigma 1}, \phi_{\sigma o}$, and $\phi_{\sigma 1}$), but we are particularly interested in the first three, which correspond to the grid point of interest. Such a model has the benefit that data from all grid points in $N(s)$ can be used to estimate these parameters of interest. For each grid point s , we estimate the seven parameters by maximum likelihood using data from the nine grid points in $N(s)$, first rewriting the model as a simple log-linear model:

$$\begin{aligned} \log \sigma_r = & \sigma_{os} + \sigma_{1s} x + \theta_{\sigma o} X_r + \phi_{\sigma o} Y_r + \theta_{\sigma 1} x X_r \\ & + \phi_{\sigma 1} x Y_r, \end{aligned} \quad (\text{B2})$$

$$\xi_r = \xi_{os},$$

where $X_r = \lambda_r - \lambda_s$ and $Y_r = \delta_r - \delta_s$. Results presented in section 5 were obtained by setting $x = 0$ in the equations above (i.e., no covariate for the scale parameter), while results presented in section 6 use the full model presented above.

REFERENCES

- Beniston, M., 2004: The 2003 heat wave in Europe: A shape of things to come? An analysis based on Swiss climatological data and model simulations. *Geophys. Res. Lett.*, **31**, L02202, doi:10.1029/2003GL018857.
- , and H. F. Diaz, 2004: The 2003 heat wave as an example of summers in a greenhouse climate? Observations and climate model simulations for Basel, Switzerland. *Global Planet. Change*, **44**, 73–81.
- , and Coauthors, 2007: Future extreme events in European climate: An exploration of regional climate model projections. *Climatic Change*, **81** (Suppl. 1), 71–95.
- Black, E., M. Blackburn, G. Harrison, B. Hoskins, and J. Methven, 2004: Factors contributing to the summer 2003 European heatwave. *Weather*, **59**, 217–223.
- 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.
- Buishand, T. A., 1984: Bivariate extreme-value data and the station-year method. *J. Hydrol.*, **69**, 77–95.
- , 1991: Extreme rainfall estimation by combining data from several sites. *Hydrol. Sci. J.*, **36**, 345–365.
- Choulakian, V., and M. A. Stephens, 2001: Goodness-of-fit tests for the generalized Pareto distribution. *Technometrics*, **43**, 478–484.
- Coles, S. G., 2001: *An Introduction to Statistical Modeling of Extreme Values*. Springer Series in Statistics, Springer, 224 pp.
- , J. Heffernan, and J. A. Tawn, 1999: Dependence measures for extreme value analyses. *Extremes*, **2**, 339–365.
- Davison, A. C., and R. L. Smith, 1990: Models for exceedances over high thresholds. *J. Roy. Stat. Soc. Ser. B. Methodol.*, **52**, 393–442.
- , and D. V. Hinkley, 1997: *Bootstrap Methods and Their Application*. Cambridge University Press, 582 pp.
- Ferro, C. A. T., 2007: A probability model for verifying deterministic forecasts of extreme events. *Wea. Forecasting*, **22**, 1089–1100.
- , A. Hannachi, and D. B. Stephenson, 2005: Simple non-parametric techniques for exploring changing probability distributions of weather. *J. Climate*, **18**, 4344–4353.
- Fowler, H. J., M. Ekström, C. G. Kilsby, and P. D. Jones, 2005: New estimates of future changes in extreme rainfall across the UK using regional climate model integrations. 1. Assessment of control climate. *J. Hydrol.*, **300**, 212–233.
- Hannachi, A., I. T. Jolliffe, and D. B. Stephenson, 2007: Empirical orthogonal functions and related techniques in atmospheric science: A review. *Int. J. Climatol.*, **27**, 1119–1152.
- 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.
- Jones, P. D., and A. Moberg, 2003: Hemispheric and large-scale surface air temperature variations: An extensive revision and an update to 2001. *J. Climate*, **16**, 206–223.
- 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.
- , and —, 2005: Estimating extremes in transient climate change simulations. *J. Climate*, **18**, 1156–1173.
- , —, and X. Zhang, 2005: Intercomparison of near-surface temperature and precipitation extremes in AMIP-2 simulations, reanalyses, and observations. *J. Climate*, **18**, 5201–5223.
- Koenker, R., 2005: *Quantile Regression*. Econometric Society Monograph Series, Cambridge University Press, 366 pp.
- Meehl, G. A., and C. Tebaldi, 2004: More intense, more frequent, and longer lasting heat waves in the 21st century. *Science*, **305** (5686), 994–997.
- Nakamura, N., T. Enomoto, and S. Yamane, 2005: A simulation study of the 2003 heatwave in Europe. *J. Earth Simul.*, **2**, 55–69. [Available online at http://www.es.jamstec.go.jp/publication/journal/jes_vol.2/pdf/JES2_nakamura.pdf.]
- Naveau, P., M. Nogaj, C. Ammann, P. Yiou, D. Cooley, and V. Jomelli, 2005: Statistical methods for the analysis of climate extremes. *C. R. Geosci.*, **337**, 1013–1022.
- 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.
- Peixoto, J. P., and A. H. Oort, 1993: *Physics of Climate*. American Institute of Physics, 564 pp.
- Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, 2003: Globally complete analyses of sea surface temperature, sea ice and night marine air temperature, 1871–2000. *J. Geophys. Res.*, **108**, 4407, doi:10.1029/2002JD002670.
- Rex, D. F., 1950: Blocking action in the middle troposphere and its effect upon regional climate. I: An aerological study of blocking action. *Tellus*, **2**, 196–211.
- Schär, C., P. L. Vidale, D. Luthi, C. Frei, C. Haberli, M. A. Liniger, and C. Appenzeller, 2004: The role of increasing temperature variability in European summer heatwaves. *Nature*, **427**, 332–336.
- Shabbar, A., and B. Bonsal, 2004: Associations between low frequency variability modes and winter temperature extremes in Canada. *Atmos.–Ocean*, **42** (2), 127–140.
- Stott, P. A., D. A. Stone, and M. R. Allen, 2004: Human contribution to the European heatwave of 2003. *Nature*, **432**, 610–614.
- Svensson, C., and D. A. Jones, 2002: Dependence between extreme sea surge, river flow and precipitation in eastern Britain. *Int. J. Climatol.*, **22**, 1149–1168.
- van den Brink, H. W., G. P. Können, and J. D. Opsteegh, 2004: Statistics of extreme synoptic-scale wind speeds in ensemble simulations of current and future climate. *J. Climate*, **17**, 4564–4574.
- Zaitchik, B. F., A. K. Macalady, L. R. Bonneau, and R. B. Smith, 2006: Europe's 2003 heat wave: A satellite view of impacts and land-atmosphere feedbacks. *Int. J. Climatol.*, **26**, 743–769.
- Zwiers, F. W., and V. V. Kharin, 1998: Changes in the extremes of the climate simulated by CCC GCM2 under CO₂ doubling. *J. Climate*, **11**, 2200–2222.