ORIGINAL PAPER

Climate change and modelling of extreme temperatures in Switzerland

Boriss Siliverstovs · Rainald Ötsch · Claudia Kemfert · Carlo C. Jaeger · Armin Haas · Hans Kremers

Abstract This study models maximum temperatures in Switzerland monitored in twelve locations using the generalised extreme value (GEV) distribution. The parameters of the GEV distribution are determined within a Bayesian framework. We find that the parameters of the underlying distribution underwent a substantial change in the beginning of the 1980s. This change is characterised by an increase both in the level and the variability. We assess the likelihood of the heat wave of the summer 2003 using the fitted GEV distribution by accounting for the presence of a structural break. The estimation results do suggest that the heat wave of 2003 is not that statistically improbable if an appropriate methodology is used for dealing with nonstationarity.

Keywords Climate change · GEV · Bayesian modelling · Great Alpine heat wave

B. Siliverstovs (✉)
ETH Zürich, KOF Swiss Economic Institute, Weinbergstrasse 35, 8092, Zurich, Switzerland
e-mail: boriss.siliverstovs@kof.ethz.ch

R. Ötsch
University of Potsdam, Potsdam, Germany

C. Kemfert · H. Kremers
DIW Berlin, Berlin, Germany

C. C. Jaeger · A. Haas
PIK, Potsdam, Germany

1 Introduction

The heat wave of 2003 observed in continental Europe, including Switzerland, has attracted much attention in the literature on climate change as the unusually high temperatures led to a number of undesirable consequences including increased population mortality (World Health Organization (WHO) 2003)—especially among the elderly—with estimated 1,000 heat-related fatalities in Switzerland (Grize et al. 2005), appearance and prolonged endurance of droughts accompanied by a shortfall in crop, an increased probability and severity of forest fires, change of vegetation cycles, and strongly reduced discharge in many rivers (De Bono et al. 2004; Fink et al. 2004; ProClim 2005). In Switzerland, the heat wave of 2003 sped up the melting of glaciers in the Alps and it resulted in avalanches and flash floods.

Ongoing research is concerned with the mechanisms contributing to the formation of extreme events like the heat wave in 2003. Della-Marta et al. (2007a, b) provide an overview about the recent literature and name relevant causes, including the Atlantic Multidecadal Oscillation (AMO), the long-term global mean temperature, deficiencies in Mediterranean winter and spring precipitation and changes in land–atmosphere interactions triggered by the greenhouse gas effect. Schär and Fischer (2008) point out the coupling of heat waves to the water cycle but also mention large uncertainties related to regional changes. García-Herreira et al. (2009) present a review on the literature of the European heat wave 2003, including a thorough discussion of causes and impacts and mentioning the factors attributable to the occurence and persistence of this event, namely blocking episodes, soil moisture deficit, and sea surface temperature.

Several papers are dedicated to the probability of this extraordinary hot summer both at the European level and

more specifically for Switzerland, see Beniston (2004), Schär et al. (2004), Trigo et al. (2005), Stott et al. (2004), among others. A common conclusion of these articles is that the heat wave of summer 2003 was a very unusual event given the pattern of temperatures observed over Europe in the past. In particular, Schär et al. (2004) concluded that such record-breaking extreme temperatures observed in Switzerland were very unlikely from a statistical point of stationarity and a shift in the distribution location alone is not sufficient for explaining the heat wave of 2003.

In this paper, we model the annual maxima of monthly mean temperatures in Switzerland with the main purpose of assessing the likelihood of occurrence of the heat wave of 2003. There is no unique definition of a heat wave but it is generally understood as a prolonged period of unusually high temperatures observed in a given region. Short-termed definitions are useful when looking at increased mortality by hot temperatures (Tamerius et al. 2007). These definitions assume thresholds that have to be exceeded on subsequent days (see Robinson 2001 for a discussion, Khaliq et al. 2007 for an application for Canada). Other effects like melting of glaciers require more time to emerge and therefore provide another definition of a heat wave. Schär et al. (2004), Stott et al. (2004), and Jaeger et al. (2008) analyse mean summer temperatures, Beniston (2004) examines daily summer maximum temperatures averaged over the summer months, whereas Trigo et al. (2005) uses 15 day averages. Since impacts of climate change are more visible by observing longer time spans, we choose in line with these publications an event of longer endurance.

We apply a Bayesian approach which is better suited for predictive purposes than the classical methodology since parameter uncertainty is directly incorporated into the forecast process, see Coles (2001). Furthermore, investigation of model parameter instability and assessment of its severity is also straightforward within the Bayesian framework. Building on the Bayesian analysis of Jaeger et al. (2008), who assessed the feasibility of different trend models under the assumption of a normally distributed error term, we employ the generalised extreme value (GEV) distribution as possibly more appropriate.

Our main finding is that a proper accounting for features of the time series considerably increases the likelihood of occurrence of the heat wave of 2003. On the basis of our estimation results, we conclude that the heat wave of summer 2003 does not appear to be such an improbable event but it rather constitutes a future pattern of things to come.

In Sect. 2 we describe the data set used in our exercise. In Sect. 3 the methodology is presented. Section 4 contains the estimation results. The last section concludes.

2 Data

We analyse the temperature measured by the Swiss Federal Office of Meteorology and Climatology[1] (Begert et al. 2005). The mean monthly data are provided for the period from 1864 until present (2007) (with exception of Chateau d'Oex (since 1901) and Davos-Dorf (since 1876)) and are collected at the following twelve locations in Switzerland (Bern-Zollikofen, Geneve-Cointrin, Lugano, Segl-Maria, Basel-Binningen, Chateau d'Oex, Chaumont, Davos-Dorf, Engelberg, Saentis, Sion, and Zurich). The series are homogenous until 2003 (inclusive) for the first four locations, while for the remaining locations inhomogeneities have been provisionally corrected. Since 2004 several stations have been reconstructed. Those time series might contain minor inhomogeneities, where the reconstruction site has not changed (Davos-Dorf, Engelberg, Lugano, Saentis, Zurich) or was moved a bit (Sion). The temperature record is likely to be less homogeneous for the Bern station due to the fact that this station has been rebuild at a completely different place.[2] Thus, the most reliable data is provided by Geneve-Cointrin and Segl-Maria, where the measurement location did not undergo major changes.

In our exercise, in contrast to Schär et al. (2004), where the data from four different stations (Basel-Binningen, Geneve-Cointrin, Bern-Zollikofen, and Zurich) were amalgamated, we model each time series of temperature observations individually. In doing so, we avoid a possible aggregation bias. By using disaggregated data, we are also able to draw a comparison between the parameter estimates obtained for each station and, therefore, to establish a degree of generalisation of our results depending on the measurement location.

The descriptive statistics of the corresponding time series are given in Table 1. First, observe that the time series are quite heterogenous. Our sample includes Saentis, where the average annual maximum temperature is around 5.6 degrees Celsius, on the one hand, and Lugano with 21.5, on the other hand. Second, the summer of 2003 was indeed the hottest summer by the historical standard for all locations where the measurement took place. So far, this record was subsequently broken in 2006 for the three stations Bern-Zollikofen, Davos-Dorf, and Sion.

Since our analysis is based on the assumption of independent observations, we check for the magnitude of autocorrelation in our data. The first autocorrelation coefficient, estimated using the whole sample available, is

---

Table 1 Descriptive statistics for the whole sample

Station	Sample	Obs.	Mean	Std. dev.	Min	Max	1st AC[a]	2003
Basel-Binningen	1864–2007	144	18.801	1.4495	16.0	23.8	0.128	23.8
Bern-Zollikofen	1864–2007	144	17.917	1.3919	15.2	22.1	0.156	21.9
Geneve-Cointrin	1864–2007	144	19.586	1.4516	16.8	24.1	0.230	24.1
Zurich	1864–2007	144	18.109	1.3983	15.4	22.7	0.136	22.7
Chateau d'Oex	1901–2007	107	15.327	1.3988	12.8	19.5	0.291	19.5
Chaumont	1864–2007	144	14.597	1.5495	11.1	19.5	0.103	19.5
Davos-Dorf	1876–2007	132	11.694	1.3039	8.8	16.1	0.280	15.7
Engelberg	1864–2007	144	14.586	1.2476	12.1	18.7	0.183	18.7
Lugano	1864–2007	144	21.476	1.1682	18.8	25.1	0.305	25.1
Saentis	1864–2007	144	5.647	1.4258	2.9	10.1	0.247	10.1
Segl-Maria	1864–2007	144	11.015	1.0726	8.6	14.7	0.220	14.7
Sion	1864–2007	144	19.131	1.3549	16.4	23.3	0.271	23.1

Temperature in °C

[a] The first autocorrelation coefficient

displayed in Table 1. It takes values in the range between 0.103 and 0.305, indicating presence of low to mild positive temporal dependence in our data. However, according to Perron (1989) the detected mild positive dependence may well be spuriously induced by the presence of an unmodelled structural break in the time series in question. Therefore, after conducting the initial analysis using the whole sample period, we investigate the structural stability of the fitted model by splitting the sample into two parts. We also calculate the first order autocorrelation for each of these subsamples. We find that the evidence of temporal dependence is substantially weakened when one allows for a structural break in the temperature time record.

3 Methodology

In this section we motivate our choice of the GEV distribution for modelling annual maximum temperatures in Switzerland and explain how its parameters can be estimated using the Bayesian approach. Second, we introduce a changepoint model that allows the model parameters to differ across the sub-samples and explain how its parameters, including an additional breakpoint timing parameter, can be easily estimated. Third, we discuss how the statistical inference on the parameters of the fitted models can be performed using the information contained in the parameter posterior distributions. We also discuss how the likelihood of future extreme events can be addressed in the Bayesian framework. In particular, we are interested in comparing the probability of observing the extreme temperature event in summer 2003 as predicted by the model with constant parameters with that predicted by the changepoint model.

Fourth, we discuss how standard tools of goodness-of-fit evaluation can be adapted to our modelling approach.

Essentially, we have chosen to model the maximum temperatures in Switzerland using the GEV distribution due to two main reasons.[3] First, as argued in Leadbetter et al. (1983), the distribution of the maximum of identically distributed random variables is asymptotically approximated by the GEV distribution. Second and more relevant, the GEV distribution is very flexible in the sense that it incorporates a wide range of tail behaviours. The latter feature allows us to trace the changes in the tail behaviour of the fitted GEV distribution across the different sub-samples as implied by our changepoint model.

The GEV distribution is characterized by the following distribution function

$$F(z; \mu, \sigma, \xi) = \exp\left\{ -\left(1 + \xi \frac{(z-\mu)}{\sigma} \right)_+^{-\frac{1}{\xi}} \right\}, \qquad (1)$$

where $y_+ = \max(y, 0)$ and μ is a location parameter, σ is a scale parameter, and ξ is a shape parameter. The shape parameter ξ determines the tail behaviour which can be sub-divided into three classes: the limit $\xi \to 0$ corresponds

---

[3] This approach is rather standard in extreme value statistics. More than one observation during a period can be considered by using r-largest-techniques (Coles 2001). Changepoint problems using the Poisson process in a Bayesian framework date back to Raftery and Akman (1986). Renard et al. (2006) quote Peak-over-Threshold-methods as a second standard approach alternative to the analysis of block maxima, either with a Poisson process or the Generalized Pareto-distribution. These also allow using more data information than just one extreme during one period. Other information could be considered by kriging (spatial information) or a variety of multivariate methods. Seneviratne et al. (2006) used Canonical Correlation Analysis (CCA), based on daily maximum temperature to investigate the occurrence of heat waves.

to the Gumbel distribution, $\xi > 0$ and $\xi < 0$ to the Fréchet and Weibull distributions, respectively. The Fréchet distribution is a "long-tailed" distribution, the Gumbel distribution is a "medium-tailed" distribution, whereas the Weibull distribution is a "short-tailed" distribution which has a finite endpoint.

Collect the three parameters of interest into the following parameter vector $\theta = (\mu, \sigma, \xi)'$. Then under the assumption of independent and identically distributed realizations of a GEV(θ) random variable $\mathbf{x} = (x_1, \ldots, x_n)$, the associated log-likelihood function is

$$\ln L(\theta|\mathbf{x}) = -n \ln \sigma - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{n} \ln \left\{1 + \frac{\xi(x_i - \mu)}{\sigma}\right\}$$
$$- \sum_{i=1}^{n} \ln \left\{1 + \frac{\xi(x_i - \mu)}{\sigma}\right\}^{-\frac{1}{\xi}}. \quad (2)$$

In Bayesian statistics all relevant information on the unknown parameters θ given the observed data $\mathbf{x}$ is contained in the posterior distribution $\pi(\theta|\mathbf{x})$, which is derived from the parameter prior distribution $\pi(\theta)$ and from the likelihood function $L(\theta|\mathbf{x})$ according to the Bayes theorem

$$\pi(\theta|\mathbf{x}) = \frac{\pi(\theta)L(\theta|\mathbf{x})}{\int_\Theta \pi(\theta)L(\theta|\mathbf{x})d\theta} \propto \pi(\theta)L(\theta|\mathbf{x}). \quad (3)$$

The prior distribution $\pi(\theta)$ assumed for the parameters of the model allows to incorporate exogenous information (e.g., experts' opinion) into the analysis. Formula (3) converts the initial beliefs about the values of the unknown parameters, expressed by $\pi(\theta)$, into a posterior distribution $\pi(\theta|\mathbf{x})$, that includes the additional information contained in the data. Hence if one uses vague prior distributions, reflecting the fact that no prior strong beliefs about the parameter values are introduced, then the posterior distributions of the parameters are primarily determined by the data at hand. The term in the denominator $\int_\Theta \pi(\theta)L(\theta|\mathbf{x})d\theta$ is the normalizing constant such that the posterior distribution of the model parameters is proportional to the product of the prior distribution and of the likelihood function.

Since evaluation of the normalizing constant in (3) is not always analytically tractable, the posterior density is usually obtained by means of computation-intensive techniques collectively known as Markov Chain Monte Carlo (MCMC) simulation. In our paper we also follow this approach. More specifically, we employ a simulation technique known as the Metropolis-Hastings (MH) algorithm in order to obtain the posterior distribution of the parameters of interest. According to the MH algorithm we generate a sequence of parameter vectors $\theta_1, \theta_2, \ldots$ according to some probability rule $q(\theta_{k+1}|\theta_k)$, called a proposal generating density. One popular choice of a proposal

generating density states that the proposed parameter vector θ^* for θ_{k+1} is generated from a standard normal distribution with mean equal to the previous vector θ_k, e.g., $(\theta^*|\theta_k) \sim N(\theta_k, 1)$. Then, at each step in the sequence, we accept the proposal value θ^* for θ_{k+1} with a probability α_k using the following rule

$$\alpha_k = \min\left\{1, \frac{\pi(\theta^*|\mathbf{x})}{\pi(\theta_k|\mathbf{x})}\right\}. \quad (4)$$

Otherwise, we leave θ_k unchanged. As a result, the next value in the sequence is determined as follows

$$\theta_{k+1} = \begin{cases} \theta^* & \text{with probability } \alpha_k, \\ \theta_k & \text{with probability } 1 - \alpha_k. \end{cases} \quad (5)$$

The simulated sequence is approximately stationary, and its marginal distribution corresponds to the posterior distribution given by (3) (Gamerman 1997).

So far, we have discussed the estimation of the model parameters θ of the GEV distribution assuming that they are constant. In order to investigate the possibility of the presence of a structural break and its severity, we extend our model by allowing for a breakpoint that endogenously splits the whole sample into two sub-samples $I_0 = [1, \kappa-1]$ and $I_1 = [\kappa, n]$.[4] We allow the parameters of the fitted GEV distribution to differ across the two sub-samples,[5] i.e., $\theta_0 = (\mu_0, \sigma_0, \xi_0)'$ and $\theta_1 = (\mu_1, \sigma_1, \xi_1)'$, and the timing of the breakpoint κ appears as a new parameter in the model. Such a modification of the model is easily implemented in the Bayesian framework as opposed to the maximum likelihood approach which becomes rather cumbersome when dealing with variable changepoints. We impose an uniform prior on our breakpoint timing parameter and allow for a minimum length of the sub-sample equal to six observations. The breakpoint timing parameter is easily included into our MH algorithm. Following Coles and Pericchi (2003) we utilise a simple discrete random walk chain as the proposal generating density for the breakpoint parameter. The (log-)likelihood function needs also to be correspondingly modified for a given value of the breakpoint timing parameter κ as follows

---

[4] An interesting extension of our work could be pursued along the lines of Perreault et al. (2000) suggesting a multivariate change-point analysis in order to establish a joint changepoint for the measurement series.

[5] The modeling of a structural change using a step-function in the temporal evolution of parameters is of course not mandatory. Smooth functions could be used instead as long as they can be implemented into the likelihood function, as well as plenty approaches for modeling one or several breakpoints exist (Carlin et al. 1992; Stephens 1994; Mohammad-Djafari and Feron 2006; Moreno et al. 2005, among others).

$$\ln L(\theta_0, \theta_1, \kappa | \mathbf{x})$$

$$= -(\kappa - 1)\ln \sigma_0 - \left(1 + \frac{1}{\xi_0}\right)\sum_{i=1}^{\kappa-1} \ln\left\{1 + \frac{\xi_0(x_i - \mu_0)}{\sigma_0}\right\}$$

$$- \sum_{i=1}^{\kappa-1} \ln\left\{1 + \frac{\xi_0(x_i - \mu_0)}{\sigma_0}\right\}^{-\frac{1}{\xi_0}} - (n - \kappa + 1)\ln \sigma_1$$

$$- \left(1 + \frac{1}{\xi_1}\right)\sum_{i=\kappa}^{n} \ln\left\{1 + \frac{\xi_1(x_i - \mu_1)}{\sigma_1}\right\}$$

$$- \sum_{i=\kappa}^{n} \ln\left\{1 + \frac{\xi_1(x_i - \mu_1)}{\sigma_1}\right\}^{-\frac{1}{\xi_1}}. \tag{6}$$

The posterior distribution of $\widetilde{\theta} = (\theta_0', \theta_1', \kappa)'$ can be easily obtained by means of an appropriately modified MH algorithm outlined above.

Since the posterior distribution $\pi(\theta | \mathbf{x})$ contains all relevant information on the unknown parameters θ given the observed data $\mathbf{x}$, we can use it in order to deduce point estimates of the unknown parameters of interest, for example. These are given by the respective means of the posterior distribution. Likewise, assessment of the probability of future events reaching a certain extreme level could also be straightforwardly investigated. For our purpose, this is a particularly appealing property since we intend to show that observing the record-breaking temperature in summer 2003, denoted as z_{2003}, does not appear that excessively improbable once the features of the underlying time series are taken into account. To this end, we intend to compare the predictive posterior probability for z_{2003} implied by a model with constant parameters with that implied by our changepoint model.

Denote $\pi(z_{2003}|\theta)$ as the density function of z_{2003} under conditional independence assumption for x_i given a GEV(θ) distribution with $\theta = (\mu, \sigma, \xi) \in \Theta$. Then we can compute the posterior predictive density of z_{2003}, given annual maxima $\mathbf{x}$ observed in the past

$$\pi(z_{2003}|\mathbf{x}) = \int_\Theta \pi(z_{2003}|\theta)\pi(\theta|\mathbf{x})\mathrm{d}\theta. \tag{7}$$

Next, define the predictive distribution of a future annual maximum as follows

$$Pr\{Z \leq z_{2003}|\mathbf{x}\} = \int_\Theta Pr\{Z \leq z_{2003}|\theta\}\pi(\theta|\mathbf{x})\mathrm{d}\theta, \tag{8}$$

where Z is a GEV(θ) distributed random variable and $Pr\{Z \leq z_{2003}|\theta\}$ is the distribution function derived from (1) evaluated at z_{2003}. Then the probability of observing a more extreme observation than z_{2003}, given the past data,—the so called exceedance probability—is given by the converse probability of (8). A similar analysis can be performed

using the changepoint model with the parameter vector $\widetilde{\theta}$. Very low exceedance probabilities would either indicate that the heat wave of summer 2003 was indeed an unusual and largely unanticipated event, or a model which delivers such low probability is misspecified, and therefore it needs to be appropriately modified.

Information contained in the posterior distribution can also be used in order to estimate the distributions of quantiles or, equivalently, return levels. Define the return level z_p as the $(1-p)$th-quantile of the GEV distribution shown in (1) with the associated return period $1/p$, such that

$$z_p(\theta) = \begin{cases} \mu - \frac{\sigma}{\xi}[1 - (-\ln(1-p))^{-\xi}], & \text{for } \xi \neq 0, \\ \mu - \sigma\ln(-\ln(1-p)), & \text{for } \xi = 0. \end{cases} \tag{9}$$

We can use the simulated sequence of the parameters θ (or, $\widetilde{\theta}$), obtained by means of the MH algorithm, in order to get a posterior distribution of a chosen return level z_p. Denote the mean of the resulting posterior distribution of z_p as $\bar{z}_p$. Then the plot of $\bar{z}_p$ against a logarithmic scale for $-\ln(1-p) \approx 1/p$ for a range of small values p is called a return level plot. The property of the return plot is that it is linear for the Gumbel distribution, i.e., for $\xi = 0$, and concave and, respectively, convex for the Fréchet ($\xi > 0$) and Weibull ($\xi < 0$) distributions.

We can also use other summary statistics, such as 0.025 and 0.975 quantiles of the posterior return level distribution. Plotting them on the same graph against the return period for a range of small values p yields a 95% credibility interval for each return level z_p. Knowing this credibility interval allows us using the return level plots as a goodness-of-fit test by comparing empirical return levels deduced from the data with those implied by the values of the estimated parameters of the model. The model produces a good fit of the data if the empirical return levels do not fall outside of the computed 95% credibility interval. This also would indicate that the chosen distribution—in our case, a GEV distribution—is able to generate data similar to those observed.

Last but not least, we used probability and quantile plots for diagnostic model checking. Given the order statistics of a sample of annual maxima $z_{(1)} \leq z_{(2)} \leq \cdots \leq z_{(n)}$, the empirical distribution function evaluated at $z_{(i)}$ reads as follows

$$\widetilde{F}(z_{(i)}) = \frac{i}{n+1}. \tag{10}$$

The corresponding model-based distribution function is given by

$$\widehat{F}(z_{(i)}) = \exp\left\{-\left(1 + \hat{\xi}\frac{(z_{(i)} - \widehat{\mu})}{\widehat{\sigma}}\right)_+^{-\frac{1}{\xi}}\right\}, \tag{11}$$

where $(\widehat{\mu}, \widehat{\sigma}, \widehat{\xi})$ denote the point estimates of the parameters of interest, e.g., deduced from the respective posterior

distributions. Then a probability plot consists of the points $\widetilde{F}(z_{(i)})$ plotted against $\widehat{F}(z_{(i)})$. Similarly, plotting $\widehat{F}^{-1}(i/(n+1))$ against $z_{(i)}$ for $i = 1,...,n$ produces a quantile plot. In case of a good fit, the empirical and model-based quantities should not deviate far from each other, i.e., they should lie close to the 45-degree line on each plot, of course, accounting for data randomness and for parameter estimation uncertainty.

4 Results

4.1 A constant parameter model

First, we report the results of fitting the GEV distribution to the data available before 2003, assuming constant parameter values for the underlying GEV distribution. As discussed in Sect. 3, the posterior distributions of the model parameters were obtained using a MH random walk algorithm (Gamerman 1997) because an exact analytical solution is unavailable. We employed proper but diffuse priors that are also independent: the prior distribution is $N(0, 10,000)$ for the location parameter μ and the logarithmic transformation of the scale parameter σ, and $N(0,100)$ for the shape parameter ξ. We generated Markov chains of 100,000 observations with an initial burn-in period equal to 20,000 observations. Furthermore, in order to reduce autocorrelation in the generated Markov chains we have applied thinning by storing every 100th generated value. The convergence of the Markov chains was assessed using the convergence criterion suggested in Geweke (1992).

Table 2 presents estimation results derived from the posterior distribution based on all observations before 2003. It is worth noting that for all stations the shape parameter ξ of the fitted GEV distribution is very likely to be negative. This suggests that the tail behaviour implied by the estimated model with constant parameters is very likely to correspond to the Weibull distribution, i.e., $\xi < 0$.

As discussed in Sect. 3 above, we can assess how likely the hot summer of 2003 was, given historical data. We calculate the predictive probability of exceeding the temperature observed in 2003, see Table 2. The associated predictive probabilities are very low suggesting that the heatwave of 2003 was indeed a rather unusual event. Thus, at this point, our conclusions conform with those of Schär et al. (2004). The next column of Table 2 contains the corresponding predictive return periods. Such periods indicate that one should expect to observe such extreme temperature values on average once within a corresponding period (Coles 2001). For all measurement stations, the return periods by far exceed our sample period of 144 years (except for Chateau d'Oex), for which temperature recordings were made.

4.2 A changepoint model

A casual examination of the time series reveals that, in the last 25 years of our sample, the average annual maximum is much higher than observed for the period up to the early 1980s. Table 3 presents the descriptive statistics for two subsamples: before 1982 and afterwards. The characteristics of these two sub-samples are quite different. We

Table 2 GEV parameter estimates (posterior mean and standard deviation) and predictions on exceeding the temperature of 2003, sample 1864–2002

	$\hat{\mu}$	Std. dev.	$\hat{\sigma}$	Std. dev.	$\hat{\xi}$	Std. dev.	2003 (°C)	Predictive probability[c]	Predictive return period
Basel-Binningen	18.175	0.118	1.256	0.085	−0.156	0.060	23.8	0.0013	774.51
Bern-Zollikofen	17.333	0.123	1.269	0.087	−0.191	0.059	21.9	0.0033	299.62
Geneve-Cointrin	18.980	0.124	1.323	0.090	−0.200	0.063	24.1	0.0018	558.53
Zurich	17.503	0.115	1.232	0.083	−0.158	0.056	22.7	0.0017	572.27
Chateau d'Oex[a]	14.676	0.132	1.212	0.099	−0.119	0.071	19.5	0.0058	172.50
Chaumont	13.956	0.135	1.445	0.095	−0.216	0.052	19.5	0.0011	919.12
Davos-Dorf[b]	11.131	0.113	1.160	0.084	−0.172	0.061	15.7	0.0025	406.45
Engelberg	14.042	0.098	1.098	0.070	−0.144	0.050	18.7	0.0021	471.44
Lugano	20.958	0.103	1.053	0.070	−0.161	0.051	25.1	0.0028	359.94
Saentis	5.029	0.118	1.271	0.086	−0.156	0.054	10.1	0.0028	356.30
Segl-Maria	10.569	0.093	0.999	0.068	−0.214	0.056	14.7	0.0008	1,310.46
Sion	18.588	0.120	1.282	0.086	−0.254	0.055	23.1	0.0011	912.24

[a] The available sample starts in 1901

[b] The available sample starts in 1876

[c] The exceedance probability of the 2003 recording

Table 3 Descriptive statistics for two subsamples

	Subsample	Mean	Std. dev.	Min	Max	1st AC[a]	Subsample	Mean	Std. dev.	Min	Max	1st AC
Basel-Binningen	1864–1981	18.503	1.247	16.0	21.3	−0.008	1982–2007	20.154	1.532	18.2	23.8	−0.278
Bern-Zollikofen	1864–1981	17.676	1.276	15.2	20.6	0.101	1982–2007	19.012	1.370	17	22.1	−0.293
Geneve-Cointrin	1864–1981	19.298	1.318	16.8	22.6	0.133	1982–2007	20.892	1.305	18.8	24.1	−0.276
Zurich	1864–1981	17.870	1.245	15.4	20.6	0.078	1982–2007	19.192	1.539	17.2	22.7	−0.210
Chateau d'Oex	1901–1981	14.898	1.149	12.8	17.4	0.096	1982–2007	16.665	1.254	14.9	19.5	−0.261
Chaumont	1864–1981	14.347	1.442	11.1	17.9	0.041	1982–2007	15.731	1.513	13.9	19.5	−0.271
Davos-Dorf	1876–1981	11.359	1.067	8.8	13.8	0.105	1982–2007	13.058	1.295	11.1	16.1	−0.211
Engelberg	1864–1981	14.336	1.072	12.1	16.9	0.085	1982–2007	15.723	1.353	13.7	18.7	−0.229
Lugano	1864–1981	21.243	1.072	18.8	25	0.179	1982–2007	22.531	0.993	20.9	25.1	0.070
Saentis	1864–1981	5.335	1.204	2.9	8.7	0.143	1982–2007	7.062	1.497	4.8	10.1	−0.249
Segl-Maria	1864–1981	10.811	0.952	8.6	12.8	0.138	1982–2007	11.938	1.104	10.1	14.7	−0.146
Sion	1864–1981	18.836	1.211	16.4	21.8	0.126	1982–2007	20.469	1.150	18.9	23.3	−0.160

Temperature in °C

[a] The first autocorrelation coefficient

observe an increase in both the average values as well as in the extremes (minimum and maximum values) in the second sub-period compared to the first one. We also find that in the second sub-period the standard deviation is larger for all but three locations. This suggests that the whole distribution underwent a structural change such that extremely hot temperatures became more likely than they used to be. This also implies that the much discussed heat wave in 2003 may not have been a separate, undesirable incident, but rather the consequence of an increase in the level of temperatures as well as in their variability that happened in the early 1980s.

Given our earlier results indicating at most mild positive temporal dependence in the underlying time series, we also investigate how accounting for a structural break influences our conclusions drawn for the whole sample period. The estimated first autocorrelation coefficient for two subsamples is presented in Table 3. As expected, splitting the sample into two parts resulted in much lower values of the first autocorrelation coefficient reported for the first subsample. It takes values in the range between −0.008 and 0.179. At the same time, it is interesting to observe that the magnitude of the first order autocorrelation observed in the second subsample largely remained in the similar range reported for the whole sample but with the opposite sign taking values in the interval between −0.293 and 0.070. At the same time they appear to be somewhat larger than those observed for the first subsample. We attribute this difference to the fact that the estimate of autocorrelation for the second subsample is based on a rather small number of observations (27 years). This naturally increases the variability of the autocorrelation estimator and therefore makes it less reliable in comparison to that based on the much larger first subsample.

In order to investigate the possibility of the presence and severity of a structural break, we extend our model by allowing for a breakpoint that endogenously splits the whole sample into two sub-samples. As discussed in Sect. 3, we allow the parameters of the fitted GEV distribution to differ across two sub-samples and the timing of the breakpoint appears as a new parameter in the model.

Table 4 presents the parameter estimates resulting from the inclusion of the break point into our model. The posterior expected value of the location parameters and the observed temperature values appear in Fig. 1. There is substantial evidence in favour of the two regimes such that in the second part of our sample the average annual maximum temperature is 1 to 1.6 degrees higher—depending on the station—than it used to be in the first part of our sample.

The posterior distributions of the parameters in each regime are displayed in Fig. 2. Observe that the posterior distributions of the second sub-sample are more disperse than those of the first sub-sample due to the four times smaller sample size of the later sub-sample. Furthermore, Fig. 2 highlights the differences between the two different regimes. The posterior distributions of the location parameters μ_0 and μ_1 do not overlap at all, suggesting that the difference is statistically significant. At the same time, the posterior distributions of the scale parameters largely do overlap, suggesting that this parameter did not undergo much of a change. We however restrain from imposing the same value of the scale parameter in both regimes, following the advice of Coles and Pericchi (2003) who argue that when dealing with extreme events the parameter uncertainty always has to be accounted for.

An interesting contrast we find in our changepoint model is provided by the substantially different distributions of the

Table 4 Parameter estimates for the model with a structural break

	First subsample						Second subsample						Breakpoint timing		
	$\hat{\mu}_0$	Std. dev.	$\hat{\sigma}_0$	Std. dev.	$\hat{\xi}_0$	Std. dev.	$\hat{\mu}_1$	Std. dev.	$\hat{\sigma}_1$	Std. dev.	$\hat{\xi}_1$	Std. dev.	10th quantile	Median	90th quantile
Basel-Binningen	18.018	0.123	1.194	0.092	−0.193	0.074	19.353	0.280	1.125	0.240	0.234	0.216	1980	1982	1984
Bern-Zollikofen	17.179	0.131	1.233	0.098	−0.200	0.079	18.296	0.263	1.075	0.214	0.128	0.182	1980	1981	1984
Geneve-Cointrin	18.771	0.126	1.243	0.100	−0.162	0.075	20.277	0.283	1.169	0.220	0.023	0.167	1981	1982	1985
Zurich	17.389	0.126	1.208	0.097	−0.204	0.084	18.383	0.323	1.161	0.270	0.226	0.231	1980	1982	1989
Chateau d'Oex	14.414	0.138	1.077	0.107	−0.142	0.110	16.006	0.248	1.039	0.197	0.106	0.205	1980	1982	1983
Chaumont	13.804	0.148	1.422	0.103	−0.222	0.062	14.919	0.258	1.067	0.235	0.256	0.214	1979	1981	1983
Davos-Dorf	10.957	0.112	1.066	0.082	−0.236	0.071	12.426	0.260	1.117	0.212	0.059	0.186	1980	1982	1983
Engelberg	13.934	0.107	1.056	0.078	−0.227	0.072	15.051	0.260	1.132	0.267	0.076	0.186	1980	1982	1983
Lugano	20.805	0.104	1.004	0.074	−0.126	0.055	22.057	0.208	0.897	0.167	0.028	0.172	1980	1982	1986
Saentis	4.867	0.119	1.179	0.085	−0.196	0.060	6.352	0.336	1.357	0.277	−0.009	0.203	1980	1982	1987
Segl-Maria	10.469	0.099	0.965	0.072	−0.281	0.073	11.401	0.219	0.968	0.165	0.002	0.161	1980	1981	1983
Sion	18.396	0.126	1.202	0.085	−0.231	0.062	19.938	0.221	0.905	0.196	0.150	0.203	1982	1983	1988

shape parameter. For the first sub-sample, the value of the shape parameter ξ_0 is almost certainly negative. This implies that the tail behaviour of the GEV distribution fitted to the first sample is well approximated by a Weibull distribution, which conforms with the estimation results reported in Table 2 for the constant parameter model. On contrary, the posterior expectations of the shape parameter ξ_1 for the second sub-sample in all measurement locations but one exceeds zero and in three locations lies even above 0.2. This suggests that it is very likely that the tail behaviour in the second sub-sample has drastically changed and now it is rather consistent with the Fréchet distribution. However, a word of caution must be uttered as the dispersion in the posterior distribution of ξ_1 is up to three times as large as that of the shape parameter ξ_0, measured by the standard deviation, see Table 4. In addition, the probability mass on negative values of ξ_1 is still considerable, so the Weibull distribution still may deserve some attention. Further information on the difference in tail characteristics for both sub-samples can be found in Fig. 3, where the probability density function of the fitted GEV distribution for each of the two sub-samples are shown. As seen, in the second period the right-hand tail (see the lower panel for the corresponding measurement station) decays at much smaller rate than that plotted for the first subsample (depicted at the upper panel). We also observe the shift of the whole probability density function towards more extreme observations, which is consistent with Fig. 1.

Figure 4 displays the return level plots associated with every regime. Not only the whole return level curve has shifted upwards, but also the uncertainty has substantially increased for large temperatures. The return levels derived from the quantiles of the observations are in line with the credibility intervals as shown in Fig. 4, supporting the choice of the GEV distribution. Additional plot diagnostics via probability and quantile plots back up this view.[6] However, observe that the return level plot for the second sub-sample is based on less than thirty observations and therefore the parameters of the GEV distribution have been estimated with a rather large degree of uncertainty which is translated also into the return level plot.

Last but not least, Fig. 5 contains the posterior distribution of the breakpoint timing denoted as the first year of the second sub-sample. The corresponding median along with the 10th and 90th percentiles are reported in Table 4. It is either 1981 (three times), 1982 (eight times), or 1983 (once). This similarity is a remarkable finding for the different and quite heterogeneous measurement locations. This observation is further strengthened by the fact that the corresponding posterior distribution is very tight. This

---

[6] In order to save space, we do not report the corresponding probability and quantile plots but make them available upon request.
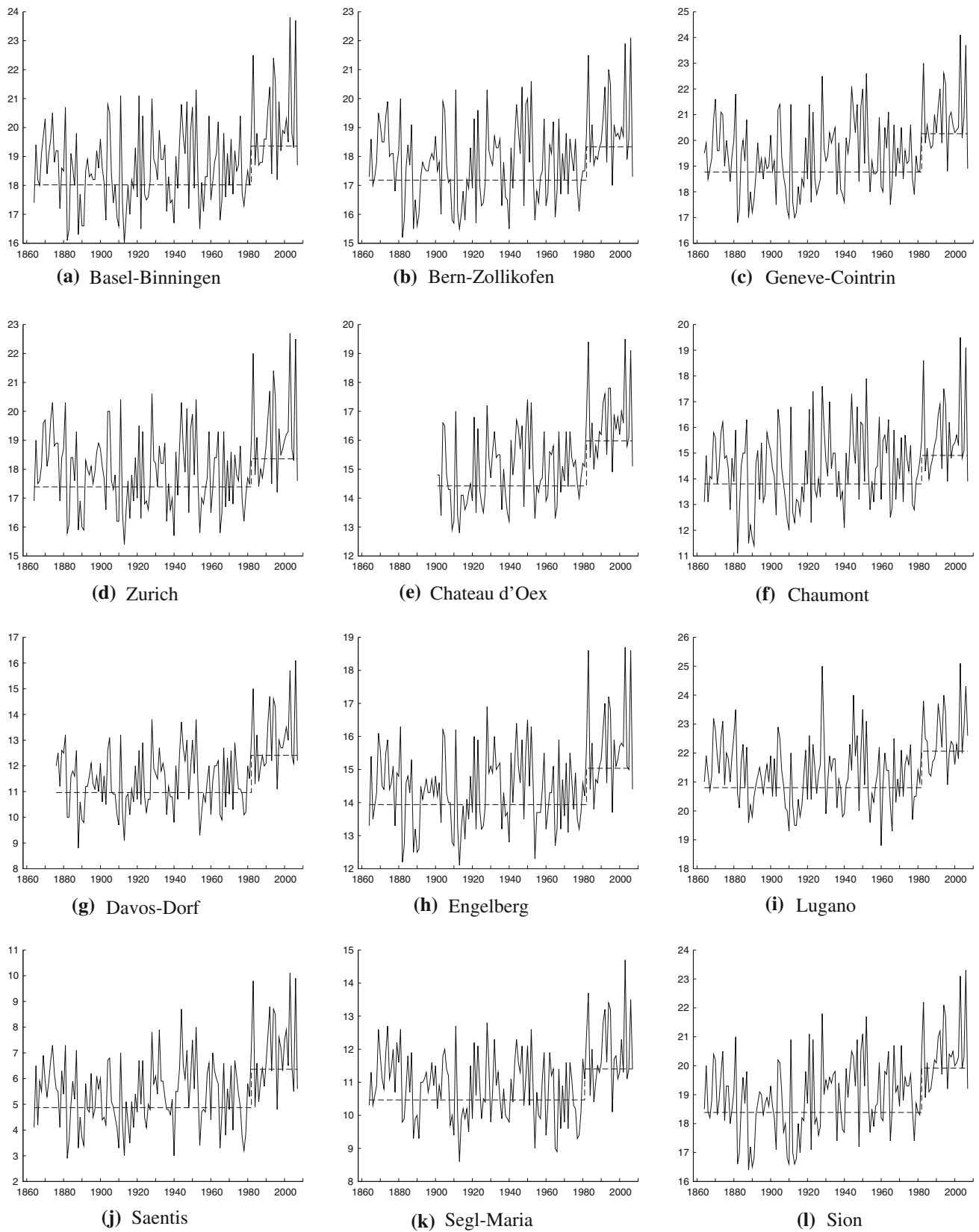
Fig. 1 Observed annual maximum of monthly mean temperature (*solid line*) and the estimated location parameter of the fitted GEV distribution (*dashed line*); the y-axis shows the temperature in °C
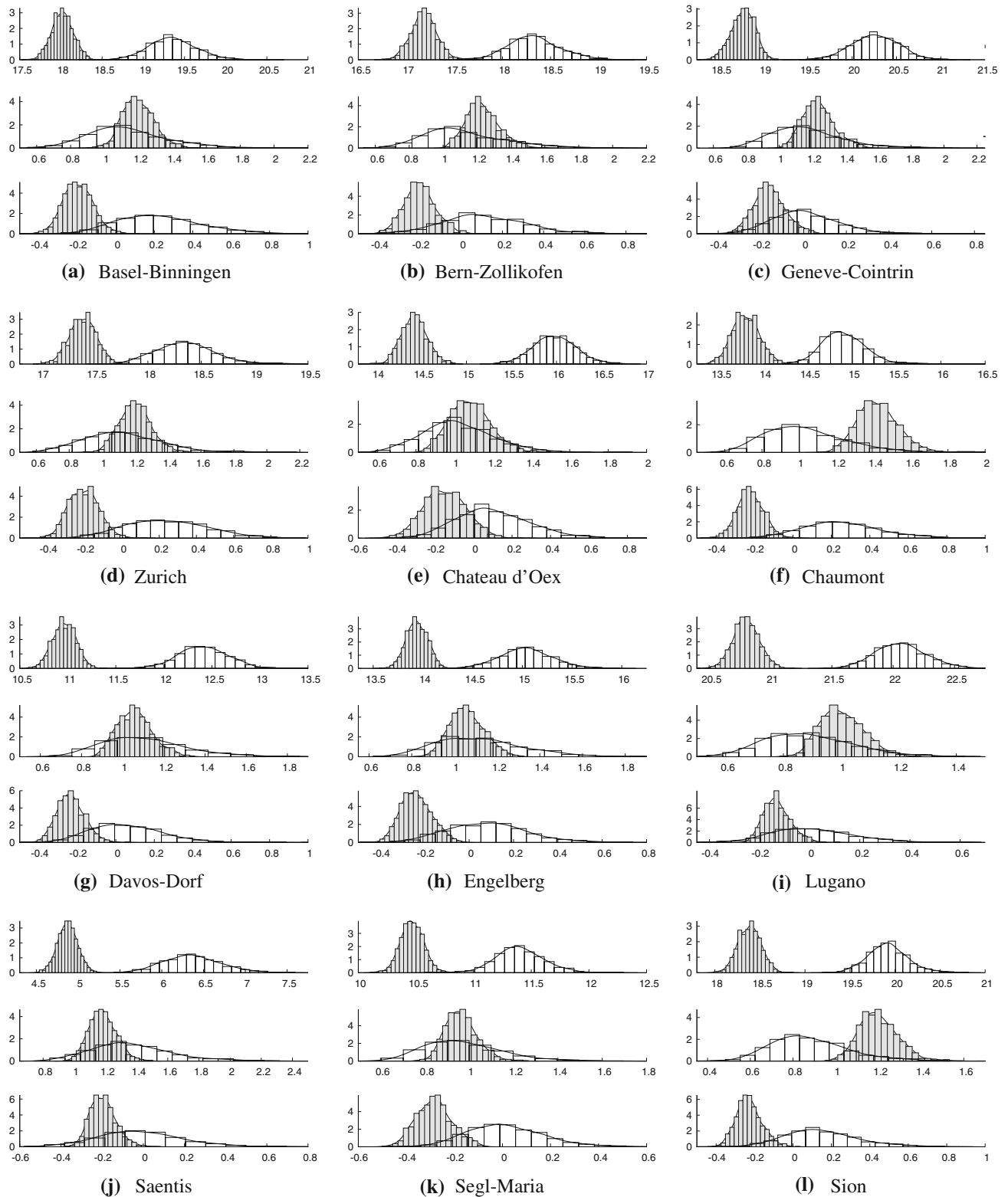
Fig. 2 Posterior distributions of the parameters of the GEV distribution fitted for the first (*shaded*) and the second (*transparent*) subsamples: μ_0 and μ_1—*upper panel*, σ_0 and σ_1—*middle panel*, ξ_0 and ξ_1—*lower panel*
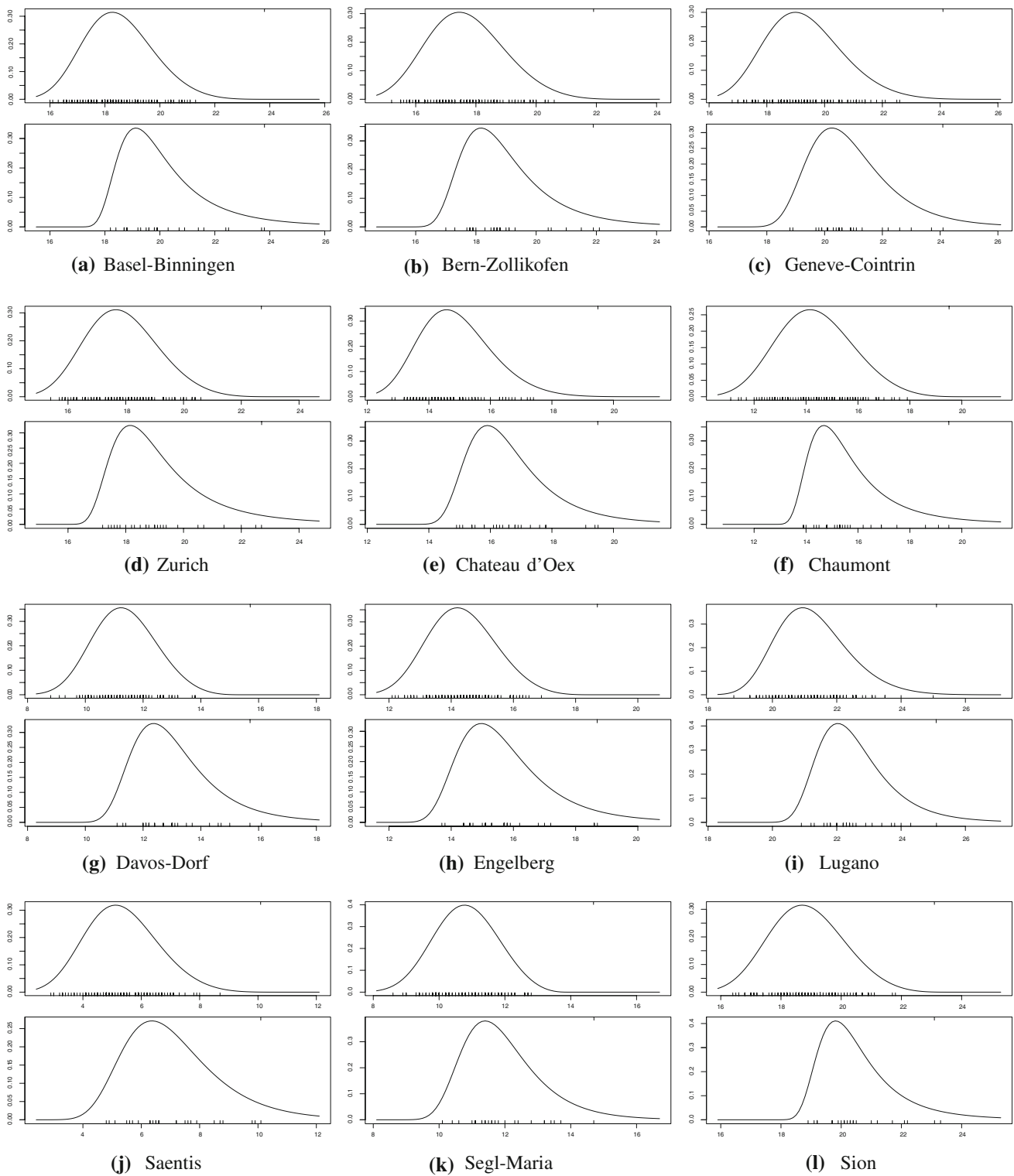
Fig. 3 Fitted GEV probability density function: the corresponding parameter estimates are given in Table 4: the first subsample (*upper panel*), the second subsample (*lower panel*). The temperature recording of summer 2003 is marked as the rug on the top of each panel. The x-axis shows the annual maximum of monthly mean temperature in °C
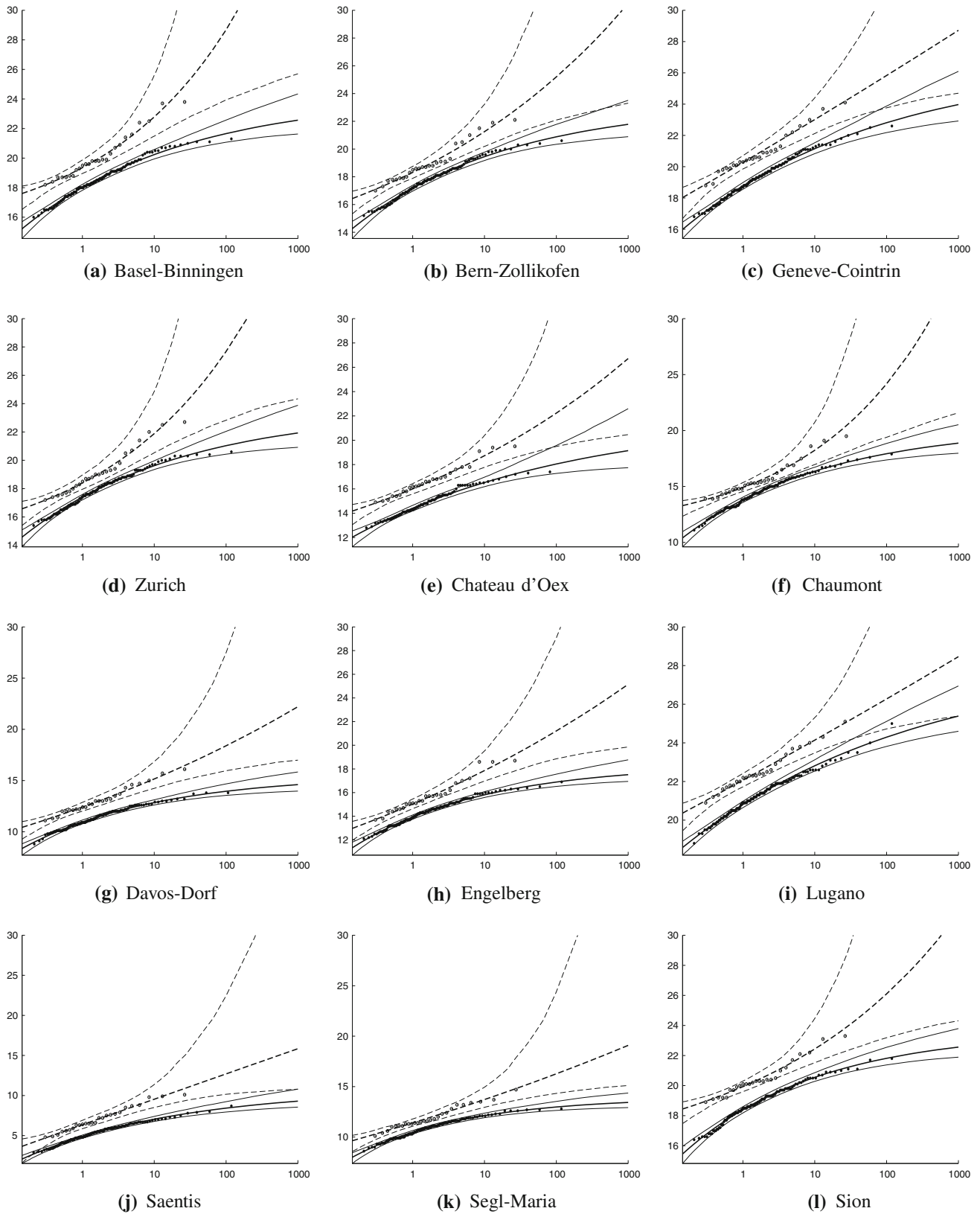
Fig. 4 Return level plot with the corresponding 95% credibility interval for the first (*solid line*) and the second (*dashed line*) subperiods: filled and empty circles correspond to the empirical estimates for the first and the second subperiods, respectively; the x-axis shows the return period (years) $1/p$ on a logarithmic scale, the y-axis shows the temperature in °C
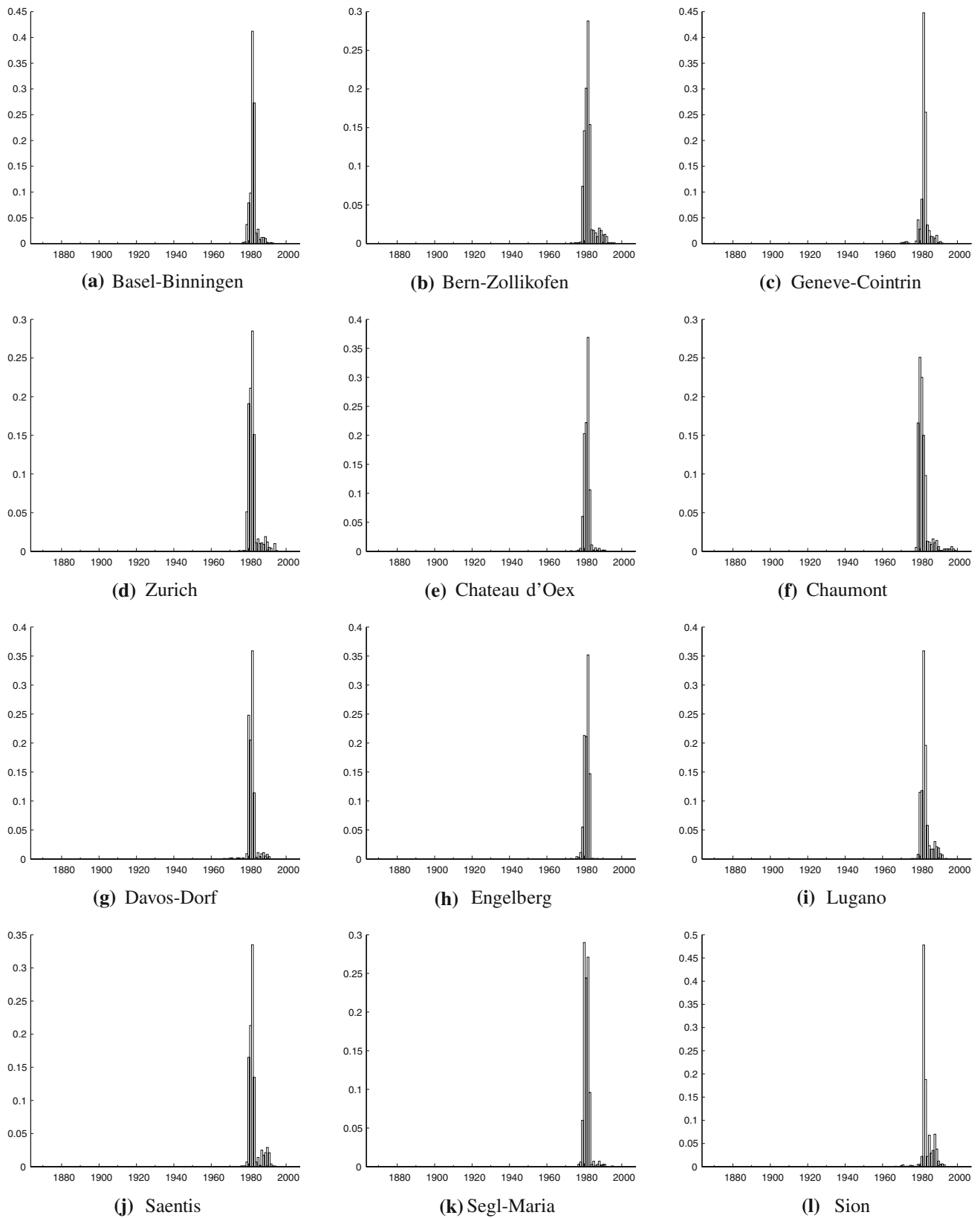
(a) Basel-Binningen

(b) Bern-Zollikofen

(c) Geneve-Cointrin

(d) Zurich

(e) Chateau d'Oex

(f) Chaumont

(g) Davos-Dorf

(h) Engelberg

(i) Lugano

(j) Saentis

(k) Segl-Maria

(l) Sion

Fig. 5 Posterior distribution of the breakpoint timing parameter

finding, in our opinion, strongly favours our hypothesis that the distribution of maximum temperatures in Switzerland underwent a significant structural change towards more extreme values in the beginning of the 1980s.

In the analysis above we used a longer-term definition of heat waves. In order to check the robustness of our results to alternative shorter-term definitions of heat waves we used daily temperature records provided by the European Climate Assessment (ECA) (Klein Tank et al. 2002). For five selected measurement locations (Basel-Binningen, Geneve-Cointrin, Lugano, Saentis, Zurich), we compared the time series of the annual maximum of daily temperature, of the annual maximum temperature exceeded on three subsequent days, and of the number of days exceeding a high threshold defined by a percentile (99.5%). Our analysis revealed that for such short-term definitions of heat waves, the time series exhibit a differing behaviour for varying measurement locations. Both increases and decreases over time as well as no obvious trends can be observed among the five measurement sites. The main findings of this paper may therefore not be valid for short-term definitions of heat waves. This point, however, will need further analysis. Della-Marta et al. (2007a, b), for example, suggest a warm-bias in instrumental recordings in the late 19th and early 20th century that might be responsible for patterns discovered in the analysis based on short-term definitions of heat waves. Correcting for this bias may result in clearer findings. After having corrected for the warm-bias in daily maximum temperatures statistics, Della-Marta et al. (2007a, b) analyse hot day and heat wave indices based on daily recordings. They find that the length of western European summer heat waves has effectively doubled and the frequency of hot days tripled over the period from 1880 to 2005.

4.3 Model comparison and prediction of the heat event 2003

In a thought experiment we want to check whether already in 2002 there was enough information available for inferring a much higher probability for events like the 2003 heatwave. For this purpose, we first replicate our analysis relying only on observations before 2003 and calculate the predictive probabilities for the extreme values exceeding the 2003 numbers under assumption of a structural break. In Table 5 we compare the predictive probability and predictive return period with the ones inferred without assuming a structural break. The introduction of the break point leads to a remarkable increase in the probability for extreme temperatures and the respective decrease of the return period as compared to a stationary climate regime. The two decades of the new regime had been long enough to produce sufficient information for detecting this probability shift in the order of a magnitude already before the 2003 heatwave actually happened.

We further analyse how recent temperature records have altered the assessment of future extremes. We redo our analysis now based on all observations up to 2007, still under assumption of a structural break. These observations lead to a further increase in exceedance probability, roughly speaking doubling them (see Table 5).

Table 5 Predictive probability and predictive return period for exceeding the temperature of 2003

	Constant parameter model		Changepoint model			
	1864–2002[a]		1864–2002		1864–2007	
	Predictive probability	Predictive return period (years)	Predictive probability	Predictive return period (years)	Predictive probability	Predictive return period (years)
Basel-Binningen	0.0013	774.51	0.0354	27.72	0.0605	16.02
Bern-Zollikofen	0.0033	299.62	0.0346	28.41	0.0612	15.83
Geneve-Cointrin	0.0018	558.53	0.0171	58.15	0.0448	21.84
Zurich	0.0017	572.27	0.0413	23.72	0.0660	14.64
Chateau d'Oex	0.0058	172.50	0.0342	28.72	0.0561	17.33
Chaumont	0.0011	919.12	0.0276	35.76	0.0553	17.57
Davos-Dorf	0.0025	406.45	0.0333	29.49	0.0661	14.63
Engelberg	0.0021	471.44	0.0325	30.31	0.0557	17.45
Lugano	0.0028	359.94	0.0224	44.15	0.0413	23.69
Saentis	0.0028	356.30	0.0360	27.27	0.0605	16.03
Segl-Maria	0.0008	1,310.46	0.0232	42.52	0.0360	27.29
Sion	0.0011	912.24	0.0274	36.00	0.0599	16.19

[a] For Chateau d'Oex and Davos-Dorf the available sample starts in 1901 and 1876, respectively.

5 Conclusions

In this paper, we model the annual maxima of monthly mean temperatures in Switzerland measured in twelve locations over the period that in most cases cover 1864 till 2007. We apply the GEV distribution whose parameters are assessed using Bayesian methods.

Our main findings are the following: First, a mechanic application of fitting a GEV distribution on data prior to 2002 suggests that the heat wave 2003 was a very unusual phenomenon. Second, a more careful examination of the time series reveals that the pattern of occurrence of an enduring heat wave may have drastically changed already in the beginning of the 1980s. In order to investigate this formally we have introduced a breakpoint parameter in our model which endogenously splits the sample into two sub-samples. We find very clear statistical evidence in favour of a structural break, convincingly supported by the fact that the posterior distribution of the breakpoint timing is very tight and centered at 1982 plus/minus 1 year at every measurement location.

Third, we show that after accounting for a shift in the parameters of the fitted GEV distribution, the event observed in 2003 appears not that improbable after all. More generally, the huge discrepancy in implications of parameter estimates of different subperiods is well illustrated with return level plots which suggest that for a given return period the likelihood of observing extreme events has increased substantially, or, equally, a certain threshold is expected to be surpassed within much shorter time periods.

Fourth, the implications of our research is that the first heat wave—now largely omitted from public discussion—occurred in 1983 which by historical standards was a year characterised by unusually high temperatures such that a new record has been established in all measurement stations but one. The conclusion of our analysis of extreme temperatures in Switzerland is that a careful examination of developments in the past combined with an appropriate statistical framework may have provided signals that could have mitigated some consequences of the heat wave observed in 2003. Moreover, such a procedure can serve as a useful tool for assessing the likelihood of more extreme things to come.

References

Begert M, Schlegel T, Kirchhofer W (2005) Homogeneous temperature and precipitation series of Switzerland from 1864 to 2000. Int J Climatol 25:65–80

Beniston M (2004) The 2003 Heat Wave in Europe: a shape of things to come? An analysis based on Swiss climatological data and model simulations. Geophys Res Lett 31:2022–2026

Carlin BP, Gelfand AE, Smith AFM (1992) Hierarchical Bayesian analysis of changepoint problems. Appl Stat 41(2):389–405

Coles S (2001) An introduction to statistical modeling of extreme values. Springer Series in Statistics. Springer, New York

Coles S, Pericchi L (2003) Anticipating catastrophes through extreme value modelling. Appl Stat 52(4):405–416

De Bono A, Giuliani G, Kluser S, Peduzzi P (2004) Impacts of summer 2003 heat wave in Europe. Environmental Alert Bulletin, UNEP

Della-Marta PM, Haylock MR, Luterbacher J, Wanner H (2007a) Doubled length of western European summer heat waves since 1880. J Geophys Res (Atmospheres) 112:D15103

Della-Marta PM, Luterbacher J, von Weissenfluh H, Xoplaki E, Brunet M, Wanner H (2007b) Summer heat waves over western Europe 1880–2003, their relationship to large-scale forcings and predictability. Clim Dyn 29:251–275

Fink AH, Brücher T, Krüger A, Leckebusch GC, Pinto JG, Ulbrich U (2004) The 2003 European summer heatwaves and drought: synoptic diagnosis and impacts. Weather 59:209–216

Gamerman D (1997) Markov Chain Monte Carlo: stochastic simulation for Bayesian inference. Chapman & Hall, London

García-Herreira R, Díaz J, Trigo R, Luterbacher J, Fischer E (2009) A review of the European summer heat wave of 2003. Crit Rev Environ Sci Technol (in press)

Geweke J (1992) Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: Bernardo JM, Berger JO, Dawid AP, Smith AFM (eds) Bayesian statistics. Oxford University Press, Oxford, pp 169–193

Grize L, Huss A, Thommen O, Schindler C, Braun-FahrlSnder C (2005) Heat wave 2003 and mortality in Switzerland. Swiss Med Wkly 135:200–205

Jaeger CC, Krause J, Haas A, Klein R, Hasselmann K (2008) A method for computing the fraction of attributable risk related to climate damages. Risk Anal 28(4):815–823

Khaliq MN, Ouarda TBMJ, St-Hilaire A, Gachon P (2007) Bayesian change-point analysis of heat spell occurrences in Montreal, Canada. Int J Climatol 27(6):805–818

Klein Tank A, Wijngaard JB, Können GP, Böhm R, Demarée G, Gocheva A, Mileta M, Pashiardis S, Hejkrlik L, Kenr-Hansen C, Heino R, Bessemoulin P, Müller-Westermeier G, Tzanakou M, Szalat S, Pálsdóttir T, Fitzgerald D, Rubin S, Capaldo M, Maugeri M, Leitass A, Bukantis A, Aberfeld R, van Engelen AFV, Forland E, Mietus M, Coelho F, Mares C, Razuvaev V, Nieplova E, Cegnar T, Antonio López J, Dahlström B, Moberg A, Kirchhofer W, Ceylan A, Pachaliuk O, Alexanders LV, Petrovics P (2002) Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. Int J Climatol 22:1441–1453

Leadbetter MR, Lindgren G, Rootzen H (1983) Extremes and related properties of random sequences and processes. Springer Series in Statistics. Springer, New York

Mohammad-Djafari A, Feron O (2006) Bayesian approach to change points detection in time series. Int J Imaging Syst Technol 16(5):215–221

Moreno E, Casella G, Garcia-Ferrer A (2005) An objective Bayesian analysis of the change point problem. Stoch Environ Res Risk Assess 19:191–204

Perreault L, Parent E, Bernier J, Bobee B, Slivitzky M (2000) Retrospective multivariate Bayesian change-point analysis: a simultaneous single change in the mean of several hydrological sequences. Stoch Environ Res Risk Assess 14(4):243–261

Perron P (1989) The great crash, the oil price shock, and the unit root hypothesis. Econometrica 57(6):1361–1401

ProClim (2005) Hitzesommer 2003 Synthesebericht. ProClim/ Schweizerische Akademie der Naturwissenschaften

Raftery AE, Akman VE (1986) Bayesian analysis of a poisson process with a change-point. Biometrika 73(1):85–89

Renard B, Lang M, Bois P (2006) Statistical analysis of extreme events in a non-stationary context via a Bayesian framework: case study with peak-over-threshold data. Stoch Environ Res Risk Assess 21(2):97–112

Robinson P (2001) On the definition of a heat wave. J Appl Meteorol 40:762–775

Schär C, Fischer E (2008) Warnsignal Klima: Gesundheitsrisiken – Gefahren für Pflanzen, Tiere & Menschen, Chapter Der Einfluss des Klimawandels auf Hitzewellen und das Sommerklima Europas, pp 50–55

Schär C, Vidale P, Lüthi D, Frei C, Häberli C, Liniger MA, Appenzeller C (2004) The role of increasing temperature variability for European summer heatwaves. Nature 427:332–336

Seneviratne SI, Luthi D, Litschi M, Schär C (2006) Land-atmosphere coupling and climate change in europe. Nature 443(7108):205–209

Stephens DA (1994) Bayesian retrospective multiple-changepoint identification. Appl Stat 43(1):159–178

Stott PA, Stone DA, Allen MR (2004) Human contribution to the European heatwave of 2003. Nature 432:610–614

Tamerius J, Wise E, Uejio C, McCoy A, Comrie A (2007) Climate and human health: synthesizing environmental complexity and uncertainty. Stoch Environ Res Risk Assess 21(5):601–613

Trigo RM, García-Herrera R, Díaz J, Trigo IF, Valente MA (2005) How exceptional was the early August 2003 heat wave in France? Geophys Res Lett 32:L10701

World Health Organization (WHO) (2003) The health impacts of 2003 summer heat-waves. Briefing note for the delegations of the fifty-third session of the WHO (World Health Organization) Regional Committee for Europe