Project no. GOCE-CT-2003-505539

Project acronym: ENSEMBLES

Project title: ENSEMBLE-based Predictions of Climate Changes and their Impacts

Instrument: Integrated Project
Thematic Priority: Global Change and Ecosystems

Deliverable D4.3.1c: Software for exploring extreme events in gridded data sets

Due date of deliverable: month 18
Actual submission date: 17 March 2006

Start date of project: 1 September 2004 Duration: 60 Months
Organisation name of lead contractor for this deliverable: UREADMM
Revision [draft, 1, 2, ..]

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	**X**
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the Consortium (including the Commission Services)	

Statistical methods for exploring extreme events

In months 1-18 of the ENSEMBLES project, UREADMM in collaboration with NERSC have developed and tested statistical methods for exploring extreme events in gridded data sets (milestone M4.3.1).

These exploratory tools have been written up in an accompanying report for ENSEMBLES deliverable D4.3.1 that we hope to publish after feedback and revision.

Software for these methods has been coded up in the R statistical language (www.r-project.org) and can be freely obtained from the R software for CLIMate analysis (RCLIM) website:

http://www.met.rdg.ac.uk/cag/rclim

We believe that these methods will be extremely useful for other partners in WP4.3 and other work packages of the project. Please contact us if you plan to use this software and then we will be able to provide you with expert assistance.

David B. Stephenson
University of Reading
(email: d.b.stephenson@reading.ac.uk)

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

C. A. S. Coelho, C. A. T. Ferro, D. B. Stephenson

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

D. J. Steinskog

Nansen Environmental and Remote Sensing Center, Bergen, Norway.

Abstract

This report reviews and introduces new tools for the analysis of extreme climate and weather events in gridded datasets. These tools allow exploratory analysis of spatial patterns of extremes, the investigation of relationships between extremes and potentially influential factors (e.g. time, ENSO), the analysis of clusters of extremes in time and also the analysis of spatial dependence of extremes (teleconnection patterns). Methods are demonstrated using Northern Hemisphere monthly mean gridded temperature for June-July-August (JJA) summers from 1870 to 2005. Results show that hot extreme temperatures have larger variability in extratropical continental regions than in oceanic and tropical regions. No clear spatially coherent large-scale patterns are found in the analysis of potentially influential factors. This analysis indicates, however, a tendency towards more extreme events with smaller variability in central Europe. Transitions from the negative to the positive phase of the Southern Oscillation are found to amplify the variability and reduce the frequency of hot extreme events in central Europe. The Atlantic and East Pacific oceans show higher clustering of hot extreme events than continental regions. Hot extreme temperatures over central Europe during August are found to be related to hot extreme temperatures in the west North Atlantic. The software developed to perform the analysis presented here is freely available from http://www.met.reading.ac.uk/cag/rclim/. This report is deliverable D4.3.1c of the work package WP4.3 (Understanding extreme weather and climate events) in the EU project ENSEMBLES (http://www.ensembles-eu.org/). Comments on this report would be most appreciated (c.a.d.s.coelho@reading.ac.uk).

1. Introduction

Historical climate and weather time series stored in large grid point arrays can be analysed in several different ways. For example, atmospheric teleconnections are often investigated by producing composition maps or maps of correlations between the time series of a grid point and the time series of all the other grid points of the array. 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). These standard methods, however, are not appropriate for exploring extreme climate and weather events in gridded datasets for several reasons: a)

these methods summarise all events (the whole distribution) and therefore mask extreme events (tail of the distribution), b) the rarity of extreme events reduces substantially the sample size of data available for analysis and c) these methods generally assume that the data is normally (Gaussian) distributed, an assumption not valid for extreme events. Given the increasing availability of gridded datasets in recent years, there is therefore the need of developments of statistical tools for exploring climate and weather events in large gridded datasets. This report aims to review and introduce new tools for exploring extreme events in gridded datasets.

Only a few studies have analysed extreme weather in gridded datasets. Zwiers and Kharin (1998), Kharin and Zwiers (2000, 2005), van den Brink *et al.* (2004), Fowler *et al.* (2005) and Kharin *et al.* (2005) fitted 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, climate model simulations, and also for station data. These studies have mainly focused on analysis of maps of changes over time in the GEV estimated parameters and also on maps of return values (e.g. the 20-year return value is the value that is estimated to be exceeded once every 20 years on average). Naveau et al. (2005) fitted the generalized Pareto (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 themohaline circulation. Beniston and Stephenson (2004) gather several articles examining climate and weather extremes in gridded datasets. Beniston *et al.* (2006) present a variety of diagnostic methods to determine how heat waves, heavy precipitation, droughts, wind storms, and storm surges change between 1961-90 and 2071-2100 regional climate model simulations produced by the EU-PRUDENCE project.

This study will explore hot extremes in monthly mean temperatures. The study will focus on the investigation of *simple extreme* events. According to IPCC (2001) a simple extreme event is defined as an individual local weather variable exceeding a critical level on a continuous scale. The peaks-over-threshold approach (Coles 2001, chapter 4) will be used to define monthly mean temperature extreme events at each grid point over the Northern Hemisphere. In this approach a high critical level is chosen and the extreme events used in the analysis are those events with values above the critical level. More precisely, the distribution of the values above the critical level at each grid point is examined.

Statistical software is freely available for performing extreme value analysis. This software, however, is not adapted specifically for analysis of gridded datasets. As part of this study a number of functions have been written so that climate researchers can easily explore climate and weather extremes in gridded datasets. These functions will allow us to address the following scientific questions:

- How the risk from extremes of different intensities varies in space and time?
- How extremes depend on time-varying factors (e.g. time, ENSO, NAO)?
- How extremes cluster in time at different locations?

- How extremes at one location are related to extremes in another location? i.e. Are there teleconnections at extreme levels?

Section 2 motivates the study of extreme monthly mean temperatures, describes the dataset used in the investigation and defines basic concepts required for the interpretation of the results. Section 3 presents the methods used to explore extremes in gridded datasets and illustrates each method with an example of application. Section 4 summarises results and suggests future developments.

2. Motivation, dataset and definitions

The European heat wave in summer 2003 drew the attention of the climate science community since this event reflected the extremes of temperature that are projected to occur in later decades of the 21st century (Beniston, 2004; Beniston and Diaz, 2004; Meehl and Tebaldi, 2004; Schär et al., 2004; Stott *et al.*, 2004). Over 20000 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 recent 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.

This study uses monthly mean gridded surface temperature data (HadCRUT2v) from the Climate 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^{o}x5^o 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 hot temperate 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 section 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 the 1870-2005 time series of summer (June-July-August) monthly mean temperatures (T) for a grid point in central Europe (12.5^{o}E, 47.5^{o}N). Each 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. The horizontal solid line is the long term (1870-2005) summer monthly mean temperature of 15.2^{o}C. The horizontal dashed line is the 75th quantile (u) of summer monthly mean temperatures of 16.2^{o}C. If one uses the quantile u as the critical threshold level to define

extreme hot events then all the values exceeding such a level are simple extreme events. The event $T > u$, i.e. when a variable T exceeds a pre-defined threshold u, is known as an *exceedance*. The amount $T - u$ by which T exceeds a pre-defined threshold u is known as a threshold *excess*. These excesses can then be used to study the distribution of hot extreme events. The solid thick line in Fig. 1 is the long term mean trend 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. This thick line therefore represents the observed decadal variability, which shows an increasing trend with recent years warmer than those years in the 19th century. Because of the existence of this time trend and also due to the existence of the annual cycle within the summer it is then more appropriate to define extreme events using a time-varying threshold that also incorporates the long term trend. Such time-varying threshold ensures: a) approximately constant exceedance frequency; b) analysis is not biased towards the warmer climate of the end of the 20th century; c) excesses are yielded relative to contemporary climate, and are therefore designed to reflect effects of similar processes at all times. This study will investigate the distribution of excesses above a time-varying threshold estimated as illustrated in section 3a.

3. Methods and results

a) Local climatology of extremes

Before studying extreme hot events it is worth describing the whole distribution of monthly mean temperature values at each grid point, which can then be related to extremes (i.e. the tail of the distribution). Figure 2 shows estimates of the first three moments (mean, standard deviation and skewness) of the distribution of summer monthly mean temperature over the Northern Hemisphere. Skewness b is given by

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

where T is a monthly mean summer temperature value, $\overline{T}$ and s are the long term 1870-2005 summer temperature mean and standard deviation, respectively, and n_s is the number of summer months.

Both long term trends and seasonal variations are not removed before estimating the mean $\overline{T}$, standard deviation s and skewness b. Therefore from Figure 2 is not possible to distinguish the effect of these two components. Temperatures at the tropics are higher and less variable than temperatures at the extratropics. East Europe, large part of Asia and most of the tropics have a positively skewed distribution (i.e. asymmetric distribution with longer tail on the right-hand side, indicating occurrence of large (hot) temperatures in these regions). Most of North America and the Atlantic sector have a negatively skewed distribution (i.e. asymmetric distribution with longer tail on the left-hand side, indicating occurrence of small (cold) temperatures in these regions).

Figure 3 shows the maximum value of summer monthly mean temperatures. The maximum temperature spatial pattern resembles the mean temperature pattern (Figure 2a)

with higher temperatures in tropical regions and lower temperatures in extratropical regions. The maximum value, however, is based on only one value and so is not representative of the distribution of hot extreme events. Besides, the maximum value is highly non-resistant to outliers. A more robust approach for investigating extremes is to produce a subsample of the original temperature data that contains excesses above a pre-defined threshold and then use these excesses to estimate parameters of the distribution of excesses. This approach allows the description of the entire distribution of excesses and also better estimates.

The starting point for such analysis is to define and estimate the threshold to be used to obtain the excesses. A procedure for defining and estimating the threshold is illustrated below. Figure 4a 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, where y is the year index and m is the month index. The horizontal dotted line is the 1870-2005 annual mean temperature of 6.8°C. 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 solid line is the quantity $M_{y,m}=L_{y,m}+ S_m$, where S_m is the mean annual cycle given by

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

where n is the number of years with available data (e.g. $n=136$ if data is available for the 1870-2005 period). The quantity $M_{y,m}$ assumes that the mean annual cycle S_m is constant for all years. This quantity is a seasonally varying fit for the time series of observed values and can be used to estimate a time-varying threshold to study hot extreme events. The time-varying threshold $u_{y,m}$ can for example be defined as $u_{y,m} = M_{y,m} + \varepsilon$, where ε is the increment necessary to have $\alpha\%$ of the observed values above $u_{y,m}$. 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 75th 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 back 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 5, 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.

When a constant threshold u is chosen (e.g. the 75th quantile as illustrated in Fig. 1) 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 4b shows the mean of 75% threshold $u_{y,m}$ for the summer months during the period from 1870 to 2005, which has a similar pattern to Figs. 2a and 3.

Figure 6 shows the mean of excesses $E(T_{y,m} - u_{y,m} \mid T_{y,m} > u_{y,m})$, the median of excesses $Med(T_{y,m} - u_{y,m} \mid T_{y,m} > u_{y,m})$ and the variance of excesses $Var(T_{y,m} - u_{y,m} \mid T_{y,m} > u_{y,m})$ above the 75% time-varying threshold $u_{y,m}$. Extratropical regions are colder and have

larger temperature variability than tropical regions (Figs. 2a and 2b). Therefore hot extreme temperatures produce larger mean excesses in extratropical regions than in tropical regions. A marked contrast with higher values of mean, median and variance of excesses over extratropical land areas and lower values over the oceans and tropical regions is observed. This contrast indicates that hot extremes are more intense 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. In tropical regions the ocean-land surface temperature contrast is much smaller, partially reflecting in smaller mean, median and variance of excesses than in extratropical latitudes.

b) Summary of extremes using extreme value theory

Two approaches based on extreme value theory (Coles, 2001) can be used to investigate extreme temperatures. Both approaches rely on asymptotic assumptions so that results are more accurate for rarer events. One approach uses the sample of block (e.g. annual) maxima values to investigate the distribution of hot extremes. In the asymptotic limit when the block of data is sufficiently large, the distribution of block maxima data approximates the GEV distribution. The other approach to study hot temperature extremes is to investigate the distribution of values above a sufficiently high threshold (e.g. 95th quantile). This second approach is known as peaks-over-threshold. For sufficiently large thresholds, the distribution of excesses $Z = T_{y,m} - u_{y,m}$ above a large threshold $u_{y,m}$ (i.e. conditional on $T_{y,m} > u_{y,m}$) approximates the Generalized Pareto (GP) distribution function

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

which is defined for $z > 0$ and $1 + \xi z/\sigma > 0$, where $\sigma > 0$ is the scale parameter and ξ is the shape parameter of the distribution. For the examples presented in this study σ and ξ have been obtained using maximum-likelihood estimation (Coles, 2001, section 2.6.3).

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 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)). For the monthly mean temperature application in this study the peaks-over-threshold approach is used.

Figures 7a shows GP scale parameter σ estimates at each grid point of summer monthly mean temperature excesses $T_{y,m}-u_{y,m}$ above the time-varying 75% threshold. The scale parameter σ provides information about the variability of the excesses. Regions with large values of scale parameter σ have higher variability of hot extreme temperatures than regions with small values of scale parameter. In accordance with the variance of

excesses shown in Fig. 6c, higher variability of extreme hot temperatures is found over extratropical continental areas when compared to oceanic and tropical regions. This accordance is observed because the variance of excesses is directly proportional to the squared value of σ, i.e. $Var(Z)= \sigma^2/(1-\xi)^2(1-2\xi)$. Note also that the scale parameter σ 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 and the median of excesses are proportional to σ, i.e. $E(Z)= \sigma/(1-\xi)$ and $Med(Z)= (\sigma/\xi)(2^{\xi}-1)$.

Figure 7b shows GP shape parameter ξ estimates at each grid point of summer monthly mean temperature excesses $T_{y,m} - u_{y,m}$ above the 75% time-varying threshold $u_{y,m}$. The shape parameter ξ tells us about the form 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. Shape parameter values above or equal to zero indicates 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 $-\sigma/\xi$, according to the GP distribution fit. 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 coloured 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 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 to the data. The goodness of fit can be examined using the Kolmogorov-Smirnov test (Massey, 1951), which for this particular application is a test of the null hypothesis that the true distribution function of temperature excesses is the same as the GP distribution with scale parameter σ and shape parameter ξ. Although our use of this test overestimates the goodness of fit – because the same sample used to estimate σ and ξ is used to perform the test – it can be used as a preliminary indicator of quality of fit. Other tests, for example, the Chi-square test or the Mann-Whitney test could be used to check the robustness of the Kolmogorov-Smirnov test results. Additionally, it is also advisable to examine if the chosen threshold is large enough so that the asymptotic assumption is respected. We have chosen a time-varying threshold such that 75% of the summer values fall below. However, one could question if the 75% choice is high enough to satisfy the asymptotic assumption. Figure 8 shows the percentage of grid points with Kolmogorov-Smirnov test probability value (p-value) less than or equal to p (a probability value between 0 and 1) for different choices of time-varying thresholds such that 75%, 80%, 85%, 90% and 95% of the summer values fall below the threshold, respectively. If the true distribution function is GP, by chance it is expected that the percentage of grid points with p-value less than or equal to p should match p exactly with all points falling on the diagonal line. Figure 8 shows that all curves are on the right-hand side of the diagonal line, indicating that the quality of the fit is good for all thresholds. Besides, the 75%, 80%, 85% and 90% threshold lines match well each other, suggesting that the asymptotic assumption is respected and that for all these thresholds the excesses are fitted well by the GP distribution.

Figure 9a shows the excesses above the 75% time-varying threshold $u_{y,m}$ during August 2003 (i.e. $T_{y,m} - u_{y,m}$ when $T_{y,m} > u_{y,m}$), the hottest ever recorded monthly mean temperature in Europe (Fig.1 and Fig 4a). Excesses up to 3°C are observed in central Europe. Figure 9b shows return period estimates $1/(1-H(z))$ for the August 2003 excesses of Fig. 9a using the GP distribution (3) with scale and shape parameter estimates of Figs. 7a and 7b, respectively. 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, suggesting that hot summers such as the August 2003 event are not as rare as suggested by Schär *et al*. (2004) that used the normal (Gaussian) distribution and obtained a return period of 46000 years over Switzerland. For example, the return period for the grid point in central Europe (12.5°E, 47.5°N) is 133 years. 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.

The relationship between extremes and factors (e.g. time and ENSO) can 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 examining how the distribution of summer temperature excesses is changing over the years and also how this distribution is related to ENSO, the following models can be used for the GP distribution fit of excesses $T_{y,m} - u_{y,m}$:

$$\log \sigma = \sigma_o + \sigma_1 t + \sigma_2 x \qquad (4)$$

$$\xi = \xi_o + \xi_1 t + \xi_2 x \qquad (5)$$

where t is a time index expressed for example in century units and x is the Southern Oscillation Index (SOI), which is an ENSO index. Note that in (4) the logarithm of σ is used instead of σ to ensure that $\sigma = exp(\sigma_o + \sigma_1 t + \sigma_2 x)$ is positive for all choices of parameter values σ_o, σ_1 and σ_2. The six GP distribution parameters on the right-hand sides of (4) and (5) can be obtained using maximum-likelihood estimation. Adding, for example, quadratic terms to (4) and (5) would capture non-linear effects.

Because of the logarithmic link function in (4) the parameters 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 t (i.e. a century) scales σ by $exp(\sigma_1)$.

Figure 10 shows maps of $exp(\sigma_1)$ (panel a) and $exp(\sigma_2)$ (panel b) estimated using summer temperature excesses $T_{y,m} - u_{y,m}$ above the 75% time-varying threshold during the period 1882-2005, which is the period of availability of the SOI. Both maps are extremely noisy without a clear large-scale pattern. Fig. 10a shows that the scale parameter σ is decreasing in central Europe (i.e. $exp(\sigma_1) < 1$), suggesting a decrease in the variability of hot extreme monthly temperature, and is increasing in nearly all the rest of Europe (i.e. $exp(\sigma_1) > 1$), suggesting an increase in the variability of hot extreme monthly temperature. Figure 10b shows that transitions from negative SOI (i.e. El Niño

conditions) and to positive SOI (i.e. La Niña conditions) amplify the variability of hot extreme temperatures over most North America and continental Europe where $exp(\sigma_2) > 1$. Transitions from positive SOI (i.e. La Niña conditions) to negative SOI (i.e. El Niño conditions) damp the variability of hot extreme temperatures over most North America and continental Europe. However, a significance test is required for a more conclusive analysis of these maps.

Figure 11 shows maps of ξ_1 (panel a) and ξ_2 (panel b). As for Fig. 10 both maps are extremely noisy without a clear large-scale pattern. Figure 11a shows that ξ_1 is positive over central Europe and the North Atlantic, indicating that over the centuries the shape parameter ξ is increasing. The tail of the distribution of monthly mean temperature excesses in these regions is becoming fatter (heavier), suggesting a tendency towards more extreme events in these regions. Figure 11b shows that ξ_2 is negative in central North America and central Europe, suggesting that in these regions when the SOI is positive the tail of the distribution of monthly mean temperature excesses is thinner (i.e. the shape parameter ξ is smaller during La Niña conditions) than when the SOI is negative (i.e. the shape parameter ξ is larger during El Niño conditions). In other words, in regions with negative ξ_2 more hot extreme temperature events are observed during El Niño years than during La Niña years. Note that Fig. 11b and Fig. 10b are mirror images of each other. Regions of $exp(\sigma_2) > 1$ in Fig. 10b coincide with regions of $\xi_2 < 0$ in Fig. 11b, suggesting that over these regions during the transition from negative SOI (i.e. El Niño conditions) to positive SOI (i.e. La Niña conditions) the variability of extremes events is amplified and the frequency of extreme events is reduced. During the transition from positive SOI (i.e. La Niña conditions) to negative SOI (i.e. El Niño conditions) the variability of extremes events is damped and the frequency of extreme events is increased. A significance test is also required here for a more conclusive analysis of these maps.

c) Clustering of extremes

The annual frequency of extreme events, e.g. the number of extreme events observed during each summer, is a measure of clustering of extremes. The average number of summer exceedances that occur in years for which there is at least one exceedance provides a measure of the average cluster size. By examining maps of this average number it is possible to identify regions where extreme events are more clustered in time (i.e. regions where there is more serial dependence).

Figure 12 shows the 1870-2005 average number of summer exceedances obtained using the time-varying 75% threshold $u_{y,m}$. A clear contrast between continental and oceanic regions is noted. Hot 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.

d) Spatial dependency of extreme events

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

$$\chi = \lim_{u \to u_+} Pr\{ T_O > u \,|\, T_E > u \} \qquad (6)$$

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 temperature at central Europe is high. If $\chi = 0$ then T_E and T_O are 'asymptotically independent'. If F_{T_O} and F_{T_E} are the marginal distribution functions of T_E and T_O, respectively, (6) can be re-written as

$$\chi = \lim_{u \to 1} Pr\{ F_{T_O}(T_O) > u \,|\, F_{T_E}(T_E) > u \}. \qquad (7)$$

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

$$\chi(u) = 2 - \frac{log\, Pr\{ F_{T_E}(T_E) < u, F_{T_O}(T_O) < u \}}{log\, Pr\{ F_{T_E}(T_E) < u \}} \qquad (8)$$

defined for $0 < u < 1$ (Coles *et al.*, 1999). Therefore, making the uniform transformations $rank(T_E)/m$ and $rank(T_O)/m$, where m is the number of observations T_E and T_O, to obtain $F_{T_E}(T_E)$ and $F_{T_O}(T_O)$ (i.e., by computing the empirical distribution functions of T_E and T_O) one can compute $\chi(u)$. For large thresholds $u \to 1$ the measure χ, which ranges from 0 to 1, provides a simple measure of extremal dependence between T_E and T_O. Larger values of χ indicate stronger dependence.

Figure 13a 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 on the west North Atlantic (67.5°W, 42.5°N). The vertical and horizontal lines are the 75th quantile of August monthly mean temperatures in the grid point in central Europe and in the west North Atlantic, respectively. The scatter plot shows that temperatures at the two grid points are positively related. Indications of extreme dependence are noticeable. Large values (above the 75th quantile) are observed simultaneously in the two grid points. The extreme dependence measure χ (6) at the 75th quantile level is obtained using the points of Fig. 13a that are located on the right-hand side of vertical 75th quantile line of August temperatures in central Europe. The χ statistics is given by the ratio between the total number of points to the right-hand side of the vertical line and the number of points on the top right-hand corner of the scatter plot (i.e. those points that are located above both the 75th quantile line of temperatures in the west North Atlantic and the 75th quantile line of temperatures in central Europe). 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 Figure 13b with $u = 0.75$ (vertical and horizontal

lines). In practice χ is computed using (8). Figures 13c and 13d show similar scatter plots of August monthly mean temperatures as in Figs 13a and 13b 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). No sign of extreme dependence is noticeable between these two grid points.

Figure 14a shows a map of χ (8) for August monthly mean temperatures for the grid point in central Europe (12.5°E, 47.5°N) with *u = 0.75*. This map allows the identification of extreme teleconnection patterns. As one should expect, grid points close to the central Europe grid point (12.5°E, 47.5°N) have large values of χ, indicating strong dependence. When hot temperatures are observed in the central Europe hot temperatures are also observed in neighbour areas. The west North Atlantic also shows some dependence, as previously illustrated in Figs. 13a and 13b. This dependence is likely to be linked to sea surface temperature conditions in the North Atlantic, but further investigation is required for a better understanding of the mechanisms behind such teleconnection. North America and west Russian show very weak dependence, as also previously illustrated in Figs. 13c and 13d.

The χ statistics provides a measure of extreme dependence for asymptotically dependent distributions. However, it fails to provide information of discrimination for asymptotically independent distributions (Coles, 2001). An alternative measure is therefore required to overcome this deficiency. Such measure is given by

$$\overline{\chi} = \lim_{u \to u_+} \overline{\chi}(u) \qquad (9)$$

where

$$\overline{\chi}(u) = \frac{2 \log \Pr\{F_{T_E}(T_E) > u\}}{\log \Pr\{F_{T_E}(T_E) > u, F_{T_O}(T_{T_O}) > u\}} - 1 \qquad (10)$$

defined for $0<u<1$ (Coles *et al.*, 1999). The $\overline{\chi}$ statistics range from –1 to 1. For asymptotically dependent variables $\overline{\chi}=1$. For independent variables $\overline{\chi}=0$. As χ provides a summary measure of the strength of dependence for asymptotically dependent variables, $\overline{\chi}$ provides a corresponding measure for asymptotically independent variables. In other words, when $\chi = 0$ (or close to zero) then $\overline{\chi}$ is a more appropriate measure of the strength of extremal dependence. As the correlation coefficient is the standard measure of association between two variables, $\overline{\chi}$ is the equivalent association measure for extreme events.

Figure 14b shows a map of $\overline{\chi}$ for August monthly mean temperatures for the grid point in central Europe (12.5°E, 47.5°N) with *u = 0.75*. As noticed in Fig. 14a, hot extreme temperatures in grid point in central Europe are strongly associated with hot extreme temperatures in neighbour grid points. Central Europe hot extreme temperatures are also confirmed to be associated with hot extreme temperatures in the west North Atlantic (Fig. 13a and 13b). Figure 14b still shows a negative association between hot temperatures in central Europe and west Russia, which is also noticeable in Fig. 13c and 13d.

4. Conclusions

This study has presented a number of tools for the analysis of extreme climate and weather events in gridded datasets. These tools allow exploratory analysis of spatial patterns of extremes, the investigation dependence of extremes on factors such as time and ENSO, the analysis of clusters of extremes and also spatial extreme dependence analysis (teleconnection patterns of extremes).

The methods and tools have been demonstrated using Northern Hemisphere monthly mean gridded temperature data from CRU covering the period from 1870 to 2005. Motivated by the recent hot summer of 2003 the demonstration of the methods has been performed using summer (June-July-August) monthly mean temperature. Extreme events have been defined as those months when the mean temperature has fallen above a 75% time-varying threshold.

These methods have allowed us to answer the questions raised in the introductory section:

- Hot extreme temperatures have larger variability in extratropical continental regions than in oceanic and tropical regions
- Hot extreme temperatures show a tendency towards more events with smaller variability in central Europe
- Transitions from negative to positive SOI amplify the variability and reduce the frequency of hot extreme events in central Europe
- Hot extreme temperatures are more clustered over the Atlantic and East Pacific oceans than over continental regions
- Hot extreme temperatures over central Europe during August are related to hot extreme temperatures in the west North Atlantic

The methods presented here could be further developed and improved by, for example, using quantile regression (Koenker, 2005) to define the threshold for obtaining the excesses instead of estimating a threshold based on the mean variability of the observed time series lifted up until the desirable percentage of points is above the threshold. 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 neighbour 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 intiative (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 EU-funded ENSEMBLES project. Functions have been written in the R statistical language (http://www.r-project.org) to allow performance of climate analysis in gridded datasets. 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. The availability of these functions will hopefully stimulate the climate science community to further investigate extreme weather and climate events in gridded datasets.

Acknowledgements

We wish to thank CRU for making available the gridded temperature dataset used in this research. CASC was supported by ENSEMBLES (GOCE-CT-2003-505539).

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*, **Vol 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).

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

Coles, S. G., 2001: An Introduction to statistical modeling of extreme values. Springer series in statistics. 208pp.

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.

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, 1998: 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.

Massey, F. J. Jr., 1951: The Kolmogorov-Smirnov test of goodness of fit. *Journal of the American Statistical Association*, **46**, 68-78.

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.

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.

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

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.

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

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.

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

Figures

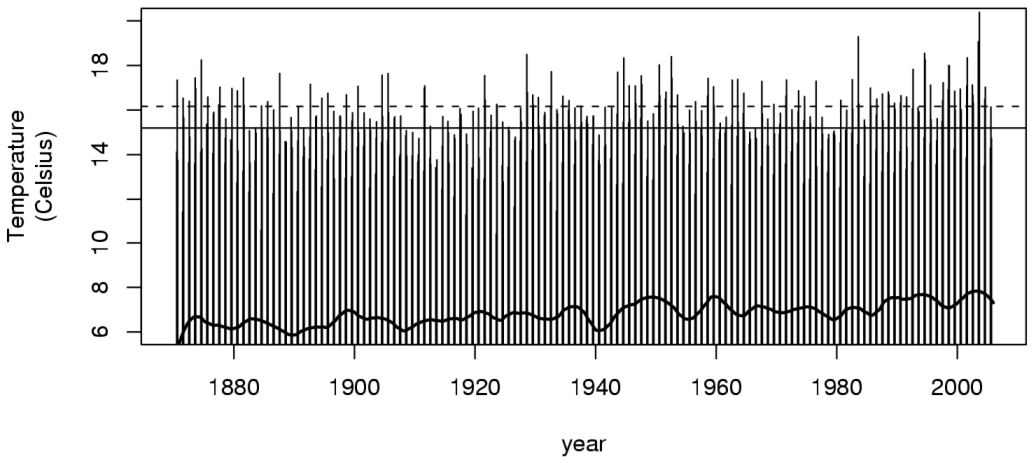


Figure 1: Summer (June-July-August) monthly mean temperatures *T* (Celsius) from 1870-2005 at a grid point in central Europe (12.5^{o}E, 47.5^{o}N). Each vertical bar is the monthly mean for a particular summer month. The horizontal solid line is the long term (1870-2005) summer monthly mean temperature of 15.2^{o}C. The horizontal dashed line is the 75th quantile of summer monthly mean temperatures of 16.2^{o}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.

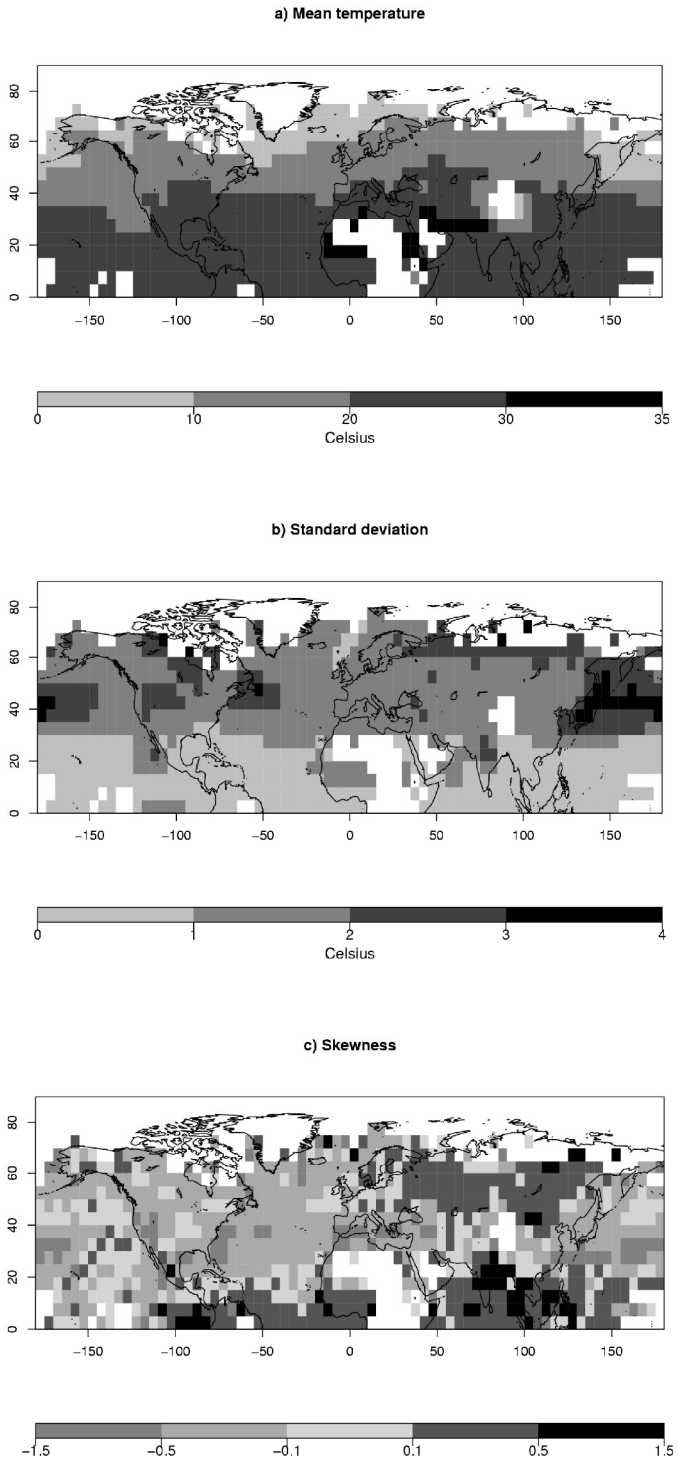


Figure 2: a) Mean ($\overline{T}$), b) standard deviation (s) and c) skewness $b = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{T_i - \overline{T}}{s}\right)^3$, of monthly mean summer temperatures estimated over the 1870-2005 period (n=408).

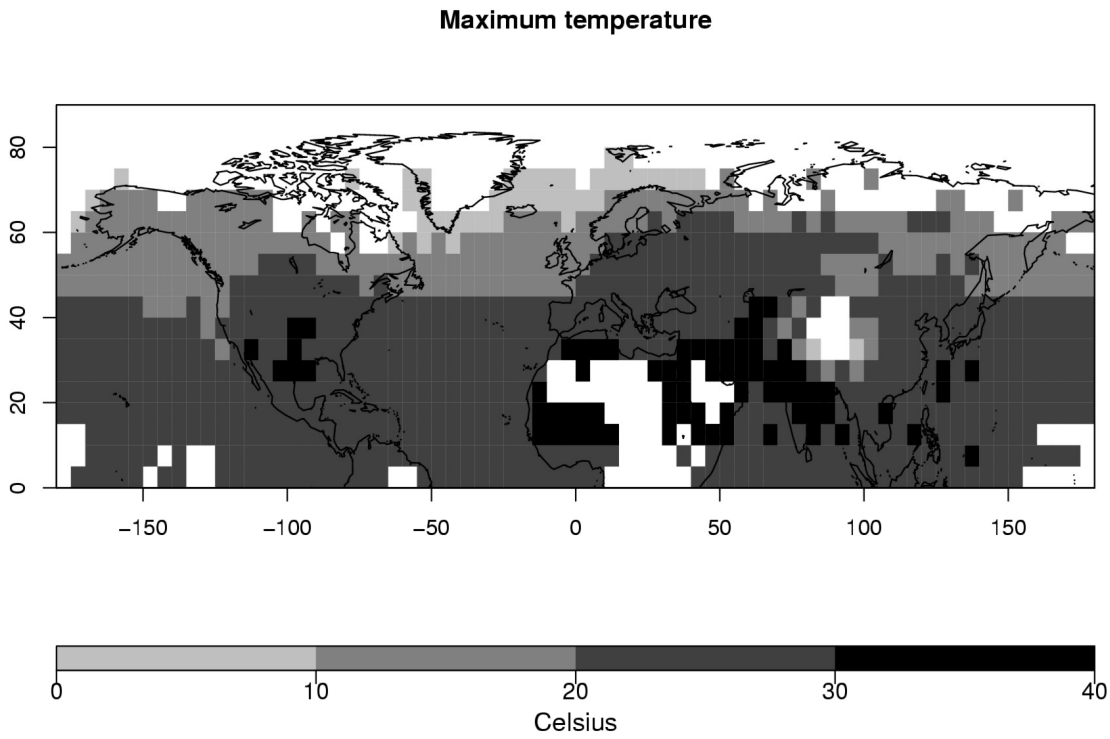


Figure 3: Maximum of monthly mean summer temperatures estimated over the 1870-2005 period.

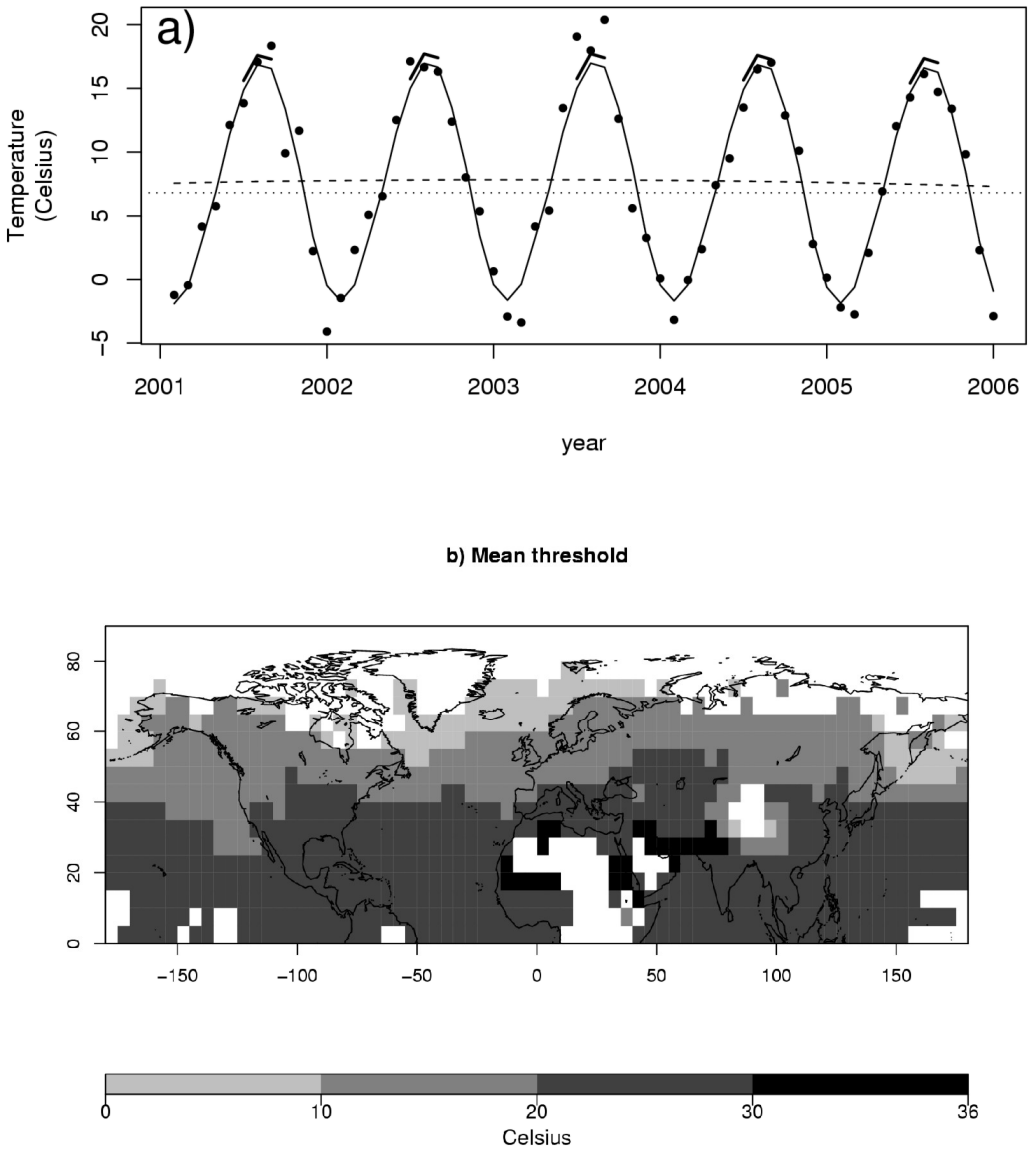


Figure 4: a) Observed monthly mean temperatures $T_{y,m}$ (black dots) for the grid point in central Europe (12.5°E, 47.5°N) during the period from 2001 to 2005. The horizontal dotted line is the 1870-2005 annual mean temperature of 6.8°C. The dashed line is the long term trend ($L_{y,m}$) that represents 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 ε is the increment necessary to have α% of the observed values above $u_{y,m}$. The thick solid segments is the 75% threshold $u_{y,m}$ for the summer months (α=25%). b) Mean of 75% threshold $u_{y,m}$ for the summer months during the period from 1870 to 2005.

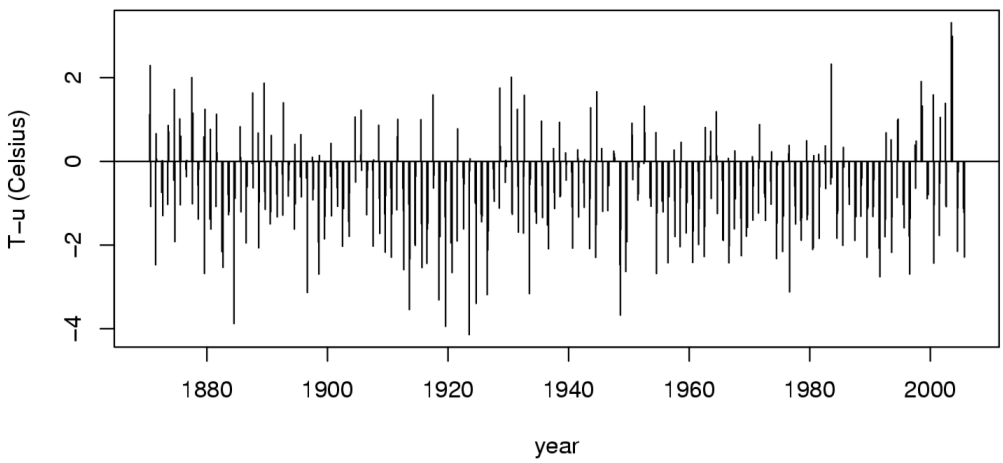


Figure 5: Summer month differences between $T_{y,m}$ and $u_{y,m}$ (75% time-varying threshold) for the grid point in central Europe (12.5°E, 47.5°N) during the period from 1870 to 2005. Vertical bars above the horizontal solid line (zero line) are excesses $T_{y,m} - u_{y,m}$ above the threshold $u_{y,m}$.

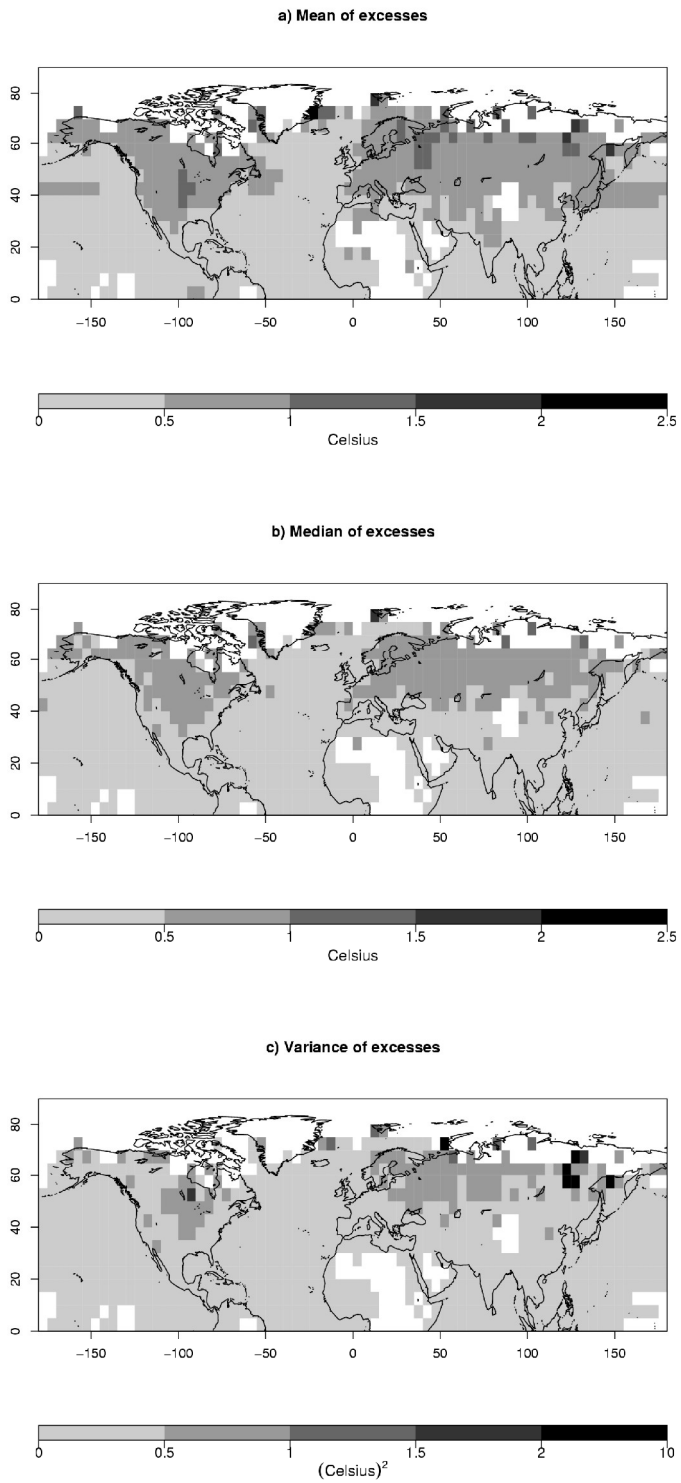


Figure 6: a) Mean of excesses $E(T_{y,m} - u_{y,m} \mid T_{y,m} > u_{y,m})$, b) the median of excesses $Med(T_{y,m} - u_{y,m} \mid T_{y,m} > u_{y,m})$ and c) variance of excesses $Var(T_{y,m} - u_{y,m} \mid T_{y,m} > u_{y,m})$ above the 75% time-varying threshold $u_{y,m}$.

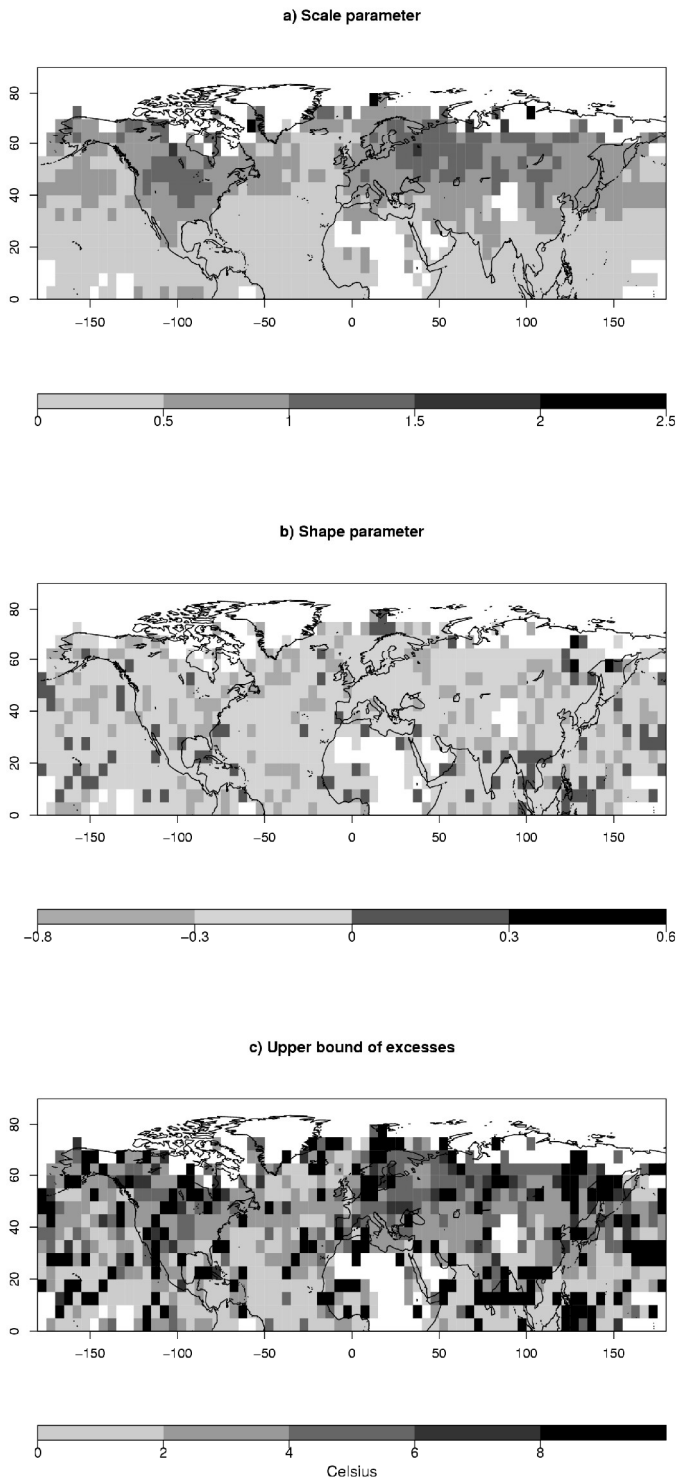


Figure 7: Summer monthly mean temperature GP a) scale parameter σ, b) shape parameter ξ, and c) upper bound $-\sigma/\xi$ of excess $T_{y,m} - u_{y,m}$ above the 75% time-varying threshold $u_{y,m}$. Grid points with $\xi \geq 0$ have infinite upper bound and coloured in black.

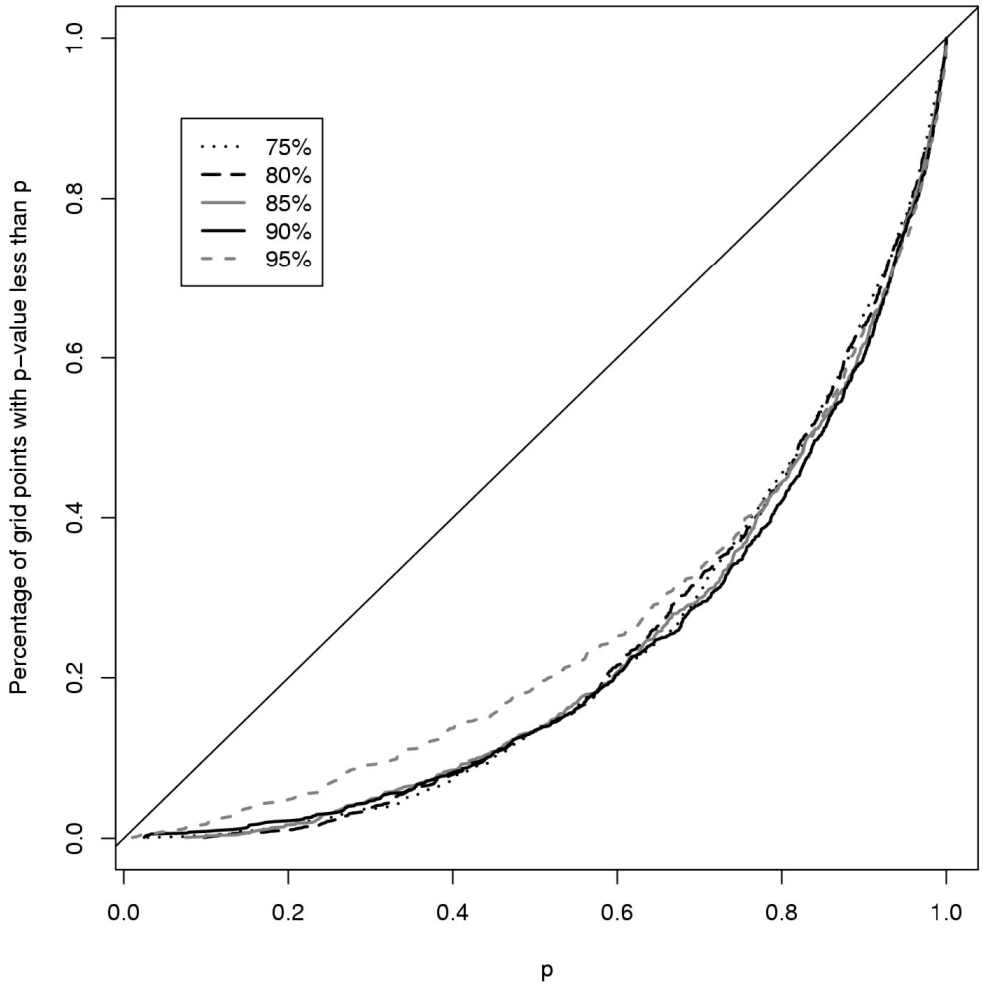
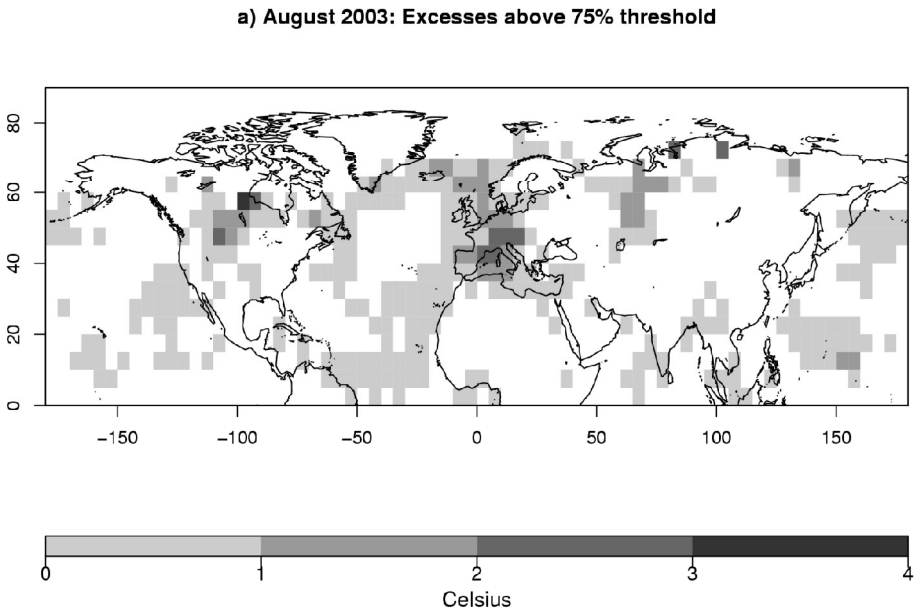


Figure 8: Percentage of grid points with Kolmogorov-Smirnov test probability value (p-value) less than or equal to p (a probability value between 0 and 1) for different choices of time-varying thresholds such that 75%, 80%, 85%, 90% and 95% of the summer values fall below the threshold.



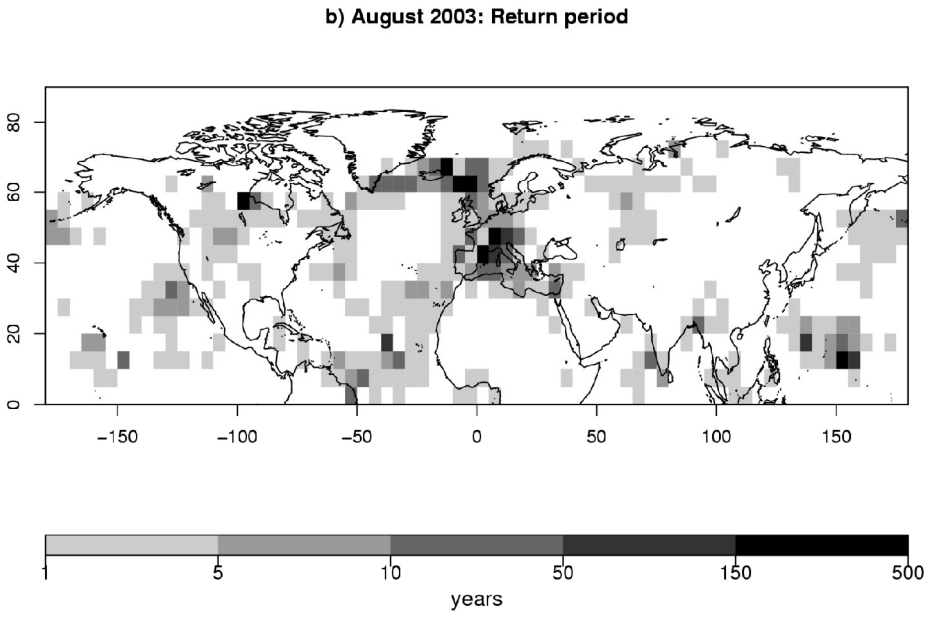


Figure 9: a) Excesses $T_{y,m} - u_{y,m}$ above the 75% threshold $u_{y,m}$ during August 2003 (i.e. $T_{y,m} - u_{y,m}$ when $T_{y,m} > u_{y,m}$). b) Return period estimates *1/(1–H(z))* for the August 2003 excesses of a) using the GP distribution with scale and shape parameters estimates of Figs. 7a and 7b, respectively.

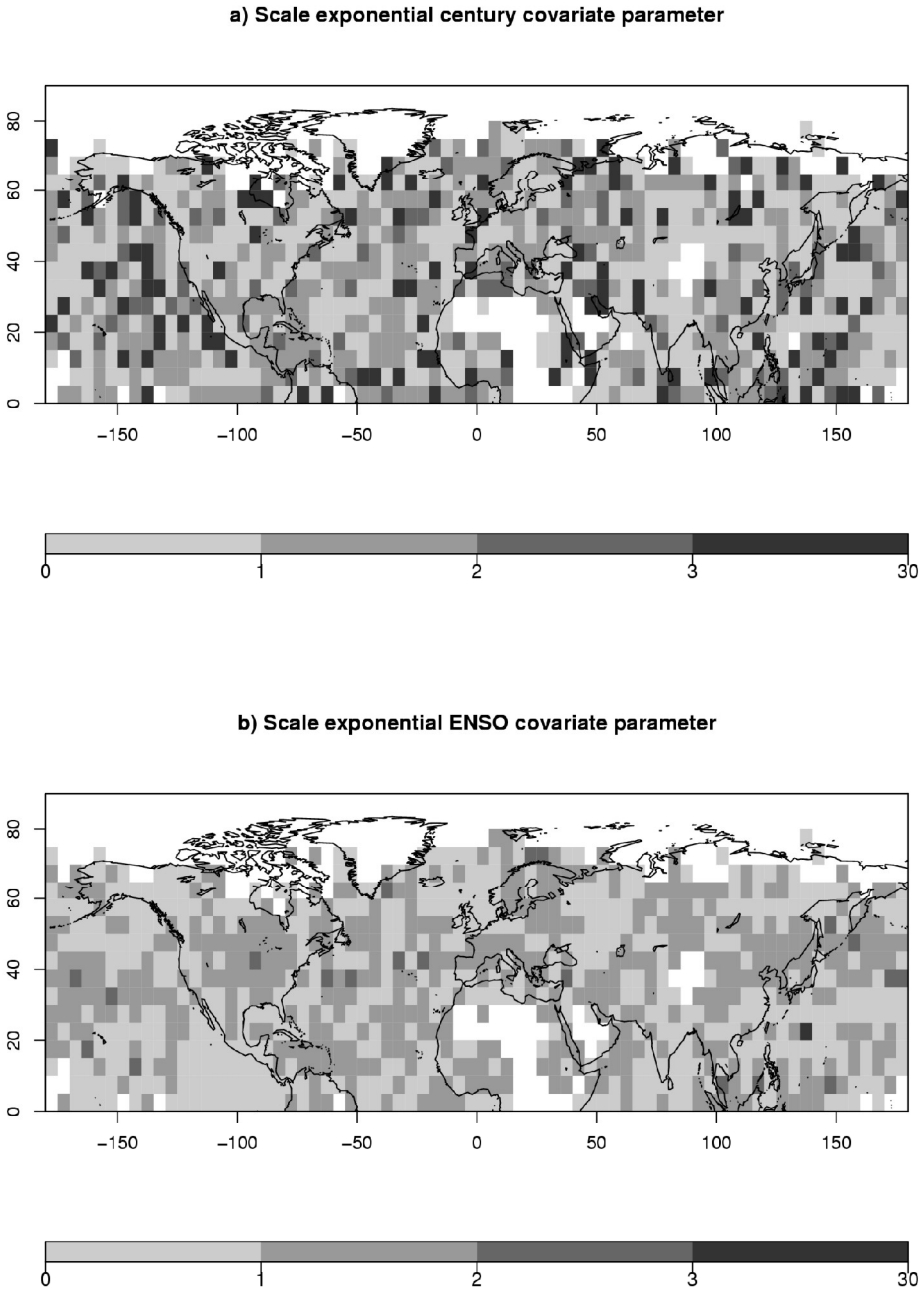


Figure 10: a) $exp(\sigma_1)$ and b) $exp(\sigma_2)$ estimated using summer temperature excesses $T_{y,m} - u_{y,m}$ above the 75% time-varying threshold during the period 1882-2005.

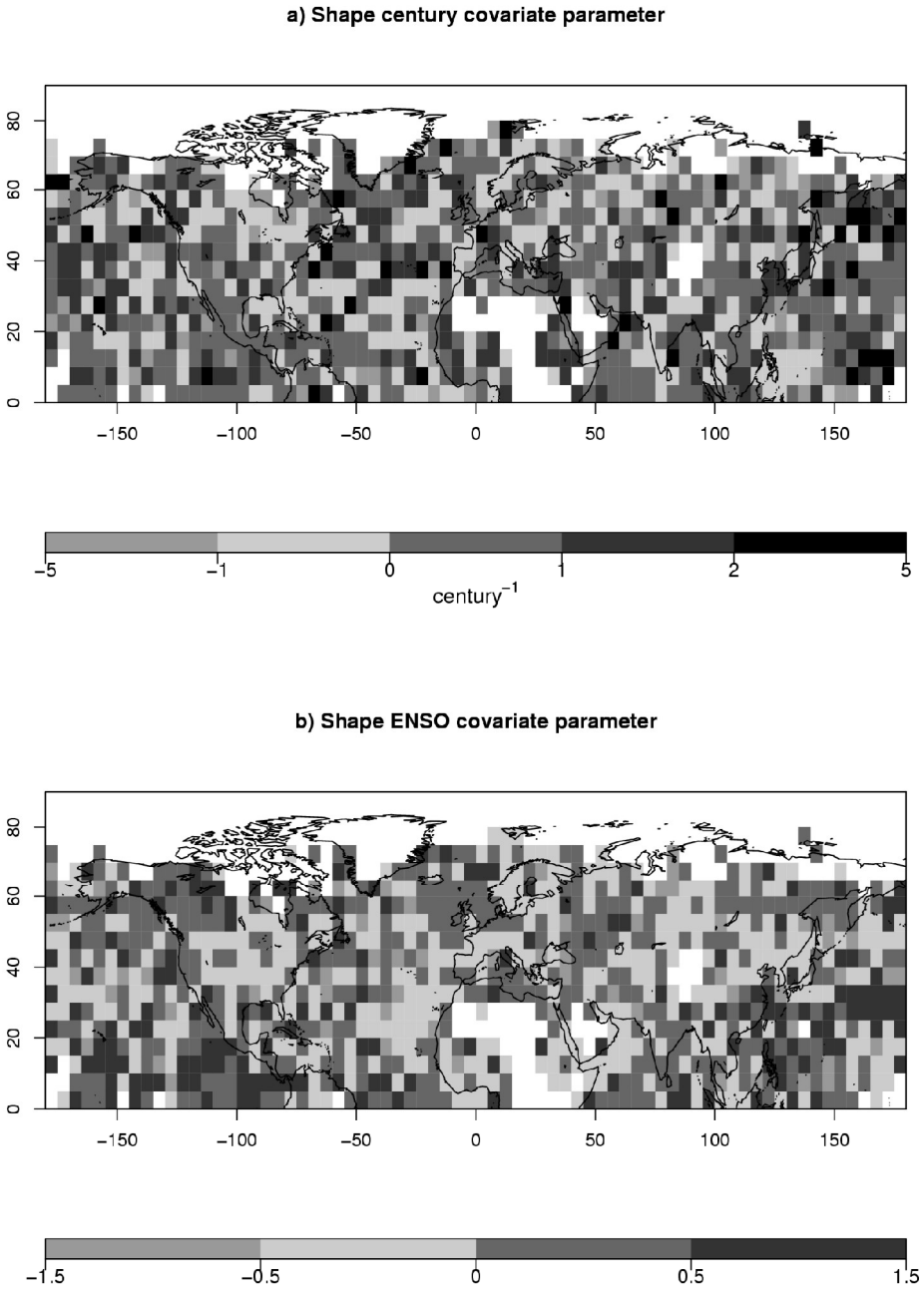


Figure 11: a) ξ_1 and b) ξ_2 estimated using summer temperature excesses $T_{y,m} - u_{y,m}$ above the 75% time-varying threshold during the period 1882-2005.

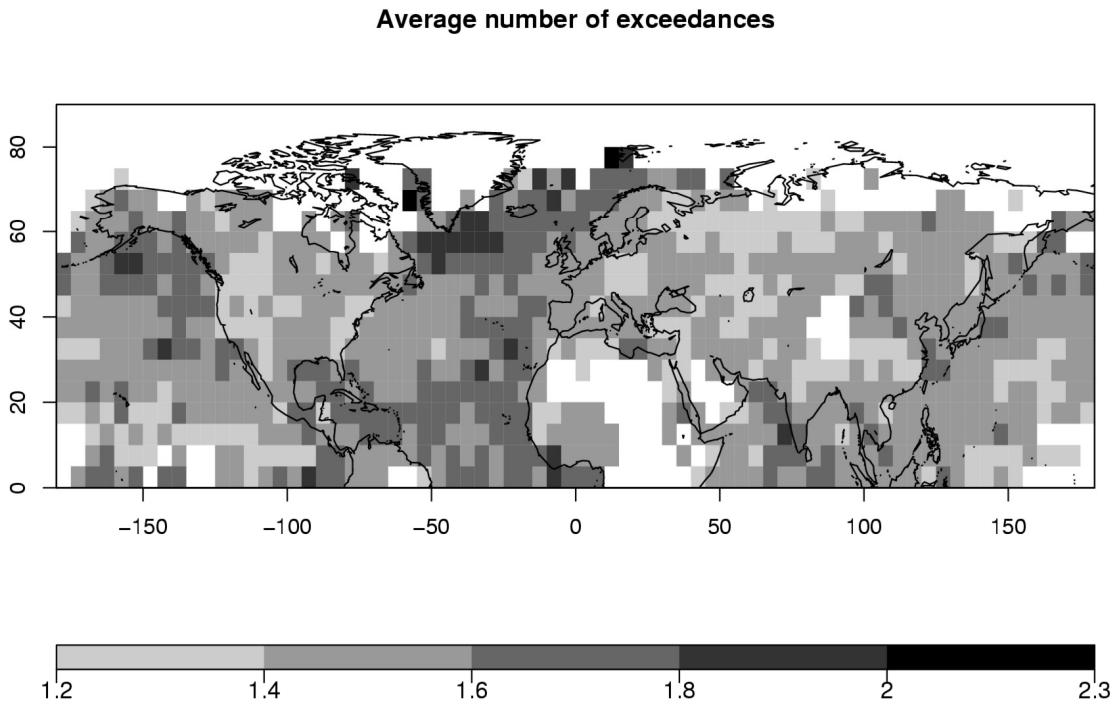


Figure 12: 1870-2005 Average number of summer exceedances obtained using the time-varying 75% threshold $u_{y,m}$.

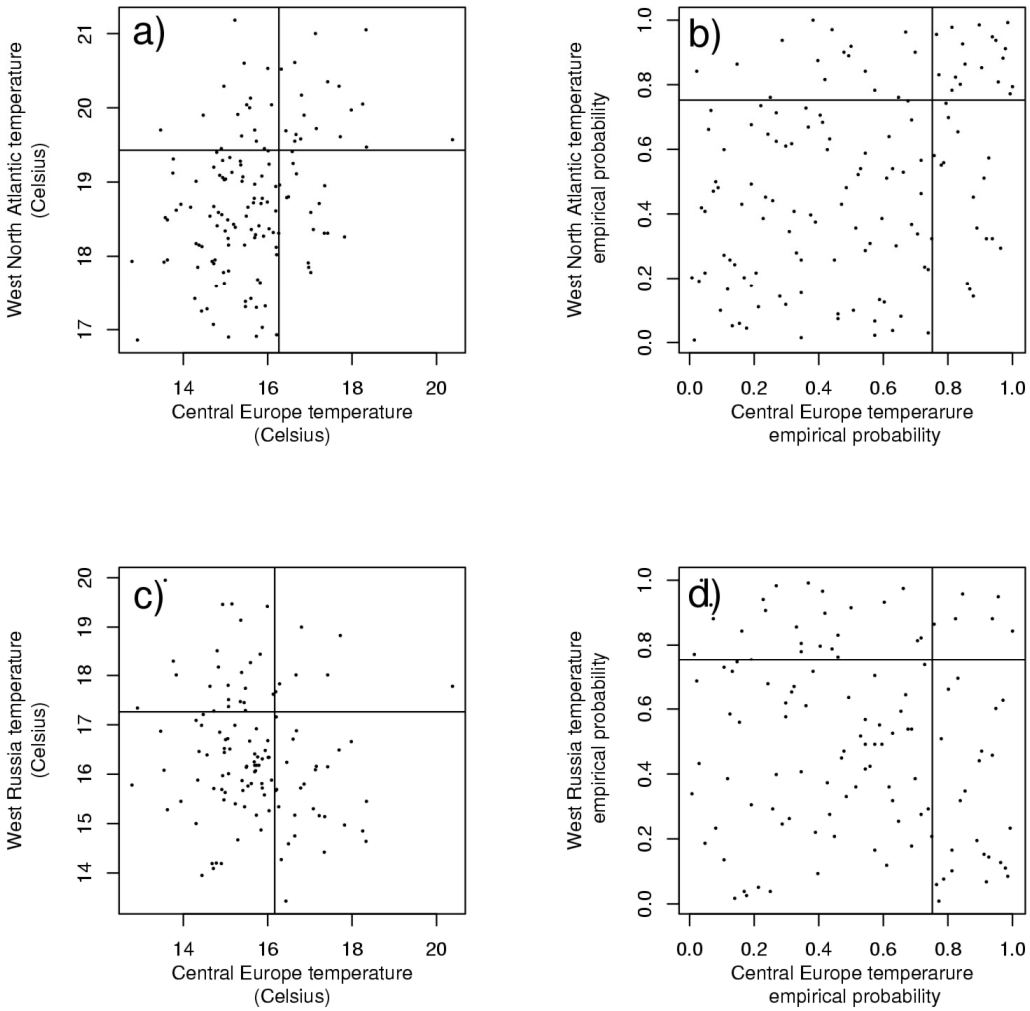
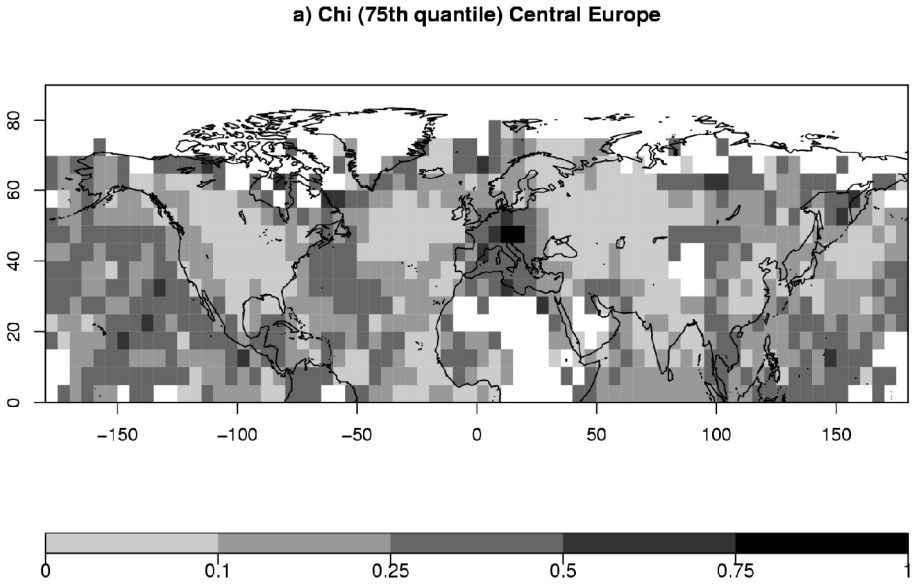


Figure 13: a) Scatter plot of August monthly mean temperatures T_E in a grid point in central Europe (12.5°E, 47.5°N) and August monthly mean temperatures T_O in a grid point on the west North Atlantic (67.5°W, 42.5°N). The vertical and horizontal lines are the 75th quantile of August monthly mean temperatures in the grid point in central Europe and in the west North Atlantic, 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).



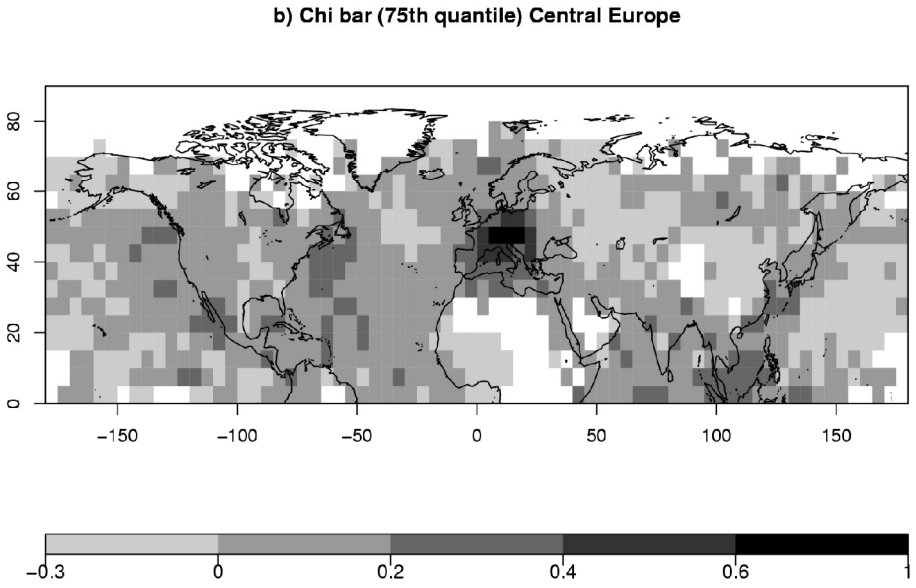


Figure 14: a) χ and b) $\overline{\chi}$ for August monthly mean temperatures for the grid point in central Europe (12.5°E, 47.5°N) with $u = 0.75$.