

## **Exploratory tools for the analysis of extreme weather and climate events in gridded datasets**

C. A. S. Coelho<sup>1</sup>

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

C. A. T. Ferro,

*NCAS-Climate, Department of Meteorology, University of Reading, Reading, U.K.*

D. B. Stephenson

*Department of Meteorology, University of Reading, Reading, U.K.*

D. J. Steinskog

*Nansen Environmental and Remote Sensing Center, Bergen, Norway.*

Submitted to Journal of Climate

16 November 2006

---

<sup>1</sup> 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, SP, Brasil.  
E-mail: caio@cptec.inpe.br

## **Abstract**

This paper reviews and introduces new tools based on Extreme Value Theory (EVT) for the analysis of extreme weather and climate events in gridded datasets. The methods allow exploratory analysis of spatial patterns of extremes, the investigation of relationships between extremes and potentially influential factors (e.g. ENSO), the analysis of temporal clustering of extremes, and also the study of teleconnection patterns of extremes (spatial extreme dependence).

The methods are illustrated using Northern Hemisphere monthly mean gridded temperatures for June-July-August (JJA) summers from 1870 to 2005. All analyses are focused on hot extreme temperature events defined by observed temperatures exceeding a pre-defined threshold. Results show that hot extreme temperatures have larger variability in extratropical continental regions than in oceanic and tropical regions. Extreme temperature variability over tropical and oceanic regions is found to be driven mainly by local processes rather than by ENSO atmospheric teleconnections. Over extratropical continental regions extreme temperature variability is found to be affected by ENSO. Larger variability of extreme temperatures is found during La Niña conditions over most of North America, some regions in eastern Europe and Scandinavia. Larger variability of extreme temperatures is found during El Niño conditions over most of Asia, Russia, and western Europe. The Atlantic and East Pacific oceans show higher temporal clustering of extreme events than continental regions, most likely due to the longer memory of oceans compared to continents. Extreme temperatures over central Europe during August are found to be related to extreme temperatures in the west North Atlantic.

## 1. Introduction

Weather and climate time series on large grid-point arrays can be analysed in many ways. For example, atmospheric teleconnections can be investigated by using composite and correlation maps between the time series of a grid-point and the time series of all the other grid-points. Principal component analysis is usually applied to gridded datasets to isolate leading patterns of climate variability, for example the El Niño-Southern Oscillation (ENSO) and the North Atlantic Oscillation (NAO). Standard methods generally summarise all events (the whole distribution) and therefore mask extreme events (tail of the distribution). Given the increasing availability of gridded datasets in recent years, there is therefore a need for developments of statistical tools designed for summarizing weather and climate extreme behaviour in large datasets.

This paper aims to review and introduce new tools based on extreme value theory (EVT) for exploring extreme events in gridded datasets. EVT is the branch of probability theory and statistical science that deals with modelling and inference for extreme values. Limit theorems characterise the statistical behaviour of processes in rare states. For example, maximum values of blocks of data (e.g. years or decades) can be studied using the generalized extreme value (GEV) distribution, and values above high thresholds can be studied using the generalized Pareto (GP) distribution (Coles 2001). EVT relies on asymptotic assumptions (e.g. large blocks or high thresholds) so that results are more accurate for rarer events, and tends to ignore the process by which values evolved into large values.

Only a few published studies have analysed extreme weather in gridded datasets. The most common practise based on EVT consists in fitting the 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 analysed maps of changes over time in the estimated GEV parameters and also maps of return values (e.g. a 20-year return value is exceeded once every 20 years on average). Naveau *et al.* (2005) fitted the GP distribution to daily temperature and precipitation climate model simulations and examined changes in 30-year return values over the Euro-Atlantic sector induced by changes in the intensity of the thermohaline circulation. The special issue of *Global and Planetary Change* edited by Beniston and Stephenson (2004) gathered several articles examining climate and weather extremes in gridded datasets. 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 and the Quasi-Biennial Oscillation) 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 centre of the distribution. Beniston *et al.* (2006) presented diagnostic methods to determine how heat waves, heavy precipitation, droughts, wind storms, and storm surges change between 1961-90 and 2071-2100 in EU-PRUDENCE project regional model simulations.

The European heat wave in summer 2003 attracted much attention from climate scientists since this event reflected the temperatures that are projected to occur in later decades of the 21<sup>st</sup> century (Beniston 2004; Beniston and Diaz 2004; Meehl and Tebaldi 2004; Schär *et al.*, 2004; Stott *et al.* 2004). Over 20,000 people are believed to have lost their lives during the summer 2003 because of the persistently hot conditions over Europe (Beniston and Diaz 2004). Motivated by this episode this

study investigates hot extremes in monthly mean temperatures during the summer months (June-July-August) in the Northern Hemisphere. A good understanding of such high-impact events can improve decision-making and disaster planning that can then help mitigate some of the losses.

The study will focus on *simple extreme* events defined as an individual local weather variable exceeding a critical level on a continuous scale (IPCC 2001). A high critical temperature level is chosen and the extreme events are those events with temperature values above the critical level. The distribution of the values above the critical level at each grid-point is then examined.

The paper will address the following scientific questions:

- How does the risk from extremes of different intensities vary in space and time?
- How do extremes depend on time-varying factors (e.g. ENSO and NAO)?
- How do extremes cluster in time at different locations?
- How are extremes at one location related to extremes at another location, i.e. are there teleconnections at extreme levels?

The methods used to explore extremes in gridded datasets are illustrated throughout the paper with examples of application. Section 2 describes the dataset used in the investigation. Section 3 presents summary statistics for the summer temperature distribution. Section 4 defines basic concepts required for the interpretation of the results. Section 5 illustrates how EVT can be used to model extreme events. Section 6 explores factors controlling extreme events. Section 7 presents temporal clustering analysis of extreme events. Section 8 investigates teleconnections between extreme events. Section 9 summarises results and suggests future developments.

## 2. Global temperature dataset: 1870-2005

This study uses monthly mean gridded surface temperature data (HadCRUT2v) from the Climatic Research Unit (CRU) (Jones and Moberg 2003; Rayner *et al.* 2003) available from <http://www.cru.uea.ac.uk/cru/data/temperature/>. 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 the longest time coverage (136 years) 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 distribution of extreme temperatures. 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 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 years of available data) are used. Anomalies in the original dataset are expressed with respect to the 1961-1990 period. To obtain the time series of actual temperatures at each grid-point before performing the extreme analysis, the climatological monthly means for the period 1961-1990 – also provided by CRU – are added to the anomaly temperature series.

Figure 1 shows summer (June-July-August) monthly mean temperatures ( $T$ ) at a grid-point in central Europe ( $12.5^{\circ}\text{E}$ ,  $47.5^{\circ}\text{N}$ ), which was chosen arbitrarily for illustrative purposes. Each of the  $n_s=408$  vertical bar 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.

### 3. Summary statistics for summer temperature distribution

Before focussing on extreme events, it is worth exploring the whole distribution of monthly mean temperature values at each grid-point. Figure 2 shows the first three sample moments of the distribution of summer temperatures. The mean is a measure of location, the standard deviation is a measure of variability, and the skewness is a measure of asymmetry of the distribution. The moment measure of skewness  $b$  is defined as

$$b = \frac{1}{n_s} \sum_{i=1}^{n_s} \left( \frac{T_i - \bar{T}}{s} \right)^3, \quad (1)$$

where  $T$  is a monthly mean summer temperature value,  $\bar{T}$  is the long-term summer mean temperature and  $s$  is the long-term standard deviation.

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 East Europe, large part of Asia and most of the tropics have a positively skewed distribution (i.e. distribution with longer tail towards higher temperatures) indicating occasional occurrence of very high temperatures in these regions. Most of North America and the Atlantic sector have a negatively skewed distribution (i.e. distribution with longer cold tail) indicating occasional occurrence of very low temperatures. The most common process responsible for high temperatures in extratropical regions is atmospheric blocking (Rex 1950). The connection between atmospheric blocking and extreme hot temperatures will be further discussed in section 5.

## 4. Definition of extreme events

Figure 3 shows the maximum values of summer monthly mean temperatures. The spatial pattern broadly resembles the mean temperature pattern (Figure 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 and may not be a very reliable summary of the distribution of hot extreme events. In addition, the maximum value is non-resistant to outliers (i.e. if a single anomalously high or erroneous temperature is recorded at a particular location where temperatures are usually mild, the use of the maximum value will strongly bias the analysis). A more robust approach for investigating extremes is to produce a subsample of the original temperature data that contains *excesses*  $T - u$  above a pre-defined threshold  $u$  and then use the sample of excesses to estimate parameters of the distribution of excesses. This is the so-called peaks-over-threshold approach (e.g. Coles 2001, chapter 4), which allows the description of the entire distribution of excesses and provides more reliable summary statistics of the distribution of extreme events.

One could question the feasibility of using the EVT block maxima approach instead of the peaks-over-threshold approach. 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) and therefore the asymptotic assumption of large blocks does not hold. For monthly mean temperatures a larger block (e.g. a few decades) would be required, and would therefore reduce substantially 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). Nevertheless, because of the existence of an annual cycle the block maxima approach would limit the investigation of hot temperature extremes to summer values.

The horizontal dashed line in Fig. 1 is the 75<sup>th</sup> quantile ( $u_{0.75} = 16.2^{\circ}\text{C}$ ) of all the summer monthly mean temperatures. Values above such a threshold (i.e. *exceedances*  $T > u_{0.75}$ ) can be used to define simple extreme events (IPCC 2001). The solid thick line in Fig. 1 is the long-term mean trend  $L_{y,m}$  estimated with a local polynomial fit with sliding window of 10 years using all (January to December) monthly mean temperature values from 1870 to 2005, where  $y$  is the year index and  $m$  is the month index. This thick line therefore represents the observed decadal variability and shows an increasing trend. Because of the existence of this time trend and also due to the existence of the annual cycle variations within June-August it is more appropriate to define extreme events using a time-varying threshold that also incorporates the long-term trend. Such a time-varying threshold ensures: a) approximately constant exceedance frequency; b) analysis is not biased towards the warmer climate of the end of the 20<sup>th</sup> century; c) excesses are yielded relative to contemporary climate, and are therefore designed to reflect effects of similar physical processes at all times. This study will investigate the distribution of excesses above a time-varying threshold estimated as described below.

The starting point for such an analysis is to define and estimate the threshold to be used to obtain the excesses. This threshold is defined as a seasonally varying fit to the time series of observed values plus a constant. A procedure for defining and estimating the threshold is illustrated in Fig. 4a that shows the observed monthly

mean temperatures  $T_{y,m}$  (black dots) for the grid-point in central Europe (12.5°E, 47.5°N) during the period from 2001 to 2005. The dashed line is the long-term trend ( $L_{y,m}$ ) that represents decadal variability (same as solid thick line in Fig. 1). The solid line is the quantity  $M_{y,m}=L_{y,m}+ S_m$  that is a seasonally varying fit to the observed values, where  $S_m$  is the mean annual cycle estimated as follows:

$$S_m = \frac{1}{n} \sum_{y=1}^n (T_{y,m} - L_{y,m}), \quad (2)$$

where  $n$  is the number of years with available data (e.g.  $n=136$  if data is available for the whole 1870-2005 period). For the quantity  $M_{y,m}$  the mean annual cycle  $S_m$  is constant for all years. The time-varying threshold  $u_{y,m}$  can then easily be defined as  $u_{y,m} = M_{y,m} + \varepsilon$ , where  $\varepsilon$  is a constant increment necessary to have  $\alpha\%$  of the observed values above  $u_{y,m}$ . The constant  $\varepsilon$  is obtained empirically by lifting  $M_{y,m}$  up until  $\alpha\%$  of the observed values are above  $u_{y,m}$ . More precisely, the procedure for obtaining  $\varepsilon$  is:

- Sort the quantity  $T_{y,m} - M_{y,m}$  in ascending order
- Store in vector  $D_j$  positive non-repeated sorted values of  $T_{y,m} - M_{y,m}$ , where  $j=1,2,3,\dots,K$  is the number of non-repeated exceedances  $T_{y,m} > M_{y,m}$
- Add  $D_1$  to  $M_{y,m}$  and check the percentage of  $T_{y,m}$  values above  $M_{y,m} + D_1$
- Add  $D_2$  to  $M_{y,m} + D_1$  and check the percentage of  $T_{y,m}$  values above  $M_{y,m} + D_1 + D_2$
- Keep repeating the procedure above until  $\alpha\%$  of  $T_{y,m}$  values are above  $M_{y,m} + D_1 + D_2 + \dots + D_J$ , where  $J < K$  is the  $j$ -th increment to  $M_{y,m}$  that ensures  $\alpha\%$  of  $T_{y,m}$  values is above  $M_{y,m} + D_1 + D_2 + \dots + D_J$
- The constant  $\varepsilon$  is then computed as  $\varepsilon = \sum_{j=1}^J D_j$

The thick solid segments in Fig 4a show the threshold  $u_{y,m}$  for the summer months with  $\alpha=25\%$ . This threshold guarantees that 25% of the observed summer temperature values are above  $u_{y,m}$ , in an equivalent way as if the constant 75<sup>th</sup> quantile had been chosen but which attempts to have constant exceedance frequency. This threshold  $u_{y,m}$  with  $\alpha=25\%$  is hereafter referred to as the 75% threshold. Exceedances (i.e. events  $T_{y,m} > u_{y,m}$ , when black dots are above the thick solid segments) are noted in the first three years shown in Fig. 4a, with the largest values observed during the summer 2003. The summer 2003 also stands out in Figure 4b, which shows summer excesses  $T_{y,m} - u_{y,m}$  above the 75% threshold  $u_{y,m}$  (vertical bars above the horizontal line) for the same grid-point during the period from 1870 to 2005. Note that exceedances occur throughout the period with fairly stable frequency. From hereon the 75% threshold has been used for all results to be presented.

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 threshold need to be examined. In order to illustrate the typical threshold value at each grid-point, Figure 5 shows the long-term time mean of the 75% 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.

Extreme events can be summarised by presenting sample statistics of the excesses  $T_{y,m} - u_{y,m}$  for the exceedance events where  $T_{y,m} > u_{y,m}$ . Figure 6 shows the sample time mean of excesses, median of the excesses and variance of the excesses. Hot extreme temperatures possess larger mean excesses in colder and more variable extratropical regions than in tropical regions. There are much higher values of the mean, median and variance of excesses over extratropical land areas than over the

oceans and tropical regions. This contrast indicates that monthly temperature excesses are on average larger and have larger variability over extratropical continental areas than over oceanic and tropical regions. The larger excesses over extratropical land regions are most likely due to the much smaller heat capacity of land compared to the oceans (Peixoto and Oort 1993).

## 5. Modelling of extreme events using extreme value theory

This section illustrates how peaks-over-threshold can be used to investigate extreme temperatures. 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

$$Pr(Z \leq z) = H(z) = 1 - \left(1 + \frac{\xi z}{\sigma}\right)^{-\frac{1}{\xi}}, \quad (3)$$

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

$$E(Z) = \frac{\sigma}{1 - \xi}, \quad (4)$$

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

$$Var(Z) = \frac{\sigma^2}{(1 - \xi)^2(1 - 2\xi)}. \quad (6)$$

For the examples presented in this study  $\sigma$  and  $\xi$  are estimated using maximum-likelihood methods (Coles 2001, section 2.6.3). L-moment estimates for  $\sigma$  and  $\xi$  (Hosking and Wallis 1987) have also been explored and provided similar results (not shown).

Figure 7a shows estimates of the scale parameter  $\sigma$  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 (6). Regions with larger values of  $\sigma$  have higher variability of extreme temperatures. In accordance with the variance of excesses shown in Fig. 6c, higher variability of extreme hot temperatures (i.e. large  $\sigma$ ) is found over extratropical continental areas when compared to oceanic and tropical regions. Note also that the scale parameter pattern (Fig. 7a) is similar to the mean of excesses (Fig. 6a) and the median of excesses (Fig. 6b). This similarity is noted because both the mean of excesses (4) and the median of excesses (5) are proportional to  $\sigma$ . The maximum variability of extreme temperatures observed over extratropical continental Europe and Asia coincides with the region where atmospheric blocking is typically observed during the summer (Black *et al.* 2004). Shorter persistence (less than 10 days) of anticyclonic (high pressure) conditions in association with warm air advection from North Africa can also contribute to increasing the variability of hot extreme temperatures over Europe (Nakamura *et al.* 2005).

Figure 7b shows estimates of the shape parameter  $\xi$  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 smaller shape parameter is thinner than in regions with 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  $-\sigma/\xi$  from the GP distribution fit. However, no clear large-scale pattern is observed. Figure 7c shows the upper bound of excesses  $-\sigma/\xi$ . Regions with

null (or nearly null) and positive shape parameters have no bound (i.e. have infinite upper tail) and are shaded in black. Larger upper bounds of excesses  $-\sigma/\xi$  (between 4°C and 8°C) are found in extratropical continental areas (e.g. north of North America, north of Europe and northeast Asia), indicating that excesses over 4°C can be observed in these regions.

When dealing with parametric distributions such as the GP it is always good practise 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), which for this particular application 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 asymptotic assumption of excesses above a sufficiently large threshold approximating a GP distribution is respected. One should critically examine whether the 75% 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, 75% and 60%. 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 p-value less than or equal to  $p$  should equal  $p$  with all points falling on the diagonal line. Figure 8 shows that all curves for the 75% threshold are close to the diagonal line, indicating that the quality of the fit is good. For thresholds lower than 75% the curves fall on the left hand side of the diagonal line as illustrated in Fig. 8 for the 60% threshold, indicating that the GP distribution does not fit well to the data and the asymptotic assumption is not valid for lower thresholds than 75%.

Figure 9a shows excesses during August 2003, the hottest ever recorded monthly mean temperature in Europe (Fig. 1 and Fig. 4a). Excesses up to 3°C are observed in central Europe. This extreme 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 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 increase temperatures locally.

Figure 9b shows the GP distribution return period estimates  $(1-H(z))^{-1}$  for the August 2003 excesses of 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 event (e.g. an excess  $z$  of 3 Celsius) to recur. Some grid-points over Europe have return period between 5 and 10 years, others between 10 and 50 years and some between 50 and 500 years. For example, the return period for the grid-point in central Europe (12.5°E, 47.5°N) is 133 years with 90% confidence interval of (52, 730) years estimated using a bootstrap resampling procedure (Davison and Hinkley 1997). The immediate left neighbour grid-point (7.5°E, 47.5°N) has a return period of 184 years, and the grid-point centred in 2.5°E, 42.5°N in the south of France has the highest return period over Europe of 316 years. These return period estimates are much smaller than the value of 46,000 years over Switzerland obtained by Schär

*et al.* (2004) that used a normal (Gaussian) distribution fitted to the mean June-August temperature during 1990-2002.

## 6. Factors controlling extreme events

The dependence of extremes on factors such as time and ENSO can easily be examined by modelling the shape and scale parameters of the GP distribution as functions of these factors. 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:

$$\log \sigma = \sigma_o + \sigma_l x \quad (7)$$

$$\xi = \xi_o \quad (8)$$

where  $x$  is an ENSO index such as the Southern Oscillation Index (SOI). Note that in (7) the logarithm of  $\sigma$  is used instead of  $\sigma$  to ensure that  $\sigma = e^{(\sigma_o + \sigma_l x)}$  is positive for all choices of parameter values  $\sigma_o$  and  $\sigma_l$ . The three GP distribution parameters  $\sigma_o$ ,  $\sigma_l$  and  $\xi_o$  can then be estimated using maximum-likelihood methods (Coles 2001).

The appropriateness of the model given by (7) and (8) can be tested by performing a likelihood ratio test (Coles 2001, section 2.6.6). The model of (7) and (8) is tested against a simpler nested model given by  $\log \sigma = \sigma_o$  and  $\xi = \xi_o$ . The need for the extra parameter  $\sigma_l$  is tested by the null hypothesis  $H_0: \sigma_l = 0$  against the alternative hypothesis  $H_1: \sigma_l \neq 0$ . Figure 10 shows the map of p-values for the hypothesis test above, where  $x$  is the SOI. 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 tropical and oceanic regions the simple model with constant shape and scale parameters is enough to fit the excesses above the 75% time-varying

threshold, and therefore there is no advantage of using the SOI factor to model hot extreme temperatures over these regions. On the other hand, Fig. 10 shows that over extratropical continental regions in North America, Europe and Asia p-values are less than 0.05 and therefore the null hypothesis can be rejected at the 5% significance level in favour to the alternative hypothesis. These results suggest that SOI is a statistically significant factor for modulating extreme temperature variability over extratropical continental regions. In other words, the variability of hot extreme temperatures over extratropical continental regions is affected by ENSO atmospheric teleconnections.

The model given by (7) and (8) has also been tested against a combination of other, more elaborate models including time  $t$  (e.g. in centuries) and also the square of SOI as covariate factors for the shape and scale parameters to investigate non-linear effects (not shown). In all tests performed the null hypothesis of the simpler model given by (7) and (8) could not be rejected, indicating that time  $t$  and  $x^2$  are less relevant for the modulation of hot extreme temperatures over Northern Hemisphere extratropical continental areas than ENSO alone.

Because of the logarithmic link function in (7) the parameters  $\sigma_o$  and  $\sigma_l$  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  $\sigma$  by  $e^{\sigma_l}$ . Figure 11 shows maps of  $e^{\sigma_o}$  (panel a),  $e^{\sigma_l}$  (panel b) and  $\xi_o$  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  $e^{\sigma_o}$  of Fig. 11a is similar to the map of  $\sigma$  of Fig. 7a and the map of  $\xi_o$  of Fig. 11c is similar to the map of  $\xi$  of Fig. 7b, where  $\sigma$  and  $\xi$  have been estimated as constant parameters. Figure 11b shows that over most of North America, some regions in eastern Europe and Scandinavia  $e^{\sigma_l}$  is larger than one, indicating that

$\sigma$  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 North America (97.5°W, 42.5°N). Figure 12a shows a scatter plot of excesses for this grid-point and the SOI with the median (solid line) and the upper and lower quartiles (dashed lines) of the GP distribution with scale parameter  $\sigma = e^{(\sigma_o + \sigma_1 x)}$  and shape parameter  $\xi = \xi_o$  superimposed. The increased variability of hot extreme temperatures can be noted for larger SOI. Figure 11b also shows that over most of Asia, Russia, and western Europe  $e^{\sigma_1}$  is smaller than one indicating that  $\sigma$  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 west Russia (52.5°E, 57.5°N).

## 7. Temporal clustering of extreme events

The annual frequency of extreme events, e.g. the number of extreme events observed during each summer, is a proxy for clustering of extremes. The average number of summer exceedances  $\bar{n}_e = \frac{1}{N} \sum_{i=1}^{n_i} e$  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 extreme event (i.e. an exceedance) is observed and  $e = 0$  if an extreme event 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 extreme events are more clustered in time (i.e. regions where there is more serial dependence).

Figure 13 shows  $\bar{n}_e$  computed for summer exceedances over the period 1870-2005. A clear contrast between continental and oceanic regions is noted. Extreme temperatures are more clustered over the Atlantic and East Pacific oceans (average of around 1.8 events per year) than over North America, Europe and Asia (average of

around 1.5 events per year). This indicates that temperatures above the threshold  $u_{y,m}$  are more clustered over the oceans than over land mostly because of the longer memory of the oceans when compared to the continents.

## 8. Teleconnections between extreme events

Association of extreme values between different locations (i.e. teleconnection at extreme levels) can be studied using an asymptotic dependence measure  $\chi$  (Buishand 1984; Coles *et al.* 1999) as follows. Suppose we are interested in investigating how extreme monthly mean temperatures at central Europe  $T_E$  are related to extreme monthly mean temperatures 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  $\chi$  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 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 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 re-written 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  $rank(T_E)/N$  and  $rank(T_O)/N$  to obtain  $F_{T_E}(T_E)$  and  $F_{T_O}(T_O)$  one can compute  $\chi(u)$ , where  $rank(.)$  is the rank of the data. For large thresholds (i.e.  $u \rightarrow 1$ ) the measure  $\chi$ , which ranges from 0 to 1, provides a simple measure of extremal dependence between  $T_E$  and  $T_O$ . Larger values of  $\chi$  indicate stronger dependence.

Figure 14a shows the scatter plot of August monthly mean temperatures  $T_E$  in a grid-point in central Europe (12.5°E, 47.5°N) and August monthly mean temperatures  $T_O$  in a grid-point in the west North Atlantic (67.5°W, 42.5°N). The scatter plot 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  $\chi$  is obtained using the points of Fig. 14a that are located on the right-hand side of the 75<sup>th</sup> threshold line of August temperatures in central Europe. The  $\chi$  statistic is given by the ratio between the number of points on the top right-hand corner of the scatter plot (i.e. those points that are located above both 75<sup>th</sup> threshold lines) and the total number of points to the right-hand side of the vertical line. The  $\chi$  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 Figure 14b with  $u = 0.75$  (vertical and horizontal lines). In practice  $\chi$  is computed using (11). Figures 14c and 14d show similar scatter

plots for the grid-point in central Europe (12.5°E, 47.5°N) and a grid-point in west Russia (52.5°E, 57.5°N). No sign of extreme dependence is noticeable between these two grid-points.

Figure 15a shows a map of  $\chi$  for extreme August monthly mean temperatures for the grid-point in central Europe (12.5°E, 47.5°N). As expected, grid-points close to the central Europe grid-point (12.5°E, 47.5°N) have large values of  $\chi$ , indicating strong dependence. The west North Atlantic also shows some dependence, as previously illustrated in Figs. 14a and 14b.  $\chi$  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  $\chi$  ranging from 0.25 and 0.5 means that if the temperature in central Europe exceeds the 75<sup>th</sup> quantile, then there is a 25 to 50% risk that temperature at the west North Atlantic regions will also exceed the 75<sup>th</sup> 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, and also 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 and 14d.

The  $\chi$  statistic provides a measure of extreme dependence for asymptotically dependent distributions (i.e. when  $\chi \neq 0$ ). However, it fails to provide information of discrimination for asymptotically independent distributions, i.e. when  $\chi = 0$  (Coles 2001). An alternative measure is therefore required to overcome this deficiency. Such measure is given by

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

where

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

defined for thresholds on the range  $0 < u < 1$  (Coles *et al.* 1999). The  $\bar{\chi}$  statistics range from  $-1$  to  $1$ . For asymptotically dependent variables  $\bar{\chi}=1$ . For independent variables  $\bar{\chi}=0$ . As  $\chi$  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 zero) then  $\bar{\chi}$  is a more appropriate measure of the strength of extremal dependence. As the correlation coefficient is the standard measure of association between two variables,  $\bar{\chi}$  is the equivalent association measure for extreme events.

Figure 15b shows a map of  $\bar{\chi}$  for extreme August monthly mean temperatures for the grid-point in central Europe (12.5°E, 47.5°N). As noticed in Fig. 15a, hot extreme temperatures in central Europe are strongly associated with extreme temperatures in neighbouring grid-points. Central Europe extreme temperatures are also confirmed to be associated with extreme temperatures in the west North Atlantic (Fig. 14a and 14b). 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 scatter plot of Fig. 14a. Figure 15b still shows a 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.

## 9. Conclusions

This study has presented a number of tools for the investigation of extreme weather and climate events in gridded datasets. These tools allow the study of:

- spatial patterns of simple extreme events;
- relationship between potentially influential factors (e.g. time and ENSO) and extreme events;
- temporal clustering of extreme events;
- teleconnection patterns of extremes.

The methods have been demonstrated using Northern Hemisphere gridded temperature data from CRU covering the period from 1870 to 2005. Motivated by the recent hot summer of 2003, the methods have been applied to summer (June-July-August) monthly mean temperatures. Extreme events have been defined as those months when the mean temperature has fallen above a time-varying threshold  $u_{y,m}$ .

The methods have revealed that hot extreme temperatures:

- have larger variability in extratropical continental regions than in oceanic and tropical regions;
- have variability over tropical and oceanic regions mainly driven by local processes rather than by ENSO atmospheric teleconnections;
- have variability over extratropical continental regions affected by ENSO atmospheric teleconnections;
- have larger variability during La Niña conditions over most of North America, some regions in eastern Europe and Scandinavia;
- have larger variability during El Niño conditions over most of Asia, Russia, and western Europe;

- are more clustered over the Atlantic and East Pacific oceans than over continental regions
- over central Europe during August are related to hot extreme temperatures in the northwestern Atlantic.

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 estimation of the parameters of the GP distribution could be made using not only the data of a single grid-point but instead also using data from neighbouring grid-points. Such an approach with an increased sample size is likely to provide better estimates for the parameters and also smoother (less noisy) spatial maps.

The software used to perform the analysis presented here has been developed as part of the RCLIM initiative (R software for CLIMate analysis) and is freely available at <http://www.met.reading.ac.uk/cag/rclim/>. This initiative has been established within work package 4.3 (Understanding Extreme Weather and Climate Events) of the European Union funded ENSEMBLES project (GOCE-CT-2003-505539). Functions have been written in the R statistical language (<http://www.r-project.org>). In addition to the functions used for the climate analysis of extremes presented here, other functions have been written for reading and writing netcdf gridded datasets, general exploratory climate analysis (e.g. compute one point correlation, principal component analysis, extract subsets of data from a dataset), and animating and plotting climate analysis of gridded datasets.

## **Acknowledgements**

CASC and DJS were supported by the EU-ENSEMBLES project (GOCE-CT-2003-505539, <http://www.ensembles-eu.org/>). CASC was also partially supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) processes number 2005/05210-7 and 2006/02497-6. 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.

## 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. *Geophysical Research Letters*, **31**, L02202, doi:10.1029/2003GL018857

Beniston, M. 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 and Planetary Change*, **44**, 73-81.

Beniston, M. and D. B. Stephenson, 2004: Extreme climate events. Special Issue. *Global and Planetary Change*, **44**, 211pp.

Beniston, M., D. B. Stephenson, O. B. Christensen, C. A. T. Ferro, C. Frei, S. Goyette, K. Halsnaes, T. Holt, K. Jylha, B. Koffi, J. Palutikof, R. Scholl, T. Semmler, and K. Woth, 2006: Future Extreme Events in European Climate: An exploration of regional climate model projections, *Climatic Change*, PRUDENCE special issue, (in press).

Black, E., M. Blackburn, G. Harrison, B. Hoskins and J. Methven, 2004: Factors contributing to the summer 2003 European heatwave. *Weather*, **59**, 217-223.

Buishand T. A., 1984: Bivariate extreme-value data and the station-year method. *J. Hydrology*, **69**, 77-95.

Choulakian, V. and M. A. Stephens, 2001: Goodness-of-fit tests for the generalized Pareto distribution. *Technometrics*, **43**, 478-484.

Coles, S. G., J. Heffernan and J. A. Tawn, 1999: Dependence measures for extreme value analyses. *Extremes*, **2**, 339-365.

Coles, S. G., 2001: An Introduction to Statistical Modeling of Extreme Values. Springer series in statistics. 208pp.

Davison, A. C. and D. V. Hinkley, 1997: Bootstrap methods and their application, Cambridge University Press, New York, 582 pp.

Ferro, C. A. T., 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. *Journal of Hydrology*, **300**, 212-233.

Hosking, J. R. M. and J. R. Wallis, 1987: Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics*, **29** (3), 339-349.

IPCC, 2001: Intergovernmental Panel on Climate Change Third Assessment Report. Available from <http://www.ipcc.ch/>

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.

Kharin, V. V. and F. W. Zwiers, 2005: Estimating extremes in transient climate change simulations. *J. Climate*, **18**, 1156-1173.

Kharin, V. V., F. W. Zwiers 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. 342pp.

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., Enomoto, T. and Yamane, S., 2005: A simulation study of the 2003 heatwave in Europe. *Journal of the Earth Simulator*, **2**, 55-69. Available at [www.es.jamstec.go.jp/esc/images/jes\\_vol.2/pdf/JES2\\_nakamura.pdf](http://www.es.jamstec.go.jp/esc/images/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. Geoscience*, **337**, 1013-1022.

Peixoto, J. P. and Oort, A. H., 1993: Physics of climate. *American Institute of Physics*. 564pp. ISBN: 0883187124.

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 on regional climate I: An Aerological study of blocking action. *Tellus*, **3**, 196-211.

Schär, C., P. L. Vidale, D. Luthi, C. Frei, C. Haberli, M. A. Lineger 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. *Atmosphere-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. doi: 10.1038/nature03089

Svensson, C. and D. A. Jones, 2002: Dependence between extreme sea surge, river flow and precipitation in eastern Britain. *International Journal of Climatology*, **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. *International Journal of Climatology*, **26**, 743-769.

Zwiers, F. W., V. V. Kharin, 1998: Changes in the extremes of the climate simulated by CCC GCM2 under CO<sub>2</sub> doubling. *J. Climate*, **11**, 2200-2222.

## List of Figures

Figure 1: Summer (June-July-August) monthly mean temperatures  $T$  (Celsius) from 1870-2005 at a grid-point in central Europe (12.5°E, 47.5°N). Each of the  $n_s = 408$  vertical bars 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 75<sup>th</sup> quantile of summer monthly mean temperatures of 16.2°C. The solid thick line is the long-term trend estimated with a local polynomial fit with sliding window of 10 years using all monthly mean temperature values from 1870 to 2005.

Figure 2: a) Mean ( $\bar{T}$ ), b) standard deviation ( $s$ ) and c) skewness ( $b$ ), of monthly mean summer temperatures estimated over the 1870-2005 period. White shading signifies missing values.

Figure 3: Maxima of monthly mean summer temperatures over the 1870-2005 period.

Figure 4: a) Observed monthly mean temperatures  $T_{y,m}$  (black dots) for the grid-point in central Europe (12.5°E, 47.5°N) during the period from 2001 to 2005. The dashed line is the long-term trend ( $L_{y,m}$ ) that represents decadal variability (same as the solid thick line in Fig. 1). The thin solid line is the quantity  $M_{y,m} = L_{y,m} + S_m$ , (see text for explanation). The thick solid segments are the time-varying threshold  $u_{y,m}$  given by  $u_{y,m} = M_{y,m} + \varepsilon$ , where  $\varepsilon$  is the increment necessary to have  $\alpha\%$  of the observed values above  $u_{y,m}$ . The thick solid segments is the 75% threshold  $u_{y,m}$  for the summer months ( $\alpha=25\%$ ). b) Summer month differences between  $T_{y,m}$  and  $u_{y,m}$  (75% time-varying threshold) for the grid-point in central Europe (12.5°E, 47.5°N) over the period from

1870 to 2005. Vertical bars above the horizontal solid line (zero line) are excesses  $T_{y,m} - u_{y,m}$  above the threshold  $u_{y,m}$ .

Figure 5: Time mean of 75% threshold  $u_{y,m}$  for the summer months over the period from 1870 to 2005.

Figure 6: a) Mean of excesses, b) the median of excesses and c) variance of excesses above the 75% time-varying threshold  $u_{y,m}$ .

Figure 7: Summer monthly mean temperature GP a) scale parameter  $\sigma$ , b) shape parameter  $\xi$ , and c) upper bound  $-\sigma/\xi$  of excess  $T_{y,m} - u_{y,m}$  above the 75% time-varying threshold  $u_{y,m}$ . Grid-points with  $\xi \geq 0$  have infinite upper bound and are shaded in black.

Figure 8: Percentage of grid-points with Anderson-Darling (AD, dotted lines), Kolmogorov-Smirnov (KS, dashed lines) and Cramér-von Mises (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, 60% (grey curves) and 75% (black curves) of the summer values falling below the threshold.

Figure 9: a) Excesses  $T_{y,m} - u_{y,m}$  above the 75% 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 scale and shape parameters estimates of Figs. 7a and 7b, respectively.

Figure 10:  $P$ -values for the likelihood ratio test for model  $\log \sigma = \sigma_o + \sigma_I x$  and  $\xi = \xi_o$  against a simpler nested model given by  $\log \sigma = \sigma_o$  and  $\xi = \xi_o$ . The need for the extra parameter  $\sigma_I$  is tested by the null hypothesis  $H_0: \sigma_I = 0$  against the alternative hypothesis  $H_1: \sigma_I \neq 0$ .

Figure 11: a)  $\exp(\sigma_o)$ , b)  $\exp(\sigma_I)$  and c)  $\xi_o$  estimated using summer temperature excesses  $T_{y,m} - u_{y,m}$  above the 75% time-varying threshold during the period 1882-2005.

Figure 12: Scatter plot of excesses  $T_{y,m} - u_{y,m}$  above the 75% time-varying threshold and SOI for two grid-points: a) over North America (97.5°W, 42.5°N) and b) over west Russia (52.5°E, 57.5°N). Median (solid line), upper and lower quartiles (dashed lines) of the GP distribution with scale parameter  $\sigma = e^{(\sigma_o + \sigma_I x)}$  and shape parameter  $\xi = \xi_o$ .

Figure 13: Average number of summer exceedances  $\bar{n}_e$  obtained using the time-varying 75% threshold  $u_{y,m}$  over the period 1870-2005.

Figure 14: a) Scatter plot of August monthly mean temperatures  $T_E$  in a grid-point in central Europe (12.5°E, 47.5°N) and August monthly mean temperatures  $T_O$  in a grid-point on the west North Atlantic (67.5°W, 42.5°N). The vertical and horizontal lines are the 75<sup>th</sup> quantile of August monthly mean temperatures in each grid-point, respectively. b) Scatter plot 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  $u = 0.75$ . Panels c)

and d) are similar to panels a) and b) but now for the grid-point in central Europe (12.5°E, 47.5°N) and a grid-point in west Russia (52.5°E, 57.5°N).

Figure 15: a)  $\chi$  and b)  $\bar{\chi}$  for August monthly mean temperatures for the grid-point in central Europe (12.5°E, 47.5°N) with  $u = 0.75$ .

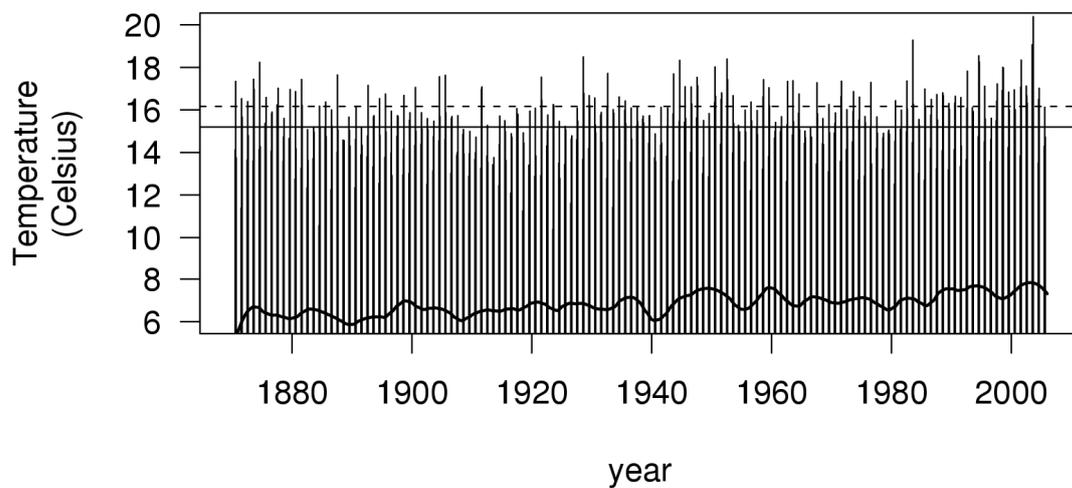


Figure 1: Summer (June-July-August) monthly mean temperatures  $T$  (Celsius) from 1870-2005 at a grid-point in central Europe ( $12.5^{\circ}\text{E}$ ,  $47.5^{\circ}\text{N}$ ). Each of the  $n_s = 408$  vertical bar is the monthly mean for a particular summer month. The horizontal solid line is the long-term (1870-2005) summer monthly mean temperature of  $15.2^{\circ}\text{C}$ . The horizontal dashed line is the 75<sup>th</sup> quantile of summer monthly mean temperatures of  $16.2^{\circ}\text{C}$ . The solid thick line is the long-term trend estimated with a local polynomial fit with sliding window of 10 years using all monthly mean temperature values from 1870 to 2005.

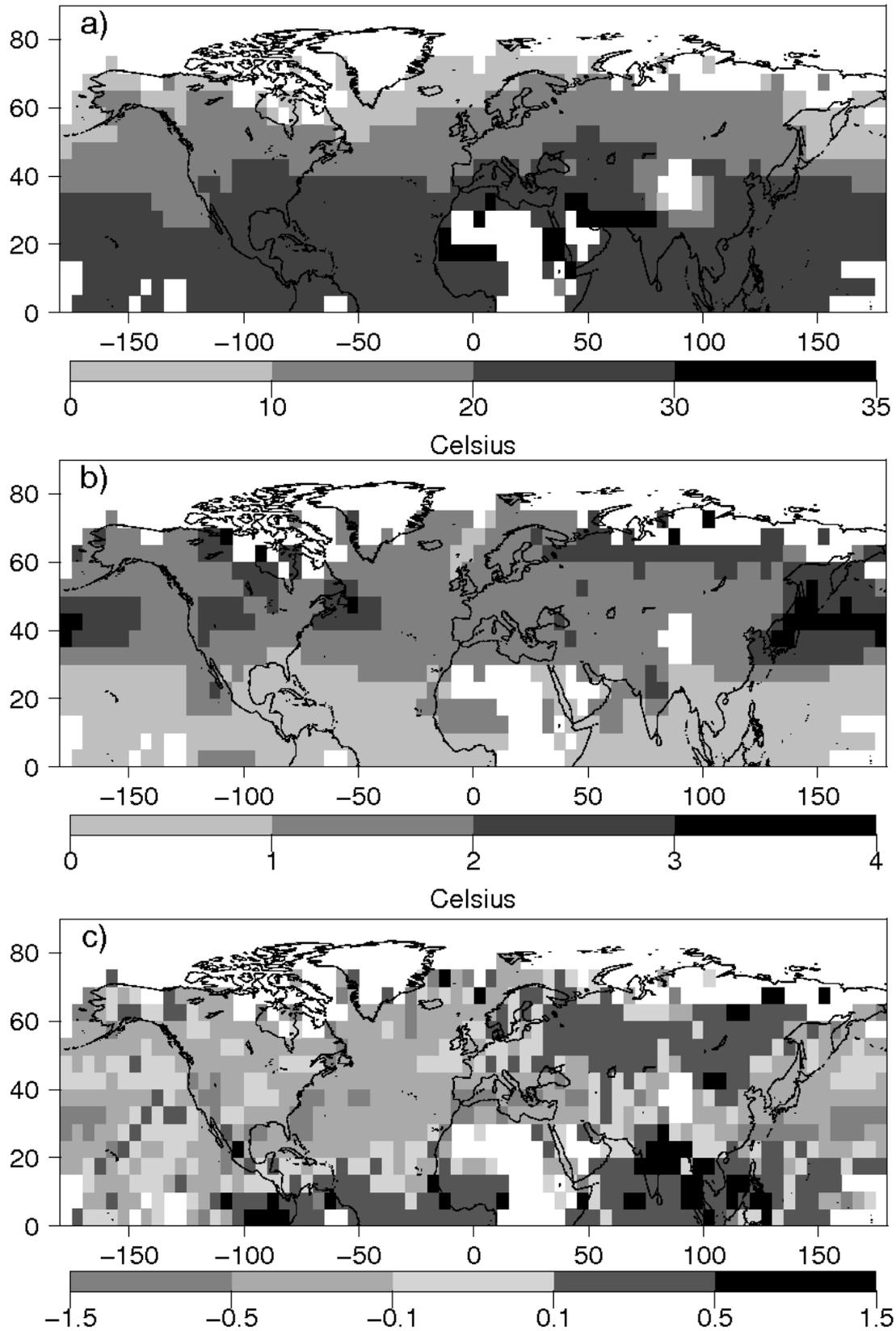


Figure 2: a) Mean ( $\bar{T}$ ), b) standard deviation ( $s$ ) and c) skewness ( $b$ ), of monthly mean summer temperatures estimated over the 1870-2005 period. White shading signifies missing values.

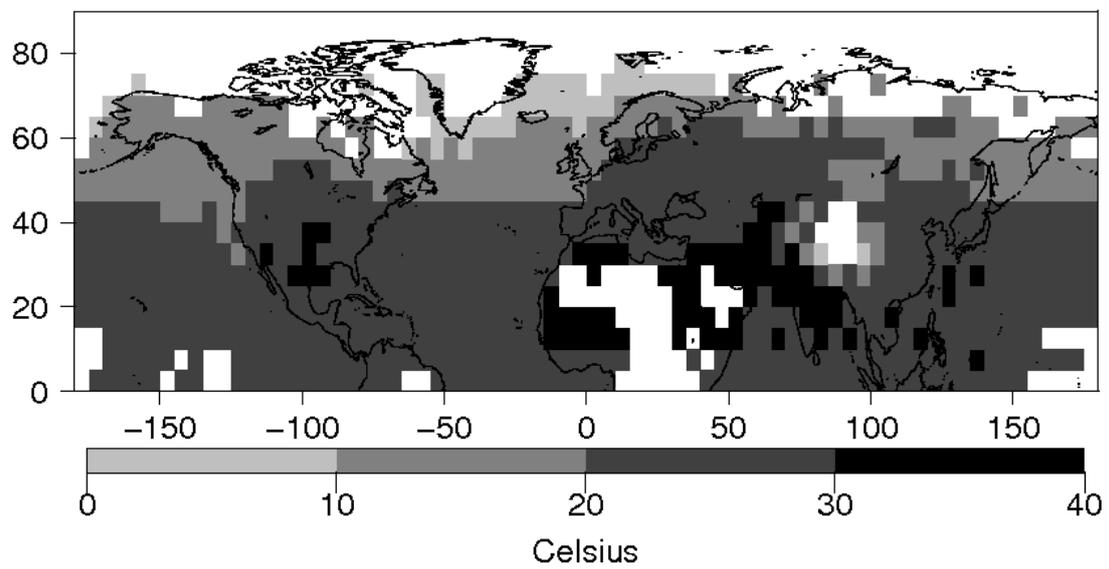


Figure 3: Maxima of monthly mean summer temperatures over the 1870-2005 period.

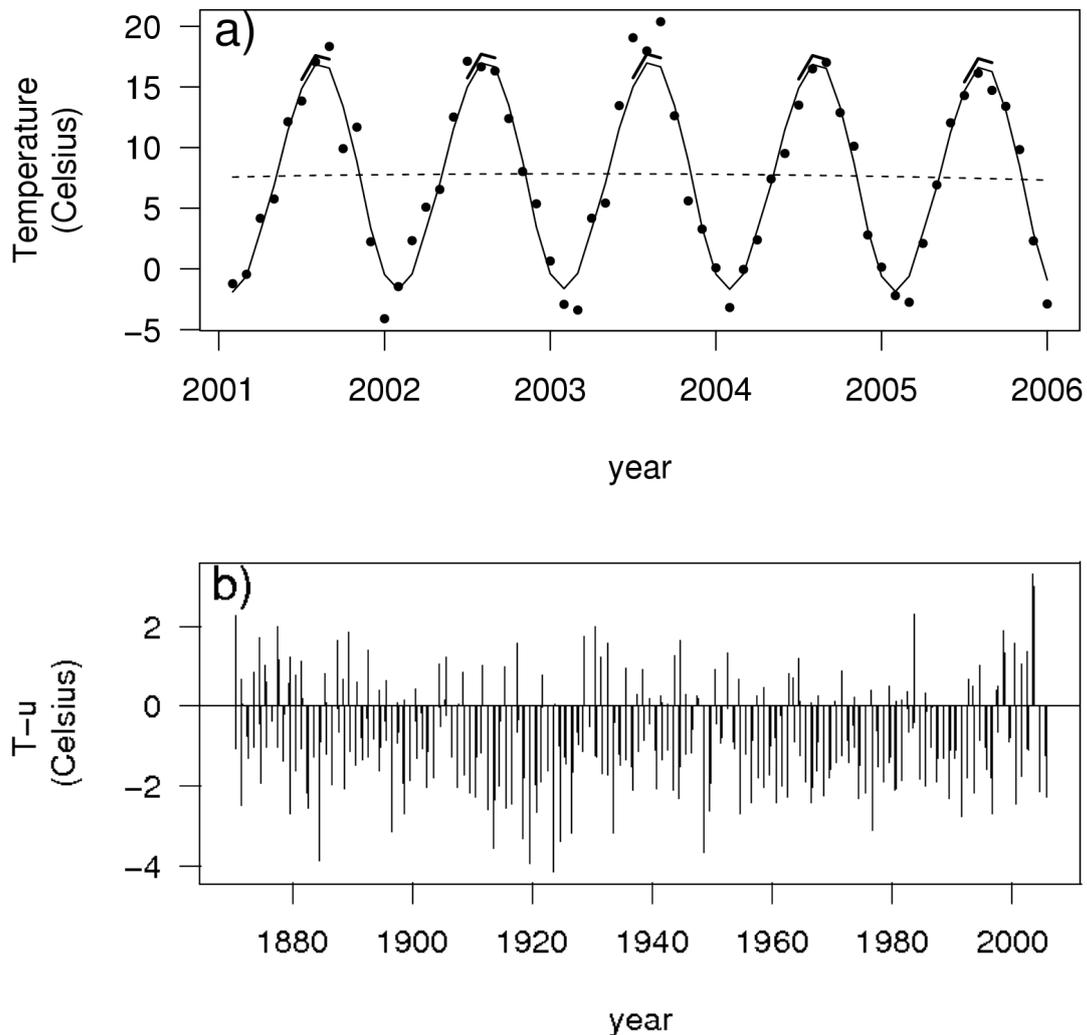


Figure 4: a) Observed monthly mean temperatures  $T_{y,m}$  (black dots) for the grid-point in central Europe ( $12.5^{\circ}\text{E}$ ,  $47.5^{\circ}\text{N}$ ) during the period from 2001 to 2005. The dashed line is the long-term trend ( $L_{y,m}$ ) that represents decadal variability (same as the solid thick line in Fig. 1). The thin solid line is the quantity  $M_{y,m} = L_{y,m} + S_m$ , (see text for explanation). The thick solid segments are the time-varying threshold  $u_{y,m}$  given by  $u_{y,m} = M_{y,m} + \epsilon$ , where  $\epsilon$  is the increment necessary to have  $\alpha\%$  of the observed values above  $u_{y,m}$ . The thick solid segments is the 75% threshold  $u_{y,m}$  for the summer months ( $\alpha=25\%$ ). b) Summer month differences between  $T_{y,m}$  and  $u_{y,m}$  (75% time-varying threshold) for the grid-point in central Europe ( $12.5^{\circ}\text{E}$ ,  $47.5^{\circ}\text{N}$ ) over the period from 1870 to 2005. Vertical bars above the horizontal solid line (zero line) are excesses  $T_{y,m} - u_{y,m}$  above the threshold  $u_{y,m}$ .

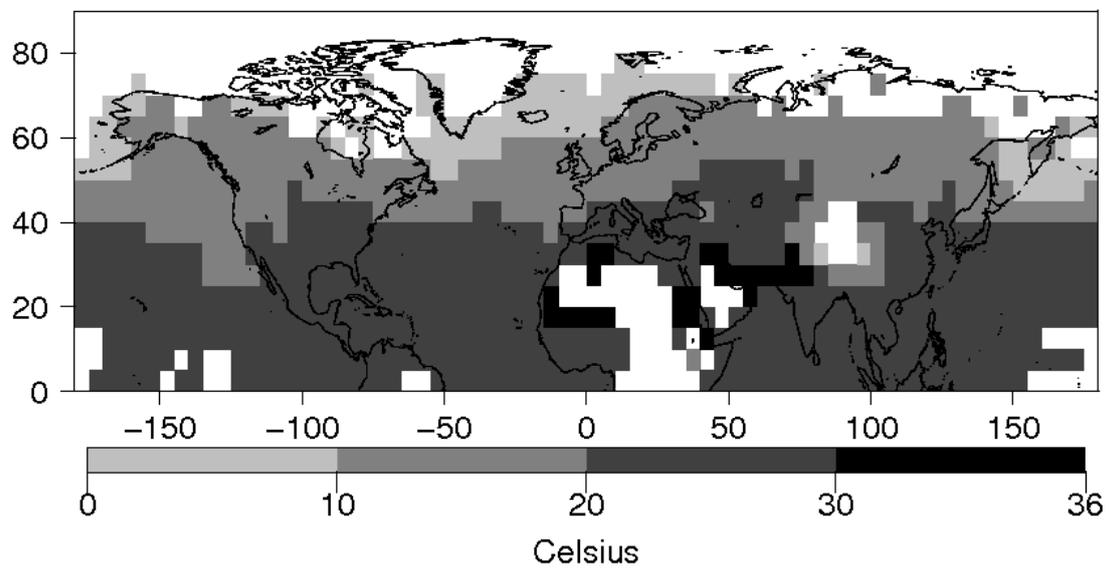


Figure 5: Time mean of 75% threshold  $u_{y,m}$  for the summer months over the period from 1870 to 2005.

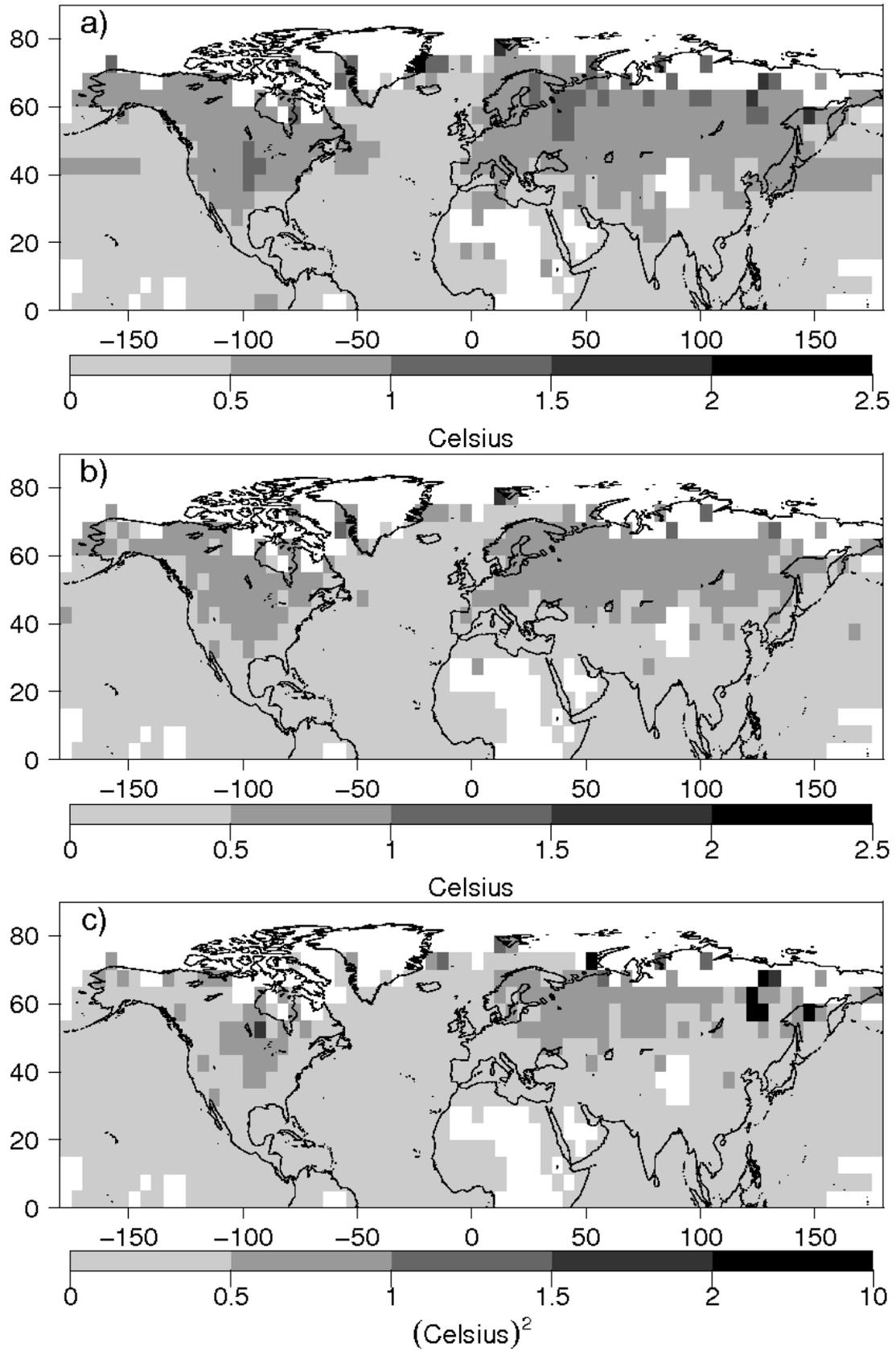


Figure 6: a) Mean of excesses, b) the median of excesses and c) variance of excesses above the 75% time-varying threshold  $u_{y,m}$ .

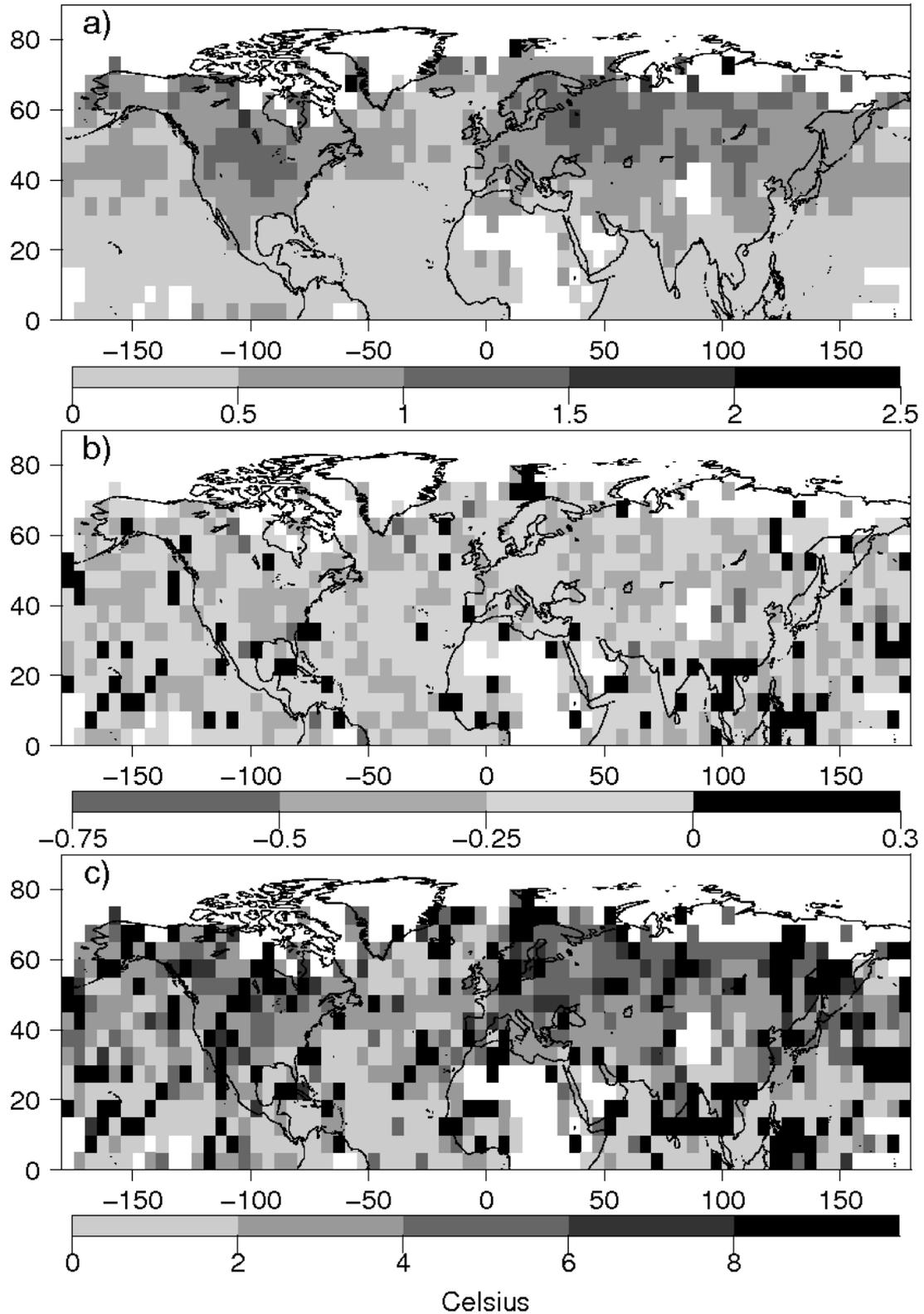


Figure 7: Summer monthly mean temperature GP a) scale parameter  $\sigma$ , b) shape parameter  $\xi$ , and c) upper bound  $-\sigma/\xi$  of excess  $T_{y,m} - u_{y,m}$  above the 75% time-varying threshold  $u_{y,m}$ . Grid-points with  $\xi \geq 0$  have infinite upper bound and are shaded in black.

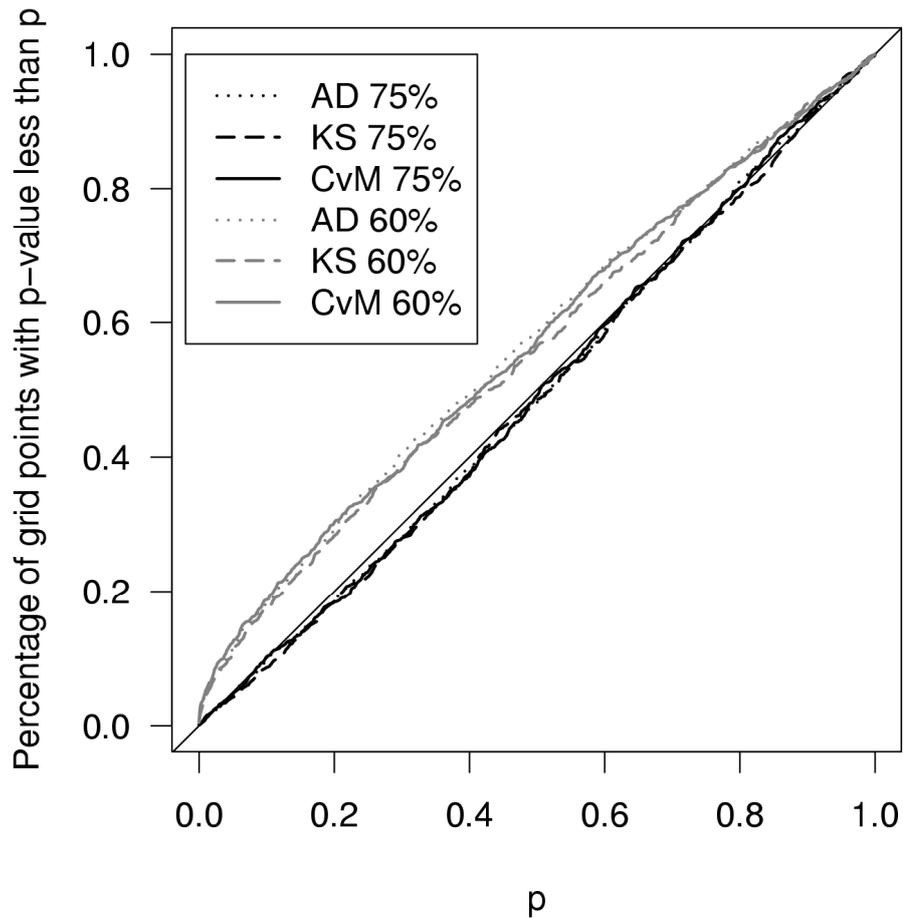


Figure 8: Percentage of grid-points with Anderson-Darling (AD, dotted lines), Kolmogorov-Smirnov (KS, dashed lines) and Cramér-von Mises (CvM, solid lines) test statistics probability value ( $p$ -value) less than or equal to  $p$  (a probability value between 0 and 1) for two choices of time-varying thresholds, 60% (grey curves) and 75% (black curves) of the summer values falling below the threshold.

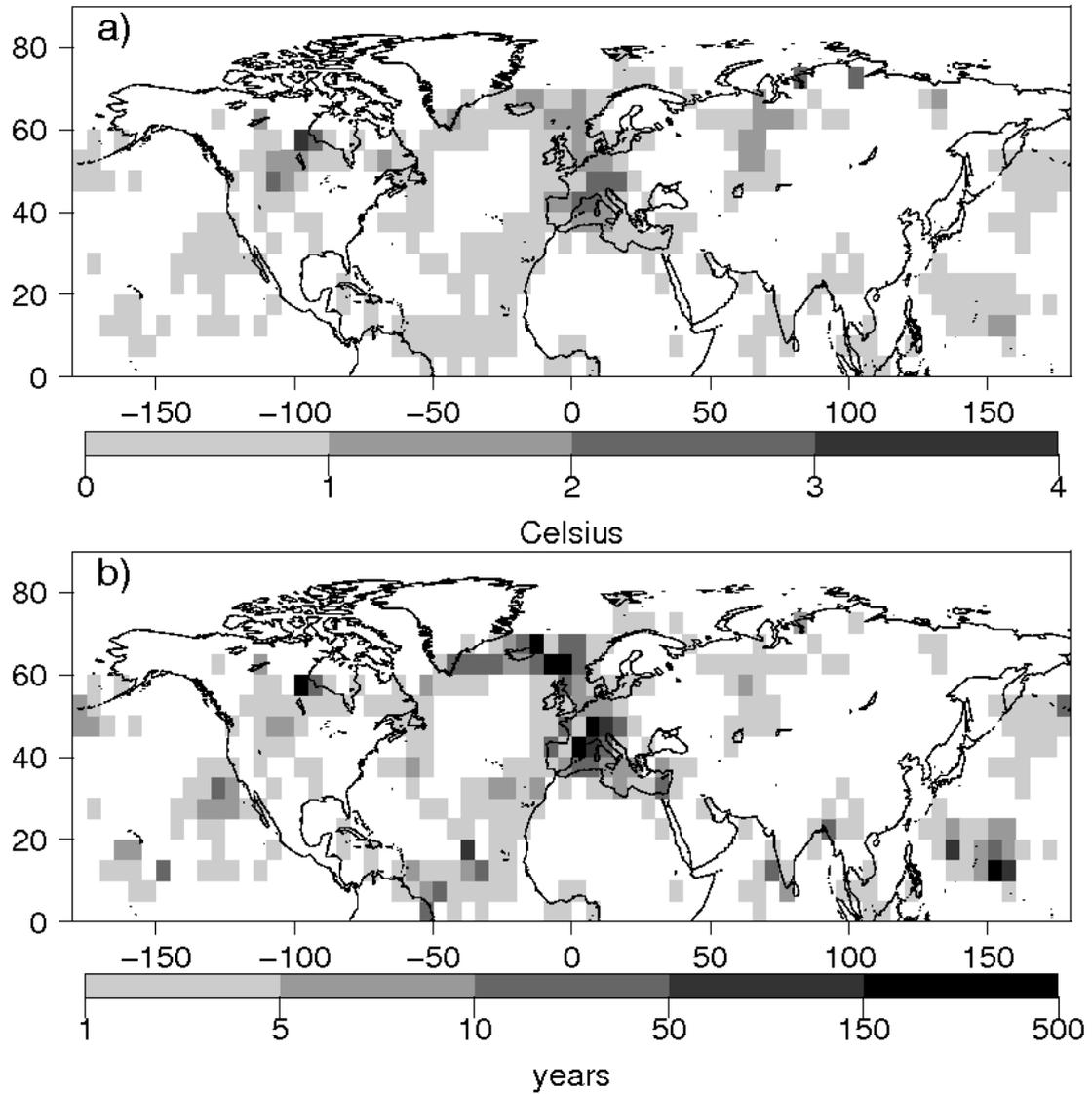


Figure 9: a) Excesses  $T_{y,m} - u_{y,m}$  above the 75% threshold  $u_{y,m}$  during August 2003 (i.e.  $T_{y,m} - u_{y,m}$  when  $T_{y,m} > u_{y,m}$ ). b) Return period estimates for the August 2003 excesses using the GP distribution with scale and shape parameters estimates of Figs. 7a and 7b, respectively.

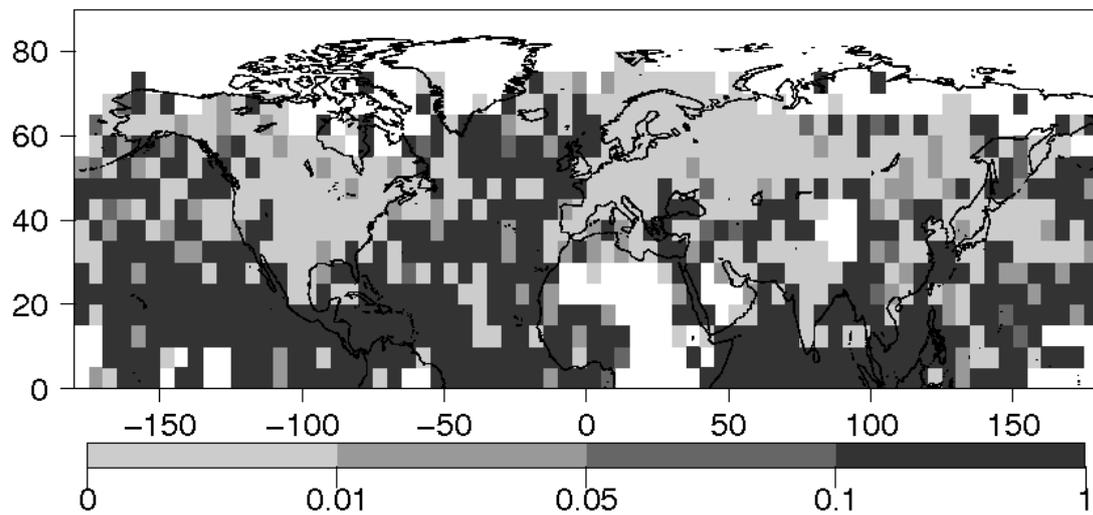


Figure 10:  $P$ -values for the likelihood ratio test for model  $\log \sigma = \sigma_o + \sigma_l x$  and  $\xi = \xi_o$  against a simpler nested model given by  $\log \sigma = \sigma_o$  and  $\xi = \xi_o$ . The need for the extra parameter  $\sigma_l$  is tested by the null hypothesis  $H_0: \sigma_l = 0$  against the alternative hypothesis  $H_1: \sigma_l \neq 0$ .

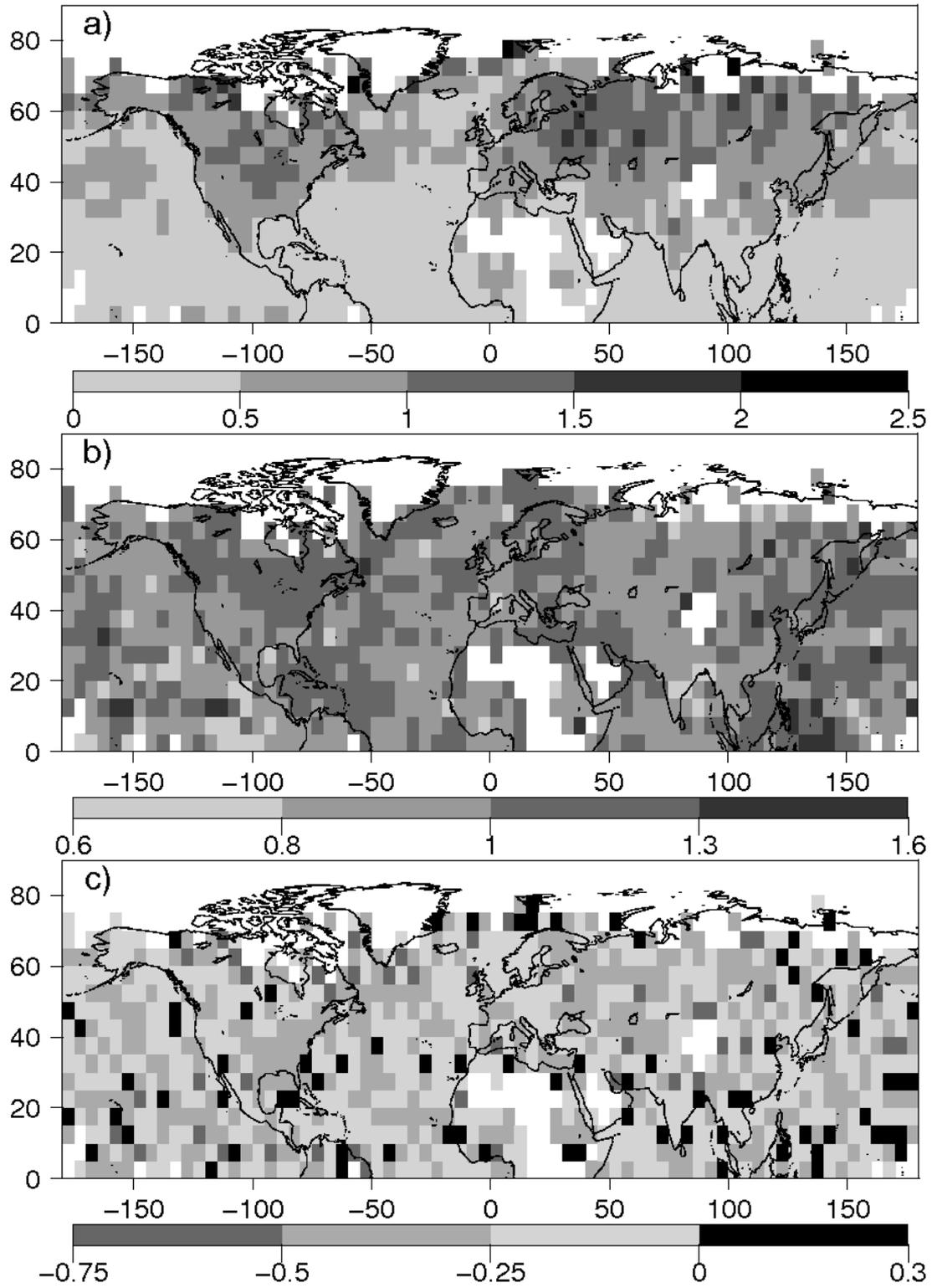


Figure 11: a)  $\exp(\sigma_0)$ , b)  $\exp(\sigma_1)$  and c)  $\xi_0$  estimated using summer temperature excesses  $T_{y,m} - u_{y,m}$  above the 75% time-varying threshold during the period 1882-2005.

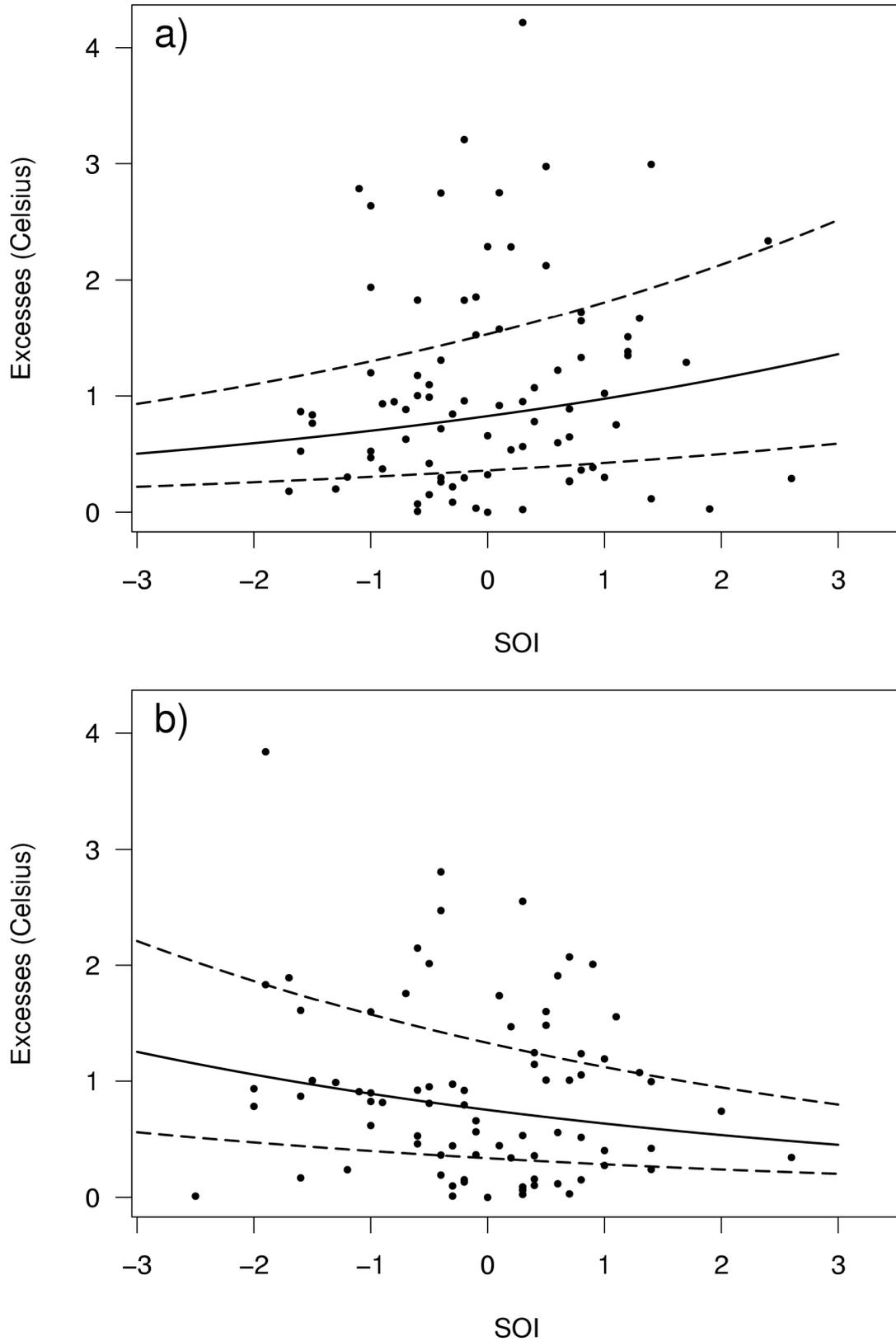


Figure 12: Scatter plot of excesses  $T_{y,m} - u_{y,m}$  above the 75% time-varying threshold and SOI for two grid-points: a) over North America (97.5°W, 42.5°N) and b) over west Russia (52.5°E, 57.5°N). Median (solid line), upper and lower quartiles (dashed lines) of the GP distribution with scale parameter  $\sigma = e^{(\sigma_o + \sigma_1 x)}$  and shape parameter  $\xi = \xi_o$ .

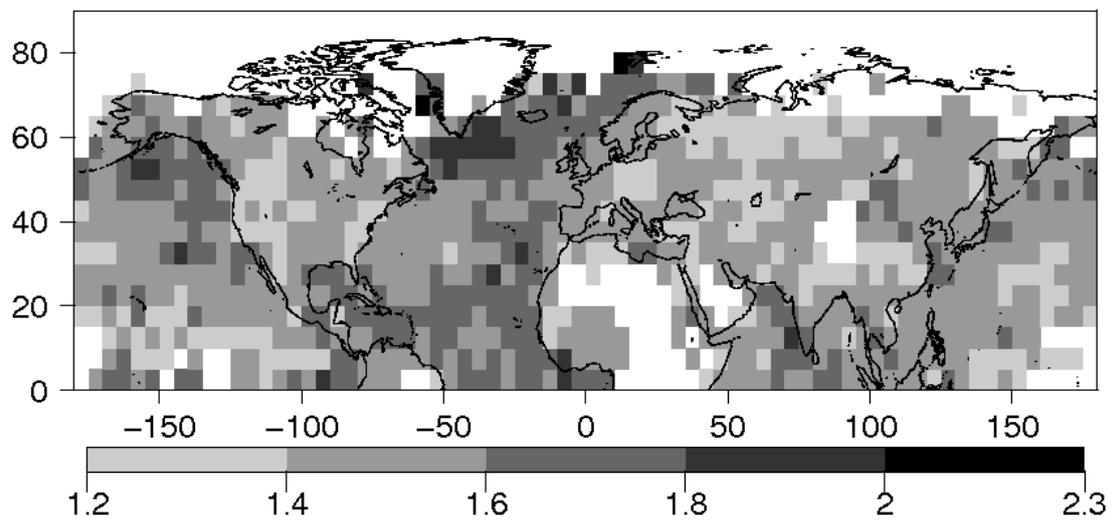


Figure 13: Average number of summer exceedances  $\bar{n}_e$  obtained using the time-varying 75% threshold  $u_{y,m}$  over the period 1870-2005.

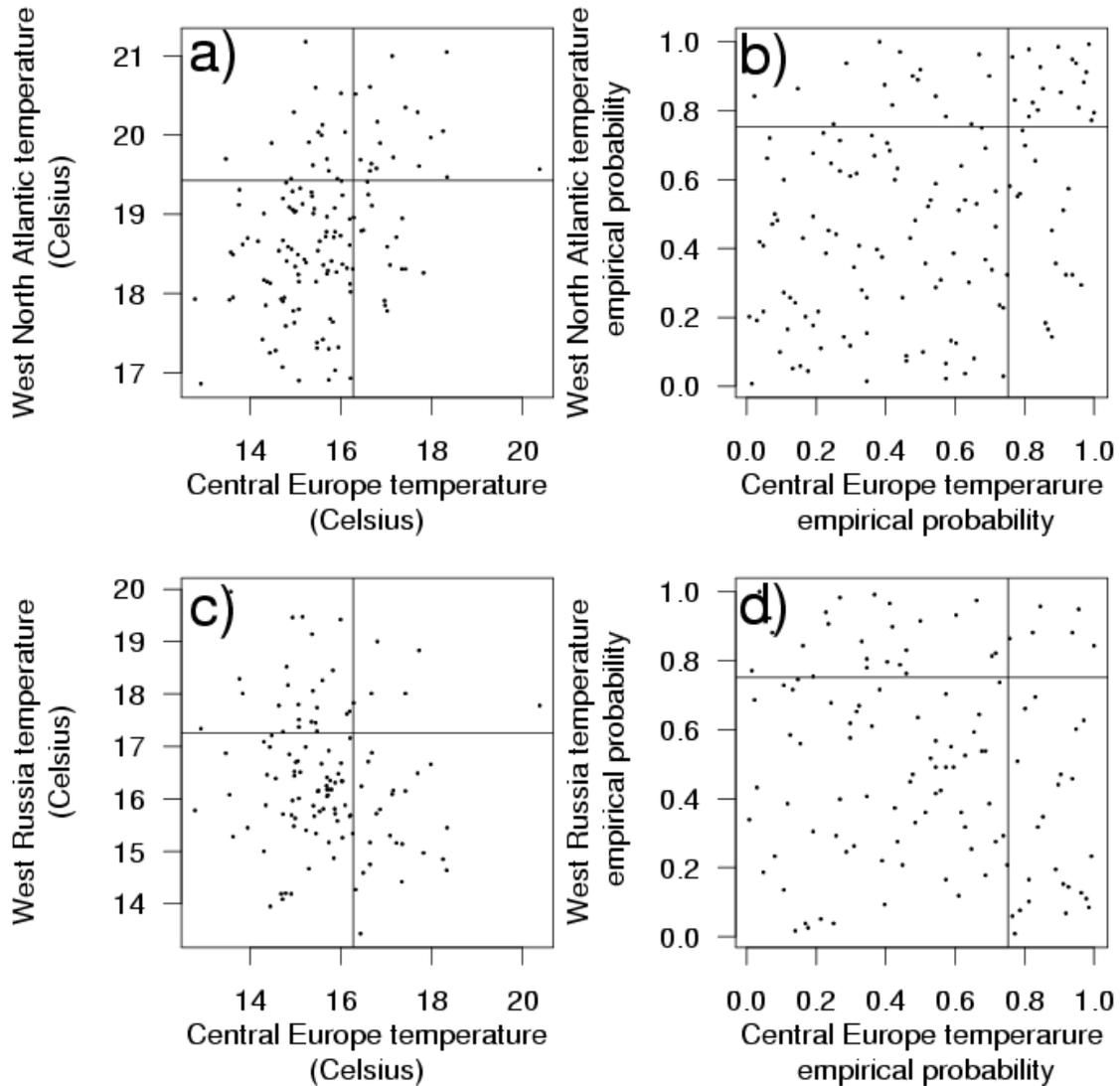


Figure 14: a) Scatter plot of August monthly mean temperatures  $T_E$  in a grid-point in central Europe ( $12.5^\circ\text{E}$ ,  $47.5^\circ\text{N}$ ) and August monthly mean temperatures  $T_O$  in a grid-point on the west North Atlantic ( $67.5^\circ\text{W}$ ,  $42.5^\circ\text{N}$ ). The vertical and horizontal lines are the 75<sup>th</sup> quantile of August monthly mean temperatures in each grid-point, respectively. b) Scatter plot 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  $u = 0.75$ . Panels c) and d) are similar to panels a) and b) but now for the grid-point in central Europe ( $12.5^\circ\text{E}$ ,  $47.5^\circ\text{N}$ ) and a grid-point in west Russia ( $52.5^\circ\text{E}$ ,  $57.5^\circ\text{N}$ ).

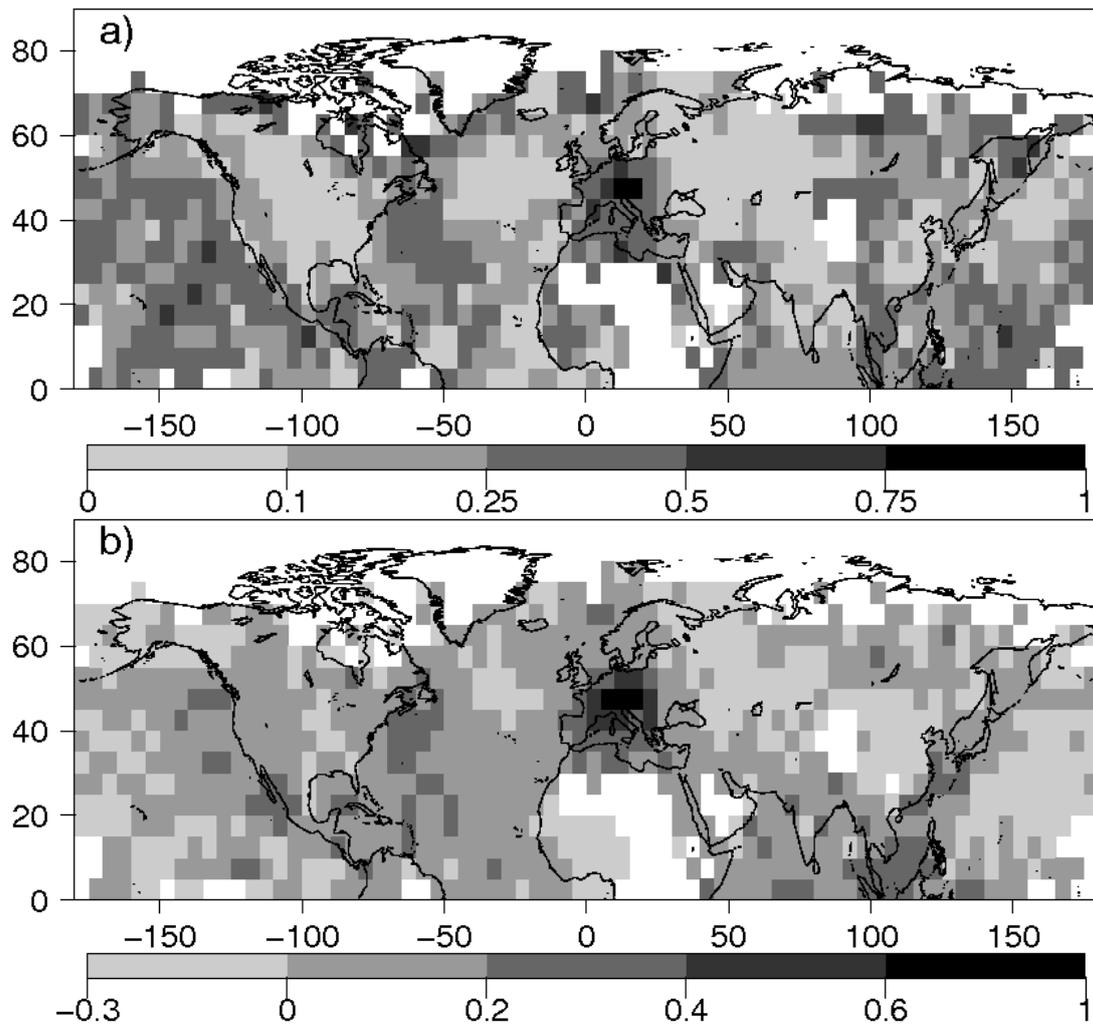


Figure 15: a)  $\chi$  and b)  $\bar{\chi}$  for August monthly mean temperatures for the grid-point in central Europe (12.5°E, 47.5°N) with  $u = 0.75$ .