Vol. 33: 243–255, 2007

CLIMATE RESEARCH
Clim Res

Published April 20

Regional growth curves and improved design value estimates of extreme precipitation events in the Czech Republic

Jan Kyselý[1,*], Jan Picek[2]

[1]Institute of Atmospheric Physics Academy of Sciences of the Czech Republic, Boční II 1401, 141 31 Prague, Czech Republic
[2]Technical University, Hálkova 6, 461 17 Liberec, Czech Republic

ABSTRACT: Regional analysis was used to improve estimates of the probabilities of extreme precipitation events in the Czech Republic. The delimitation of 4 homogeneous regions was based on statistical procedures (cluster analysis of site characteristics and subsequent tests for regional homogeneity). The regions also reflect climatological differences in precipitation regimes and synoptic patterns that cause heavy precipitation. The Generalized Extreme Value (GEV) distribution was identified as the most suitable distribution for modelling maximum annual 1 to 7 d precipitation amounts according to the L-moment ratio diagram and goodness-of-fit tests. Only in the northeast region (which is most prone to the occurrence of extremely high precipitation totals) was the Generalized Logistic (GLO) distribution preferred. The regional approach considerably lessened the between-site variation of estimates of the shape parameter of the GEV/GLO distribution compared to at-site procedures, and the estimates of high quantiles (e.g. 50 yr return values) were more reliable and climatologically consistent in individual regions. Different growth curve shapes are characteristic of the 4 regions, the between-region variability being larger for multi-day than 1 d events. Particularly noteworthy is the heavy tail of distributions of 5 and 7 d annual maxima in the northeast region, reflected also in the inapplicability of the general 4-parameter kappa distribution in regional homogeneity tests. A nonparametric statistical test on the value of the tail index supports a hypothesis that data in this flood-prone area may be drawn from a distribution with a very heavy tail for which L-moments do not exist.

KEY WORDS: Extreme precipitation event · Regional frequency analysis · L-moments · Design values · Central Europe · Czech Republic

1. INTRODUCTION

Extreme precipitation events in central Europe, in particular those that resulted in the massive summer floods of July 1997 (Odra basin) and August 2002 (Elbe basin), have recently been the subject of many studies (e.g. Brázdil 1998, Květoň et al. 2002, Brázdil et al. 2004, Engel 2004, Kundzewicz et al. 2005, Řezáčová et al. 2005). One of the most serious limitations of previous statistical modelling of precipitation extremes in the Czech Republic and surrounding central European countries is that design values (i.e. precipitation amounts corresponding to fixed return periods) have generally been derived from analyses of separate datasets at individual sites.

A regional approach to frequency analysis consists in substituting space for time by using observations from different sites in a region to compensate for short time-series records at individual sites. Taking into account measurements at neighbouring locations, or from particular regions is beneficial from the statistical point of view — frequency estimates of rare events become more reliable and less sensitive to random fluctuations. The method is climatologically sound only if one can assume that basic statistical characteristics of extremes do not differ among sites (regions that meet such conditions are termed 'homogeneous'). The approach is most advantageous for variables such as precipitation that exhibit high and largely random spatial variability.

*Email: kysely@ufa.cas.cz

Regional frequency analysis has become a widely-used tool in hydrological (e.g. Pilon & Adamowski 1992, Adamowski 2000, Gottschalk & Krasovskaia 2002, Kjeldsen & Rosberg 2002) as well as climatological studies (e.g. Guttman et al. 1993, Naghavi & Yu 1995, Smithers & Schulze 2001, Fowler & Kilsby 2003). As hinted above, it employs observations from a suite of measuring sites in order to estimate probability distributions at any particular location. An example is the 'index storm' procedure; the basic assumption is that the frequency distributions at sites over a homogeneous region are identical except for a site-specific scaling factor, the index storm. The advantage of regional over single-site estimation is greater at the distribution tails, which are of interest in many practical applications, including planning for weather-related emergencies and design and operation of water reservoirs.

In this study, a regional frequency analysis based on L-moments (e.g. Hosking & Wallis 1997) was used to construct regional growth curves and improve estimates of the design values of extreme precipitation events in the Czech Republic. After a brief description of the data and precipitation characteristics of the area under study (Section 2) and the methods used (Section 3), the results of the 3 main steps of the regional algorithm are dealt with in Sections 4 to 6—these are the formation of homogeneous regions, the choice of frequency distributions, and the estimation of the parameters and quantiles of the fitted distributions. Two statistical issues that point to very heavy tails of distributions of precipitation extremes and that have not yet been dealt with in the climatological literature—a failure of regional homogeneity tests with the kappa distribution under special conditions, and the results of a nonparametric test on the tail index—are discussed in Section 7. Benefits of the regional approach and concluding remarks are summarized in Section 8.

2. DATA AND STUDY AREA

2.1. Input datasets

Daily precipitation totals measured at 78 stations operated by the Czech Hydrometeorological Institute were used (Fig. 1). The observations span the period 1961–2000; there are no missing values in the dataset. The altitudes of the stations range from 158 to 1324 m above sea level (a.s.l.). Samples of maximum annual 1,

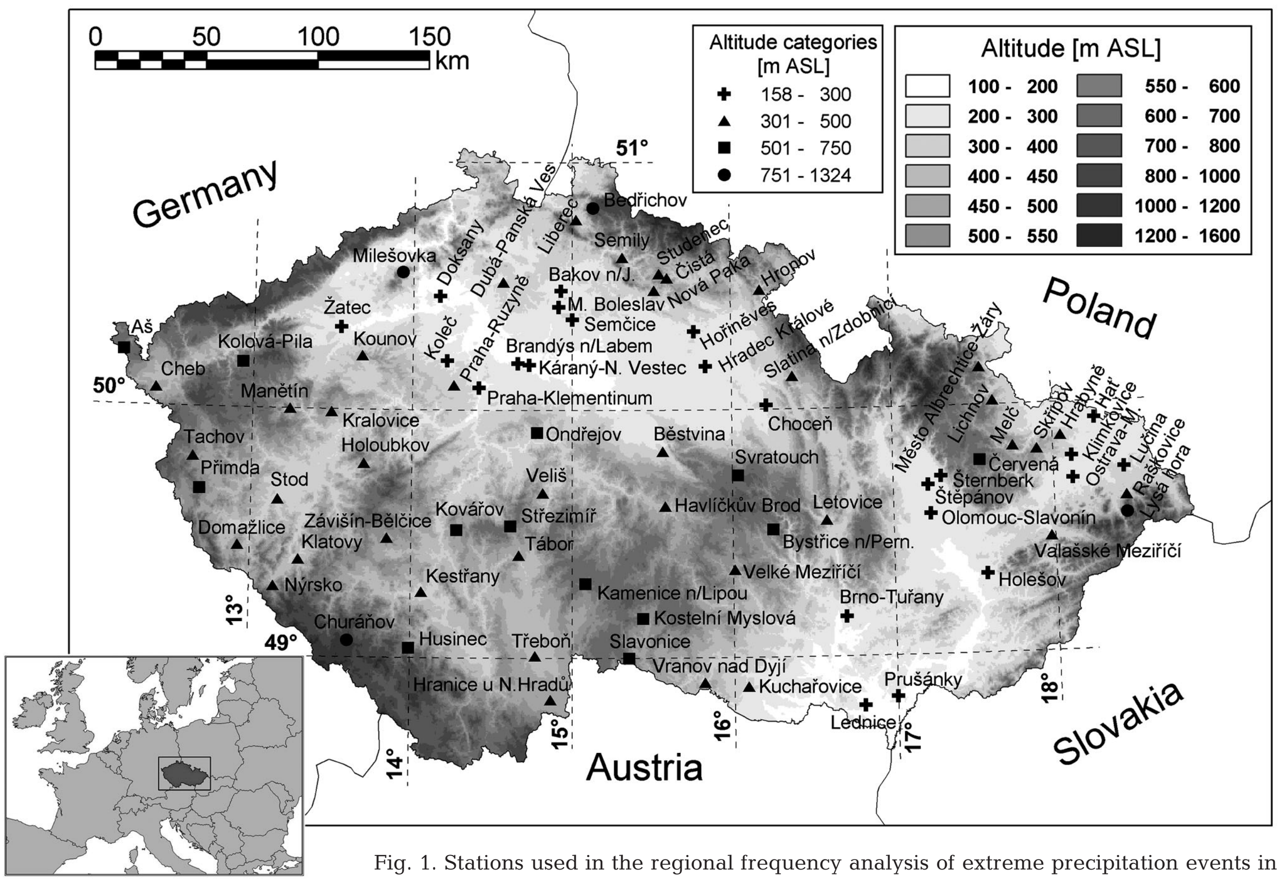

Fig. 1. Stations used in the regional frequency analysis of extreme precipitation events in the Czech Republic. Altitude categories (m above sea level, ASL) are indicated by symbols

3, 5 and 7 d precipitation amounts were drawn from each station's records and examined as extreme precipitation events.

The data underwent standard quality checking for gross errors (Coufal et al. 1992). Sites discordant with the group as a whole were identified using a discordancy measure based on L-moments (Hosking & Wallis 1993), and examined for errors or sources of potential unreliability in measurements. However, all values of the discordancy measure larger than the critical value originated from observed outliers (mainly related to the heavy rainfall event resulting in the 1997 floods) and we did not find any physical grounds (based on precipitation patterns related to e.g. orography and atmospheric circulation) to exclude them from further analysis (cf. Fowler & Kilsby 2003).

2.2. Precipitation patterns in the Czech Republic

Relatively large spatial and temporal variability is typical for precipitation in the Czech Republic. In the winter half-year (October to March), precipitation is linked to passages of cyclones and frontal systems, and usually falls from stratiform clouds. Although winter precipitation events are frequently long-lasting (compared to showers and storms in summer) and the number of rainy or snowy days is greater, monthly precipitation amounts are smaller in winter than in summer. Almost two thirds of the total annual precipitation falls in the summer half-year (April to September); produced largely by convective clouds that may or may not be related to atmospheric fronts. At most sites, January–February are the driest and June–July the wettest months (Tolasz 2007).

Heavy precipitation events result from an inflow of moist maritime air from the North Atlantic (west to northwest advection) and, most importantly, an inflow of warm moist air from the Mediterranean Sea. The influence of the latter tends to be more pronounced in the eastern part of the Czech Republic, in Moravia and Silesia, and cyclones of Mediterranean origin are the frequent cause of multi-day precipitation maxima that may result in flood events (e.g. Šamaj et al. 1983, Štekl et al. 2001).

Spatial variability of precipitation is linked to atmospheric circulation as well as orographic features, with mountain ranges in the north and northeast parts of the Czech Republic receiving the largest totals (the mean annual amounts exceeding 1400 mm). The driest lowland areas receive only about 400 mm in mean annual precipitation. The huge spatial and temporal variability of precipitation in the Czech Republic is well illustrated by the fact that storms in the summer season (lasting not longer than several hours) may bring up to 50% of the mean annual precipitation, and that the record-breaking 3 d totals in mountainous areas exceed mean annual precipitation in lowland regions (e.g. Štekl et al. 2001).

3. METHODS

3.1. L-moments

Several steps of the regional analysis incorporate L-moments, an alternative set of scale and shape statistics of a data sample or a probability distribution (e.g. Hosking 1990, Ulrych et al. 2000). Their derivation is based on order statistics obtained by sorting the sample $\{X_1, X_2, ..., X_n\}$ of n independent realisations of variable X in ascending order $\{X_{1:n}, X_{2:n}, ..., X_{n:n}\}$; the subscript $k{:}n$ denotes the kth smallest number in the sample of length n. L-moments λ_k are defined as expectations of linear combinations of these order statistics,

$$\lambda_k = \frac{1}{k}\sum_{j=0}^{k-1}(-1)^j\binom{k-1}{j}E\left(X_{k-j:k}\right) \tag{1}$$

where E denotes the expectation operator. L-moment ratios are the L-coefficient of variation $\frac{\lambda_2}{\lambda_1}$ (L-CV), the L-skewness $\frac{\lambda_3}{\lambda_2}$ (τ_3) and the L-kurtosis $\frac{\lambda_4}{\lambda_2}$ (τ_4); except for some special cases of small samples, they take values between –1 and +1.

The kth sample L-moment λ_k ($k \le n$) can be estimated as

$$l_k = \sum_{l=0}^{k-1}(-1)^{k-l-1}\binom{k-1}{l}\binom{k+l-1}{l}b_l \tag{2}$$

where

$$b_l = \frac{1}{n}\sum_{i=1}^{n}\frac{(i-1)(i-2)\ldots(i-l)}{(n-1)(n-2)\ldots(n-l)}X_{i:n},\ l \ge 1,\ \text{and}\ b_0 = \frac{1}{n}\sum_{i=1}^{n}X_{i:n}$$

3.2. Generalized Extreme Value (GEV) distribution

The cumulative distribution function of the GEV distribution with parameters ξ (location), α (scale) and k (shape) is (e.g. Hosking & Wallis 1997)

$$F(x) = e^{-e^{-y}}$$

where

$$y = -\frac{\ln\left(1-\frac{k(x-\xi)}{\alpha}\right)}{k},\ k \ne 0$$
$$y = \frac{x-\xi}{\alpha},\ k = 0 \tag{3}$$

The distribution is bounded at $\xi + \frac{\alpha}{k}$ from right (left) if $k > 0$ ($k < 0$). L-moments are defined for $k > -1$; the first 3 population L-moments are

$$\lambda_1 = \xi + \alpha\frac{1-\Gamma(1+k)}{k},\ \lambda_2 = \alpha\frac{(1-2^{-k})\Gamma(1+k)}{k},$$
$$\text{and } \lambda_3 = \alpha\frac{\Gamma(1+k)(-1+3.2^{-k}-2.3^{-k})}{k} \tag{4}$$

where Γ denotes the gamma function.

The method of L-moments fits the GEV distribution by choosing its parameters so that the first 3 L-moments λ_1, λ_2, λ_3 match the corresponding estimates l_1, l_2, l_3. The resulting L-moment estimators of k, α and ξ are given by

$$k = 7.8590z + 2.9554z^2,\ \alpha = \frac{l_2 k}{(1-2^{-k})\Gamma(1+k)},$$
$$\text{and } \xi = l_1 + \alpha\frac{\Gamma(1+k)-1}{k} \tag{5}$$

where $z = \frac{2}{3+\frac{l_3}{l_2}} - \frac{\ln 2}{\ln 3}$

3.3. Generalized Logistic (GLO) distribution

The cumulative distribution function of the GLO distribution with parameters ξ (location), α (scale) and k (shape) is

$$F(x) = \frac{1}{1+e^{-y}} \tag{6}$$

where

$$y = -\frac{\ln\left(1-\frac{k(x-\xi)}{\alpha}\right)}{k},\ k \neq 0$$
$$y = \frac{x-\xi}{\alpha},\ k = 0$$

The distribution is bounded at $\xi + \frac{\alpha}{k}$ from right (left) if $k > 0$ ($k < 0$). L-moments are defined for $-1 < k < 1$; the first 3 population L-moments are

$$\lambda_1 = \xi + \alpha\left(\frac{1}{k} - \frac{\pi}{\sin k\pi}\right),\ \lambda_2 = \frac{\alpha k\pi}{\sin k\pi},\ \text{and } \lambda_3 = -k\lambda_2 \tag{7}$$

The L-moment estimators of k, α and ξ are given by

$$k = -\frac{l_3}{l_2},\ \alpha = \frac{l_2 \sin k\pi}{k\pi},\ \text{and } \xi = l_1 - \alpha\left(\frac{1}{k} - \frac{\pi}{\sin k\pi}\right) \tag{8}$$

3.4. Testing for regional homogeneity

The regional homogeneity tests applied to samples of annual maxima of 1 to 7 d precipitation amounts were those of Lu & Stedinger (1992) and Hosking & Wallis (1993); refer to these studies for a mathematical description. (Note that in the index storm approach 'regional homogeneity' means that probability distributions at individual locations in a region are identical, apart from a site-specific scaling factor.) Generally, the tests are based on a quantity that measures a selected aspect of the frequency distribution (a 10 yr event in the Lu-Stedinger test, and L-moment ratios in the Hosking-Wallis tests), and compare the 'at-site' estimates with the regional estimate of this quantity. Simulations of homogeneous regions with sites having record lengths the same as the observed data are necessary to estimate the variance of 90% sample quantiles at individual sites (in the Lu-Stedinger test), and the mean and variance of a dispersion measure (in the Hosking-Wallis tests). The number of realisations we performed was 500 in all Monte Carlo experiments; the GEV and kappa distributions (Hosking 1994) were used in the simulations. Three versions of the Hosking-Wallis tests were applied, based on L-CV, L-skewness and L-kurtosis and their combinations.

3.5. Goodness-of-fit test based on L-kurtosis

Goodness-of-fit of various candidate 3-parameter probability distributions was evaluated in terms of the difference between L-kurtosis τ_4 of the fitted distribution and the regional average L-kurtosis τ_4^R (Hosking & Wallis 1997). A comparison with the sampling variability of τ_4^R is performed to assess the significance of this difference. The test statistic is

$$Z = \frac{\tau_4 - \tau_4^R + B_4}{\sigma_4} \tag{9}$$

where B_4 denotes the bias and σ_4 the standard deviation of τ_4^R, both obtained by simulations of a homogeneous region with the kappa distribution. The number of replications was 500. The distribution was rejected at the 0.10 (0.05) level if $|Z| > 1.64$ ($|Z| > 1.96$).

3.6. Test on the tail index

Inference on values of the tail index of heavy-tailed distributions (including point estimates and testing of hypotheses) is a topical issue in statistical literature (e.g. Hill 1975, Cheng & Peng 2001, Hasofer & Wang 1992). (Note that the tail index usually corresponds to shape parameter k of a 3-parameter distribution.) Since L-moments do not exist for some heavy tailed distributions (e.g. the GEV distribution with $k \leq -1$), it may not always be possible to use procedures based on L-moments, and the application of an L-moment based algorithm may yield biased results. A nonparametric test on the tail index proposed by Jurečková & Picek (2001, 2004) is employed to verify whether the right tail of an underlying distribution may be as heavy as the tail of the Pareto distribution with shape parameter k_0 or heavier. The test is based on the empirical distribution function of maxima of subsamples; for mathematical details see Appendix. The test is nonparametric and works under weaker conditions than usual point estimators of the tail index.

4. REGIONALIZATION

In climatologically homogeneous areas with a simple orography (e.g. Belgium; Gellens 2002) the issue of regionalization is easily handled, and all the available data may usually be considered to be drawn (after rescaling) from the same population. Since an area with spatially variable mechanisms leading to heavy precipitation amounts (Šamaj et al. 1983, Hanslian et al. 2000, Štekl et al. 2001) and with a relatively complex orography (Fig. 1) is under study, the basic step of the regional analysis consisted in the formation of regions that are homogeneous according to the statistical characteristics of precipitation extremes.

In accordance with common practice (e.g. Hosking & Wallis 1997, Smithers & Schulze 2001, Kjeldsen et al. 2002), a cluster analysis of ‘site characteristics’ (longitude, latitude, elevation, mean annual precipitation, mean ratio of summer half-year to winter half-year precipitation, and mean annual number of dry days) yielded a number of preliminary partitionings into groups of sites. The most promising originated from Ward’s method of cluster analysis using 4 clusters; it produced relatively distinct clusters of a reasonable size, 2 of them forming large homogeneous areas comprising more than 80% of sites. (Unlike the average linkage algorithm, Ward’s method does not suffer from the undesirable snowball effect that results in one big cluster to which smaller clusters are stuck, more and more dissimilar from the mean.) A number of adjustments and reallocations were tested to improve the homogeneity of regions and to make them geographically and climatologically coherent; altogether 35 different partitionings of stations into regions were examined in terms of the Hosking-Wallis and Lu-Stedinger regional homogeneity tests (Section 3.4); for details of the procedure see Kyselý et al. (2005).

The final partitioning recognizes 4 homogeneous regions (Table 1, Fig. 2) that reflect climatological differences in precipitation regimes and synoptic patterns associated with heavy precipitation. The 2 large regions distinguish between lowland (Region 1) and higher-elevated (Region 2) locations in most of the area of the Czech Republic; while the 2 smaller regions possess distinctly different precipitation regimes. Enhanced mean annual precipitation as well as heavy precipitation amounts due to orographic effects and the greater influence of Mediterranean cyclones is characteristic of Region 3 in the NE part of the Czech Republic. Enhanced mean annual precipitation, a small number of dry days and increased precipitation (including extremes) in winter due to the larger influence of cloud belts and atmospheric fronts associated with Atlantic cyclones, are particular features of Region 4 in the N part of the Czech Republic.

The fact that the regions do not depend on the duration of events (in accord with results of the statistical tests, they are identical for 1 to 7 d precipitation totals) is useful from the practical point of view (cf. Werick et al. 1993, Smithers & Schulze 2001). The most elevated and easternmost station, Lysá hora (in the Beskydy Mountains in the NE part of the Czech Republic), cannot be classified into any of the regions; its inclusion in Region 3 (to which it might be geographically allocated) distorts the regional homogeneity. It cannot be concluded without additional precipitation data from the complex terrain of the NE region whether this is only due to sampling variability or reflects real different features of the distribution of precipitation extremes.

Taking into account the area of the Czech Republic (78.9×10^3 km^2), the number of clusters is reasonable compared to, for example, the 9 regions entering the regional frequency analysis of precipitation extremes in the UK (244×10^3 km^2; Fowler & Kilsby 2003) or the 3 regions delineated in Slovakia (49×10^3 km^2; Gaál 2006).

The stability of the results of the regional homogeneity tests on the final partitioning was examined by a Monte Carlo simulation that consisted of repeatedly removing a given portion of the data (stations) from each region, and performing the tests on the remaining part of the regional sample. These experiments fully supported the homogeneity of the regions formed; there were no occurrences of values of the test statis-

Table 1. Geographical and precipitation characteristics of the 4 homogeneous regions in the Czech Republic. ASL: above sea level. Regions — 1: main lowland; 2: higher-elevated west-central; 3: northeast; 4: north

Characteristic	Region 1	Region 2	Region 3	Region 4
No. of stations	32	28	12	5
Altitude range (m ASL)	158–468	429–1118	220–750	370–495
Mean annual precipitation (mm)	565	674	753	801
Mean maximum annual 5 d precipitation (mm)	60.4	66.0	80.0	68.7
Mean annual no. dry days	223	203	210	201
Ratio summer half-year to winter half-year precipitation	1.91	1.67	1.99	1.27
Percentage maximum annual 5 d precipitation Nov–Mar	6.6	12.3	6.3	32.0

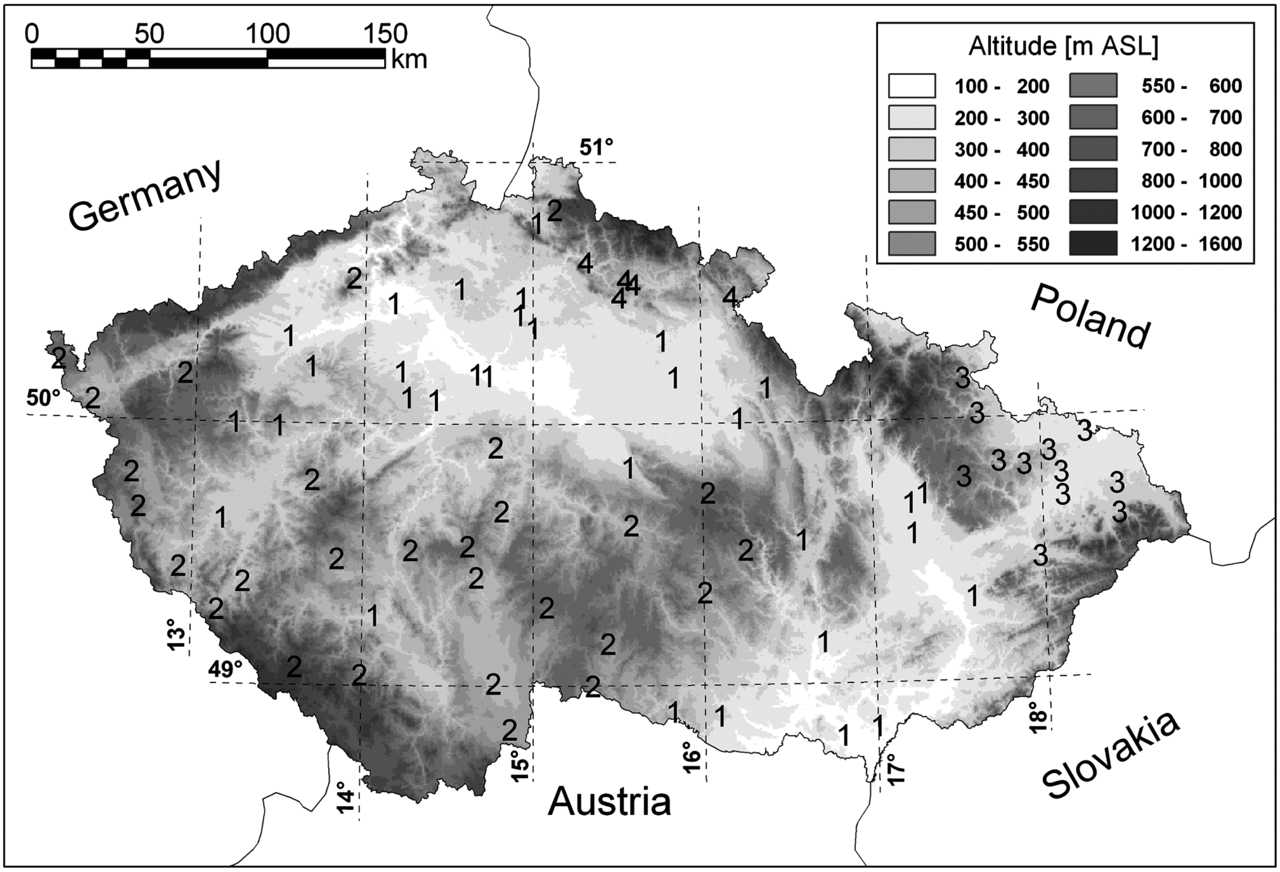

Fig. 2. Final formation of homogeneous regions (1–4) for the regional frequency analysis of precipitation extremes in the Czech Republic

tics of the Hosking-Wallis test based on L-CV ≥2 ('definite heterogeneity') throughout the perturbed samples for all variables (1 to 7 d amounts) and in all regions.

5. IDENTIFICATION OF REGIONAL DISTRIBUTIONS

L-moment ratio diagrams and goodness-of-fit tests were applied to various candidate extreme value distributions. They included GEV, GLO (described in Sections 3.2 and 3.3, respectively) and other frequently used 3-parameter models, lognormal (LN3) and Pearson Type III (PE3) distributions.

The L-moment ratio diagrams for maximum annual 1 and 5 d precipitation amounts are depicted in Fig. 3; the curves show theoretical relationships between L-kurtosis and L-skewness for the candidate distributions. Although the between-site variations in L-moment ratios are large, it can be clearly observed that the shape of the GEV distribution curve is followed most tightly, while the other distributions are likely to be suitable models in particular regions only.

The applicability of individual distributions was evaluated in terms of *Z*-statistics based on a difference between the L-kurtosis of the fitted distribution and the regional average of the L-kurtosis (see Section 3.5). The GEV distribution is appropriate for most durations of extreme precipitation events and in most regions, and it is not rejected for any region-duration pair at the 0.10 significance level (Table 2). All the other distributions are rejected for at least 36% of region-duration pairs at the 0.10 level, and at least 21% of pairs at the 0.05 level. Comparison of the absolute values of the *Z*-statistics for individual distributions also supports the superiority of the GEV distribution, with an exception of Region 3 where the GLO distribution is preferred (obvious from Fig. 3 as well). The tests with the kappa distribution were impracticable in Region 3 for 5 d and 7 d events (see Section 7.1); if the GEV or GLO distribution is employed instead of the kappa distribution in the tests, values of the *Z*-statistics support the superiority of the GLO distribution for this region. Note that the PE3 distribution is unsuitable for modelling extreme precipitation amounts in the Czech Republic, the only exception being Region 4 in the north (for multi-day

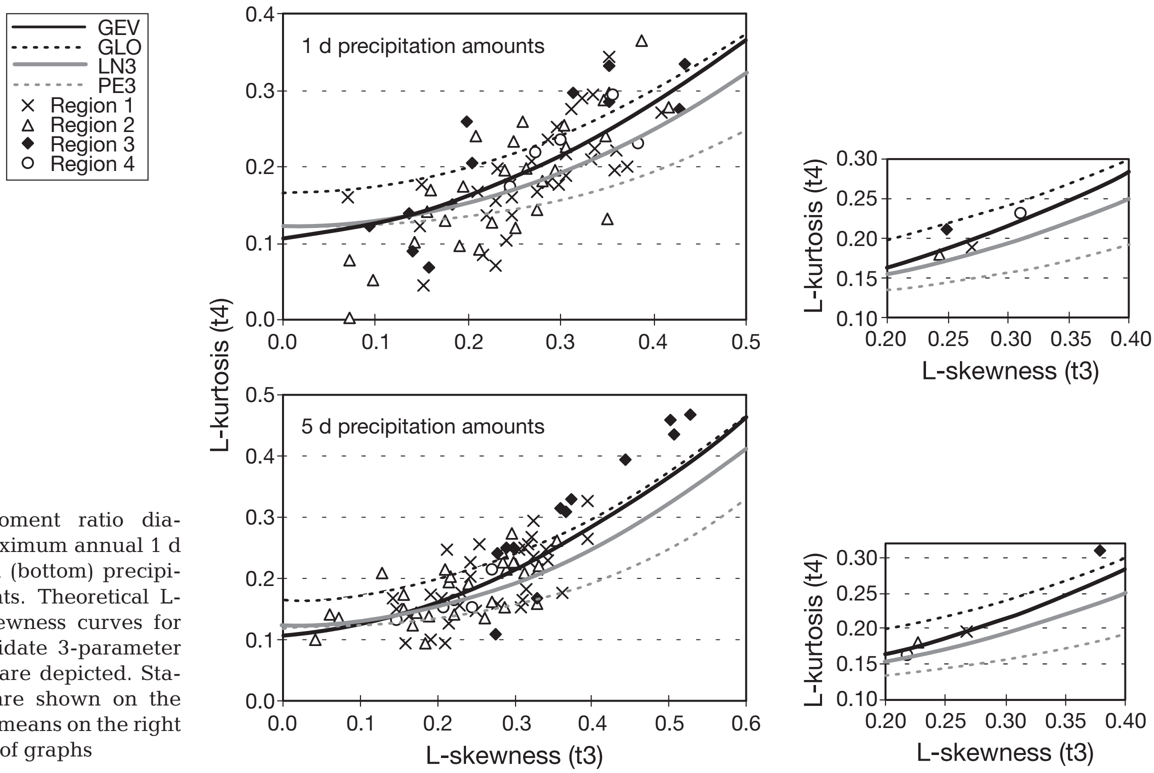

Fig. 3. L-moment ratio diagrams for maximum annual 1 d (top) and 5 d (bottom) precipitation amounts. Theoretical L-kurtosis/L-skewness curves for various candidate 3-parameter distributions are depicted. Station values are shown on the left; regional means on the right pair of graphs

Table 2. Values of *Z*-statistics for candidate 3-parameter distributions and 1, 3, 5 and 7 d precipitation amounts (denoted R1, R3, R5, R7) in the 4 regions. Distribution rejected at: *0.10 level; **0.05 level. The tests for R5 and R7 in Region 3 were impracticable with the kappa distribution (see Section 7.1). Distributions—GEV: Generalized Extreme Value, GLO: Generalized Logistic, LN3: Lognormal, PE3: Pearson Type III

Region	Variable	GEV	GLO	LN3	PE3
1	R1	0.263	2.322**	−0.983	−3.166**
	R3	−0.956	0.828	−2.251**	−4.492**
	R5	−0.443	1.639	−1.671*	−3.827**
	R7	−0.810	1.500	−1.998**	−4.107**
2	R1	0.130	2.361**	−0.868	−2.671**
	R3	−0.619	1.443	−1.583	−3.313**
	R5	−0.739	1.676*	−1.608	−3.230**
	R7	−0.740	1.953*	−1.562	−3.149**
3	R1	−1.545	−0.187	−2.196**	−3.362**
	R3	−1.222	−0.608	−2.192**	−3.847**
	R5[a]				
	R7[a]				
4	R1	−0.479	0.121	−1.025	−1.963**
	R3	0.500	1.749*	0.141	−0.562
	R5	0.130	1.223	−0.221	−0.891
	R7	1.307	2.856**	1.002	0.332

[a]Sample L-moment ratios violate restricted parameter space conditions (see Section 7.1)

events). Since the PE3 distribution was applied in a previous study on probabilities of 1 d precipitation maxima in the former Czechoslovakia (Šamaj et al. 1982), following a general recommendation of the World Meteorological Organization (WMO), the results published therein (and cited in several recent studies on extreme precipitation events) should be considered unreliable.

6. REGIONAL GROWTH CURVES AND ESTIMATES OF DESIGN VALUES

To estimate parameters of regional distributions we utilized a regional algorithm based on L-moments (Hosking & Wallis 1997). Sample L-moment ratios calculated from the data at different sites were combined to give regional average L-moment ratios; analogously to the method of conventional (product) moments, the first 3 L-moments were used to estimate parameters of a given distribution. Two variants of the scaling factor (index storm) were tested, the mean and median of at-site distributions. The accuracy of estimates was determined using a Monte Carlo simulation (bootstrap resampling; e.g. Park et al. 2001), taking into account the inter-site dependence in terms of correlation matrices (Hosking & Wallis 1997). 10 000 realisations of all regions were made, and the GEV distribution was used to fit the generated data in Regions 1, 2 and 4 while the

Table 3. RMSE for estimated regional growth curves based on the GEV distribution (Regions 1, 2 and 4) and the GLO distribution (Region 3). F: cumulative probability; $x(F)$: quantile function; R1 and R5 denote 1 d and 5 d precipitation amounts respectively

	F	$x(F)$	RMSE
Region 1 (GEV)			
R1	0.500	0.906	0.009
	0.900	1.486	0.010
	0.980	2.149	0.032
	0.990	2.483	0.041
	0.999	3.874	0.080
R5	0.500	0.912	0.011
	0.900	1.456	0.013
	0.980	2.075	0.038
	0.990	2.386	0.049
	0.999	3.672	0.093
Region 2 (GEV)			
R1	0.500	0.916	0.009
	0.900	1.479	0.010
	0.980	2.080	0.030
	0.990	2.370	0.039
	0.999	3.505	0.076
R5	0.500	0.928	0.009
	0.900	1.439	0.013
	0.980	1.962	0.036
	0.990	2.206	0.045
	0.999	3.127	0.083
Region 3 (GLO)			
R1	0.500	0.929	0.018
	0.900	1.398	0.022
	0.980	1.982	0.063
	0.990	2.307	0.083
	0.999	3.881	0.160
R5	0.500	0.878	0.050
	0.900	1.439	0.034
	0.980	2.332	0.135
	0.990	2.908	0.189
	0.999	6.360	0.415
Region 4 (GEV)			
R1	0.500	0.901	0.024
	0.900	1.443	0.024
	0.980	2.130	0.077
	0.990	2.500	0.108
	0.999	4.191	0.222
R5	0.500	0.948	0.012
	0.900	1.327	0.018
	0.980	1.706	0.053
	0.990	1.881	0.071
	0.999	2.528	0.146

GLO distribution was employed in Region 3. The regional relative RMSE of the estimated growth curves are shown in Table 3.

The regional approach considerably reduces between-site variability of the estimates of the shape parameter (k) of the GEV/GLO distribution. While the at-site analysis (estimation of parameters independently site-by-site) leads to estimates of the shape parameter of the GEV distribution for 1 d precipitation amounts in a broad range between –0.37 and +0.16 (yielding different extreme value types with $k < 0$ / $k > 0$), the values

Table 4. Estimates and 95% confidence intervals (CI) of the shape parameter k of the GEV distributions (Regions 1, 2 and 4) and GLO distribution (Region 3) of 1, 3, 5 and 7 d precipitation amounts (R1, R3, R5 and R7, respectively)

	Variable	k	95% CI	
Region 1 (GEV)	R1	–0.151	–0.082	–0.183
	R3	–0.176	–0.096	–0.216
	R5	–0.147	–0.077	–0.190
	R7	–0.129	–0.061	–0.180
Region 2 (GEV)	R1	–0.112	–0.042	–0.145
	R3	–0.117	–0.047	–0.160
	R5	–0.087	–0.011	–0.137
	R7	–0.069	–0.002	–0.124
Region 3 (GLO)	R1	–0.249	–0.158	–0.314
	R3	–0.374	–0.211	–0.514
	R5	–0.379	–0.208	–0.520
	R7	–0.342	–0.192	–0.475
Region 4 (GEV)	R1	–0.210	–0.052	–0.326
	R3	–0.063	0.076	–0.181
	R5	–0.074	0.059	–0.190
	R7	–0.030	0.110	–0.144

of k in the 4 regions lie in a relatively narrow band between –0.21 and –0.11. The differences among estimates at individual sites become even larger for multi-day extremes; again the regional algorithm efficiently lessens the variations although they become more pronounced due to real climatological differences among regions. Regional estimates of k are negative for all regions and durations, and except for multi-day events in Region 4 their 95% confidence intervals (CI) do not include zero (Table 4). A particularly conspicuous deviation in Region 3 appears for distributions of multi-day (3 to 7 d) precipitation amounts, the tails of which are much heavier (corresponding to pronounced negative values of k) compared to other parts of the Czech Republic. While in all other regions the upper tails of regional distributions are heavier for 1 d rather than multi-day events, the opposite pattern is observed in Region 3 (Table 4).

Use of the median as the scaling factor yields underestimated design values (as expected from the generally heavy tails of the fitted distributions), and the mean is preferred as the index storm. The relationship between the index storm and mean annual precipitation is approximately linear in Regions 1, 2 and 3, the slope being largest (smallest) in Region 3 (Region 1). The existence of this relationship enables design values at ungauged sites to be estimated from mean annual precipitation. In Region 4, the dependence is not observed and values of the index storm are almost identical at all stations.

The main benefits of the regional, compared to the at-site, approach are reduced uncertainty in the design value estimates, and reduced between-site variability that stems from random fluctuations. This is true particularly for the large Regions 1 and 2 where differences between design values (e.g. 50 yr return values) do not reflect climatological features, but almost entirely reflect sampling variability. For example, in Region 1 (formed by 32 stations and covering lowland areas in most parts of the Czech Republic), the at-site analysis results in 50 yr return values of 1 d precipitation amounts between 56 and 101 mm, while the regional approach halves the range of estimates (69 to 92 mm) and makes them more directly related to mean precipitation pattern. The reduction of site-by-site differences (that cannot be related to climatological peculiarities and stem from random sampling variability only) occurs in all regions and for all examined durations of precipitation events.

The superiority of the regional approach is conspicuous when probabilities of given precipitation depths are estimated (Table 5). Estimates of return periods of, for example, 1 d amounts exceeding 80 mm based on the at-site analysis are in the order of thousands to millions of years at 10% of sites, although it is easy to reveal (even without any statistical analysis) that real probabilities of such events are much larger in central Europe (the observed 1 d maxima over the period 1961–2000 exceed 80 mm at 63% of examined sites; cf. Šamaj et al. 1983; Štekl et al. 2001). Such an underestimation reflects properties of the particular samples that do not support heavy tails of distributions. The regional procedure leads to estimates of return periods of 1 d totals exceeding 80 mm at all stations in a range of 5 to 137 yr (Table 5). Similar results hold true for the probabilities of multi-day precipitation amounts where the percentage of sites with unrealistically large return periods (based on at-site data only) of 5 d totals exceeding 150 mm is even higher.

Regional growth curves (derived using the GEV distribution in Regions 1, 2 and 4, and GLO in Region 3) for 1 d and multi-day extremes are depicted in Fig. 4; error bounds for the curves are not plotted for the sake of clarity, but they are indicated in Table 3. (In Region 3, the GEV distribution was applied for comparison, too.)

Table 5. Comparison of estimates of return periods (in yr) of selected precipitation depths based on the single-site and regional analysis. R1: 1 d; R5: 5 d precipitation amounts

Station	R1 > 80 mm		R5 > 150 mm	
	at-site	regional	at-site	regional
Region 1				
Stod	45	78	66	155
Manětín	42	90	81	247
Kralovice	41	49	89	159
Žatec	57	105	115	207
Kounov	64	62	42	98
Kestřany	34	46	70	84
Doksany	77	84	133	203
Koleč	24	44	40	100
Praha-Ruzyně	40	49	65	127
Praha-Klementinum	49	82	77	180
Dubá-Panská Ves	>1000000	73	357	127
Brandýs n/Lab	33	40	73	102
Káraný-Nový Vestec	31	48	67	143
Mladá Boleslav	49	62	131	164
Bakov n/Jizerou	45	64	355	170
Semčice	27	49	75	145
Liberec	21	26	40	41
Běstvina	52	32	347	65
Hořiněves	67	54	>1000	143
Vranov n/Dyjí	122	65	>1000	151
Hradec Králové	24	31	88	93
Kuchařovice	128	59	693	166
Choceň	36	46	71	88
Slatina n/Zdobnicí	66	30	66	43
Letovice	212	60	466	143
Brno-Tuřany	119	87	208	189
Lednice	221	83	267	238
Prušánky	133	86	503	205
Štěpánov	>1000	79	>1000	136
Olomouc-Slavonín	111	52	353	119
Šternberk	67	56	>1000	94
Holešov	52	42	71	63
Region 2				
Aš	282	76	>1000000	181
Cheb	169	111	>1000000	454
Tachov	101	103	532	309
Přimda	>10000	137	758	187
Domažlice	105	71	196	116
Kolová-Pila	94	132	917	382
Nýrsko	68	50	40	80
Klatovy	29	44	51	113
Churáňov	13	13	15	13
Holoubkov	44	57	70	136
Závišín-Bělčice	32	64	103	185
Milešovka	55	104	86	276
Husinec	18	24	40	66
Kovářov	66	85	92	193
Střezimíř	65	55	81	149
Tábor	>10000	114	>1000	378
Třeboň	27	44	96	108
Ondřejov	29	45	43	95
Veliš	56	71	125	188
Hranice u N.Hradů	84	53	>1000000	100
Kamenice n/Lipou	>100000	85	>1000	232
Bedřichov	7	8	7	8
Slavonice	73	51	153	137
Kostelní Myslová	53	53	167	156
Havlíčkův Brod	30	46	>1000	132
Velké Meziříčí	>1000	69	383	244
Svratouch	55	32	35	49
Bystřice n/Pern.	53	62	340	284
Region 3				
Červená	>1000	76	46	32
Město Albrechtice	23	37	22	24
Lichnov	44	72	61	45
Melč	59	39	46	26
Skřipov	31	41	22	23
Valašské Meziříčí	16	21	20	21
Hrabyně	57	45	26	27
Klimkovice	122	59	41	36
Ostrava-Mošnov	134	30	30	26
Hať	152	81	45	41
Lučina	23	21	16	17
Raškovice	4	5	4	5
Region 4				
Semily	106	51	242	232
Nová Paka	33	47	>1000	318
Studenec	41	46	168	333
Čistá	38	46	276	305
Hronov	46	34	314	355

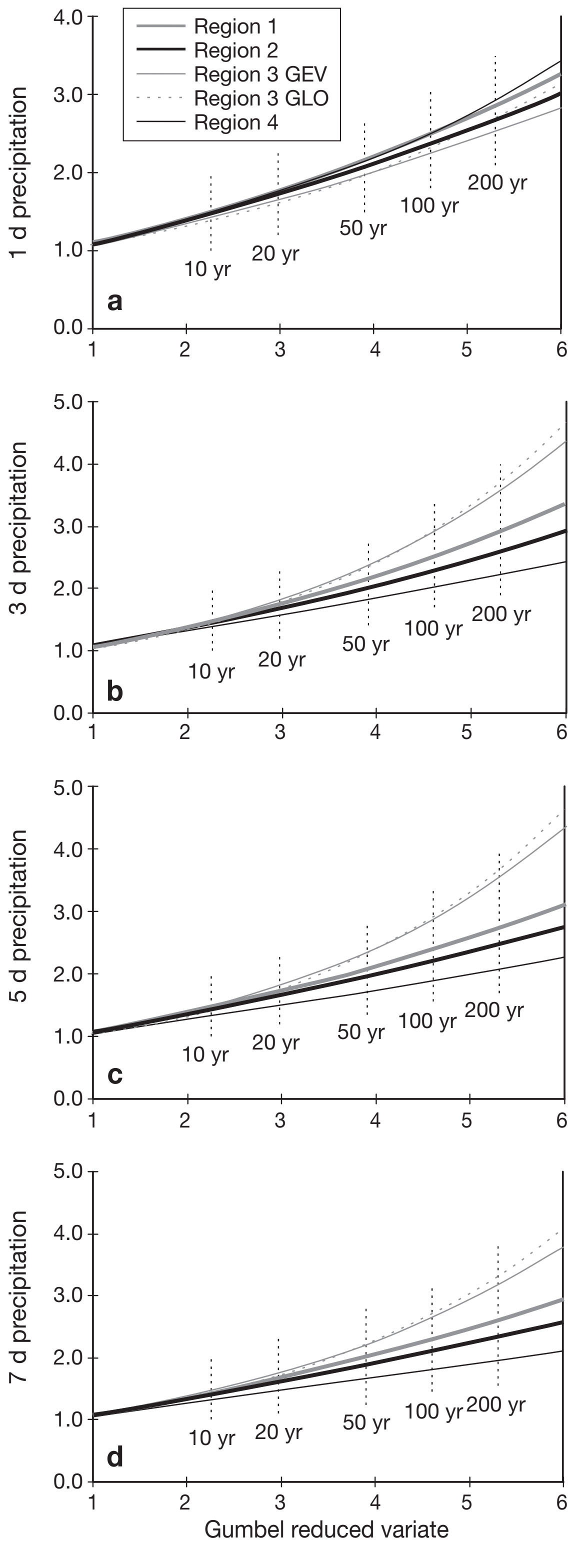

Fig. 4. (a–d) Regional growth curves of 1 to 7 d precipitation amounts based on the GEV distribution; for Region 3, the GLO growth curves are also depicted. Gumbel reduced variate: –ln[–ln(F)], where F is cumulative probability. Values corresponding to return levels of 10, 20, 50, 100 and 200 yr are depicted by vertical lines

Note that the growth curves are dimensionless and station quantiles can be obtained by multiplication with the station's mean annual maximum of a given duration. The between-region differences in the shapes of the growth curves are very similar for 3 to 7 d amounts, and upper tails of the distributions are much heavier in Region 3 compared to other regions. In Region 4, multi-day extremes are almost Gumbel-distributed (corresponding to zero value of the shape parameter k, and a straight line of a growth curve); Region 4 is the only region where the Gumbel distribution is not rejected at the 0.05 level (cf. 95% CI of the estimates of k in Table 4). Deviations between the GLO and GEV distribution in Region 3 are—in the range of return periods useful in a practical implementation of the regional analysis; the 10, 20, 50, 100 and 200 yr return levels are shown by dashed vertical lines—much smaller than differences among regions.

Variations in the shapes of the regional growth curves are relatively minor for 1 d annual maxima (Fig. 4). Particularly, the growth curves in Regions 1 and 4 are not distinct from each other if the accuracy of the estimates is measured by the RMSE, i.e. the curves overlap within RMSE bounds in a large range of return values. For multi-day precipitation amounts, the regional growth curves are distinct from each other at return levels between 20 and 100 yr (cf. Table 3). The regional similarity of growth curves for 1 d maxima was expected since processes leading to heavy 1 d amounts, mostly related to individual storms in summer season, are spatially less variable in the area under study compared to synoptic patterns causing extreme high multi-day totals (e.g. Štekl et al. 2001).

7. DISCUSSION

7.1. Impracticability of tests with the kappa distribution under special conditions

An interesting statistical issue arose in Region 3: the Hosking-Wallis regional homogeneity tests (Section 3.4) as well as the goodness-of-fit test (Section 3.5) were impracticable for 5 and 7 d precipitation amounts because the L-moment ratios estimated from the data were incompatible with any set of parameters of the general 4-parameter kappa distribution (Hosking 1994) used in the simulations.

To enable the 4 parameters to be estimated from the L-moments ratios, the parameter space must be restricted, and certain conditions ensure the existence of the L-moments and the uniqueness of the parameters, given the first 4 L-moments. This restricted parameter space corresponds to the condition $\tau_4 < (1+5\tau_3^2)/6$ that was not satisfied for 5 and 7 d events in Region 3.

After a detailed examination of extreme precipitation data in Region 3 we found that the problem was introduced by the occurrence of one year (1997) with extraordinarily enhanced multi-day precipitation totals (related to a flood episode; record-breaking 5 d amounts in July 1997 led to massive floods, affecting the Odra river catchment in particular; e.g. Řezáčová et al. 2005). When the largest observation at each site in the region was omitted, the Hosking-Wallis tests confirmed the regional homogeneity of the reduced data. The homogeneity of the region according to statistical qualities of precipitation extremes, with all data retained, was also supported by the results of the Lu-Stedinger test and by the sample L-moments ratios which were very similar at all stations in Region 3 (at each site they lie within 95% CI of L-moment ratios at all other sites; CI were constructed using a bootstrap method). Very heavy distribution tails of multi-day extremes may be responsible for the failure of the Hosking-Wallis tests; this point is discussed below.

7.2. Nonparametric test on tail index

As shown in Section 6, all the regional distributions possess heavy tails with $k < 0$. Particularly noteworthy is the heavy tail of multi-day extremes in the NE Region 3, reflected also in the inapplicability of the kappa distribution in the Hosking-Wallis regional homogeneity tests. A nonparametric test on the tail index of Jurečková & Picek (2001; see Section 3.6 and Appendix) was applied to estimate the weight of the right tail of the underlying distribution. The results (Table 6) show that heavy 5 d and 7 d events in this area may be drawn from a distribution with the tail index corresponding to a very heavy tail of the distribution of maxima (tail index $m \leq -1.1$). According to simulation experiments, the results of the test depend only moderately on the chosen values of the parameters n and δ (note that small δ leads to slower convergence to the asymptotic distribution but larger threshold $a_{N,m}$; see Appendix for the explanation of parameters) and all other tested reasonable combinations of n and δ yield the same findings in Region 3.

Table 6. Results of test H_0: $m \leq -1.1$ against H_1: $m > -1.1$; with test parameters $n = 5$, $\delta = 0.25$; and $n = 4$, $\delta = 0.48$ in Region 3. Non-rejected cases at the: *0.05 level; **0.10 level. R1, R3, R5, R7: 1, 3, 5 and 7 d precipitation amounts. For details of the test see Appendix. H_0 is rejected at the 0.05 level in other regions for all variables $[1 - F_N^{\star}(a_{N,m_0}) = 0]$

Variable	$n = 5$, $\delta = 0.25$		$n = 4$, $\delta = 0.48$	
	$1 - F_N^{\star}(a_{N,m_0})$	T_m	$1 - F_N^{\star}(a_{N,m_0})$	T_m
R1	0		>0	13.04
R3	>0	1.659	>0	1.623*
R5	>0	−0.631**	>0	−0.562**
R7	>0	−1.000**	>0	−1.027**

Since the tail index and shape parameter k are identical for the GEV and GLO distributions, the test indicates that shape parameter k may be less than −1 for multi-day events in Region 3. This points to the possibility that extreme precipitation amounts in Region 3 may follow a heavy-tailed distribution for which L-moments do not exist, and the regional procedures must be modified. The test on the tail index has not yet been utilized in climatological literature, and may become a useful tool in inferences concerning statistical models of extreme values.

8. CONCLUSIONS

A regional methodology is shown to be beneficial compared to an at-site approach to the frequency analysis of extreme precipitation events. It leads to estimates of design values that are less uncertain, and spatial variability of which (related mostly to random fluctuations) is favourably reduced. The 4 homogeneous regions (in the Czech Republic) ensuing from statistical procedures (cluster analysis of site characteristics and subsequent tests for regional homogeneity) also reflect climatological differences in precipitation regimes and synoptic patterns that cause heavy precipitation, and their future applications may not be limited to the frequency analysis of rainfall extremes.

The Generalized Extreme Value (GEV) distribution was identified as the most suitable distribution for modelling maximum annual 1 to 7 d precipitation amounts, according to the L-moment ratio diagram and goodness-of-fit tests. This finding is similar to many other parts of the world where the GEV distribution was found useful in modelling precipitation extremes (e.g. Naghavi & Yu 1995, Alila 1999, Kharin & Zwiers 2000, Brath et al. 2001, Smithers & Schulze 2001, Sveinsson et al. 2001, Gellens 2002, Fowler & Kilsby 2003, Semmler & Jacob 2004). Only in Region 3 in the NE (which is most prone to the occurrence of high precipitation totals) is the GLO distribution preferred. Negative regional estimates of shape parameters of both distributions reflect heavy tails of precipitation extremes. Note that the GEV and GLO distributions with a negative value of the shape parameter are distributions with the same weight of the (heavy) upper tail, and their probability density functions converge to zero more slowly than those of other candidate 3-parameter distributions examined (LN3 and PE3). The regional approach considerably lessens the between-site variation of estimates of the shape parameter of the GEV/GLO distribution compared to at-site procedures, and the estimates of design values are more reliable and climatologically consistent in the individual regions. Differences between distributions (in cases when more than one statistical model is appropriate) are generally smaller than differences between the regional and at-site approaches, which further supports the superiority of the regional algorithm.

Considerable differences between the 4 regions in the shapes of the growth curves also indicate that the homogeneous regions are useful and reasonable for modelling probabilities of precipitation extremes. The between-region variability is almost identical for 3 to 7 d totals, and generally larger for multi-day than 1 d events. This is because the latter are related to mechanisms (mostly convective storms in the summer half-year), the spatial variability of which is relatively minor (Štekl et al. 2001). Unlike the 1 d extremes, heavy multi-day precipitation is usually associated with slowly moving cyclones over central Europe, an influence of which tends to be enhanced in the NE part of the Czech Republic.

It is likely that appropriate modifications of the regional algorithm can improve its performance and the reliability of design values (e.g. Sveinsson et al. 2001). Future directions and challenges involve incorporation of peaks-over-threshold methodology and covariates (time-dependency) into regional extreme value models (e.g. Coles & Dixon 1999, Katz et al. 2002) and the development of a region-of-influence approach (Holmes et al. 2002). The issue of time-dependency is particularly opportune since an increase in the frequency and severity of heavy precipitation is expected and/or observed over large parts of Europe (e.g. Booij 2002, Frich et al. 2002, Christensen & Christensen 2004, Pal et al. 2004), and the currently disastrous impacts of high precipitation amounts and floods on human society may become even more pronounced in a future climate. The present analysis constitutes bases for more sophisticated regional models of extremes in a non-stationary climate.

Acknowledgements. The study was supported by the Grant Agency of AS CR under project B300420601. Thanks are due to J. Hošek for assistance in drawing the figures, O. Prosová for help in preparing Section 2.2, and L. Gaál, R. Huth and I. Nemešová for comments on previous versions of the manuscript. A number of valuable suggestions from the editor and 3 anonymous reviewers helped to improve the paper.

LITERATURE CITED

Adamowski K (2000) Regional analysis of annual maximum and partial duration flood data by nonparametric and L-moment methods. J Hydrol 229:219–231

Alila Y (1999) A hierarchical approach for the regionalization of precipitation annual maxima in Canada. J Geophys Res 104:31645–31655

Booij MJ (2002) Extreme daily precipitation in western Europe with climate change at appropriate spatial scales. Int J Climatol 22:69–85

Brath A, Castellarin A, Montanari A (2001) At-site and regional assessment of the possible presence of non-stationarity in extreme rainfall in northern Italy. Phys Chem Earth B 26:705–710

Brázdil R (1998) The history of floods on the rivers Elbe and Vltava in Bohemia. Erfurter Geogr Studien 7:93–108

Brázdil R, Dobrovolný P, Kotyza O (2004) Floods in the Czech Republic during the past millennium. Houille Blanche 5: 50–55

Cheng S, Peng L (2001) Confidence intervals for the tail index. Bernoulli 7:751–760

Christensen OB, Christensen JH (2004) Intensification of extreme European summer precipitation in a warmer climate. Global Planet Change 44:107–117

Coles SG, Dixon MJ (1999) Likelihood-based inference for extreme value models. Extremes 2:5–23

Coufal L, Langová P, Míková T (1992) Meteorological data in the Czech Republic in the period 1961–1990. National Climatic Programme, Czech Hydrometeorological Institute, Prague (in Czech)

Engel H (2004) The flood event 2002 in the Elbe river basin—causes of the flood, its course, statistical assessment and flood damages. Houille Blanche 6:33–36

Fowler HJ, Kilsby CG (2003) A regional frequency analysis of United Kingdom extreme rainfall from 1961 to 2000. Int J Climatol 23:1313–1334

Frich P, Alexander LV, Della-Marta P, Gleason B, Haylock M, Klein Tank AMG, Peterson T (2002) Observed coherent changes in climatic extremes during the second half of the 20th century. Clim Res 19:193–212

Gaál L (2006) Estimation methods of statistical properties of short-term to several-day design precipitation in Slovakia. PhD thesis, Comenius University, Bratislava (in Slovak)

Gellens D (2002) Combining regional approach and data extension procedure for assessing GEV distribution of extreme precipitation in Belgium. J Hydrol 268:113–126

Gottschalk L, Krasovskaia I (2002) L-moment estimation using annual maximum (AM) and peak over threshold (POT) series in regional analysis of flood frequencies. Norw J Geogr 56:179–187

Guttman NB, Hosking JRM, Wallis JR (1993) Regional precipitation quantile values for the continental United States computed from L-moments. J Clim 6:2326–2340

Hanslian D, Brázdil R, Štekl J, Kakos V (2000) The influence of Mediterranean cyclones on high precipitation at Milešovka and Lysá hora in the period 1961–1995. Meteorol Zpr 53:33–41 (in Czech with English summary)

Hasofer AM, Wang Z (1992) A test for extreme value domain of attraction. J Amer Statist Assoc 87:171–177

Hill BM (1975) A simple general approach to inference about the tail of a distribution. Ann Statist 3:1163–1174

Holmes MGR, Young AR, Gustard A, Grew R (2002) A region of influence approach to predicting flow duration curves within ungauged catchments. Hydrol Earth Syst Sci 6:721–731

Hosking JRM (1990) L-moments: Analysis and estimation of distributions using linear combinations of order statistics. J Roy Stat Soc 52B:105–124

Hosking JRM (1994) The four-parameter kappa distribution. IBM J Res Develop 38:251–258

Hosking JRM, Wallis JR (1993) Some statistics useful in regional frequency analysis. Water Resour Res 29:271–281

Hosking JRM, Wallis JR (1997) Regional frequency analysis. An approach based on L-moments. Cambridge University Press, Cambridge

Jurečková J, Picek J (2001) A class of tests on the tail index. Extremes 4:165–183
Jurečková J, Picek J (2004) Estimates of the tail index based on nonparametric tests. In: Hubert M, Pison G, Struyf A, Van Aelst S (eds) Theory and applications of recent robust methods. Birkhauser, Basel, p 141–152
Katz RW, Parlange MB, Naveau P (2002) Statistics of extremes in hydrology. Adv Water Res 25:1287–1304
Kharin VV, Zwiers FW (2000) Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere-ocean GCM. J Clim 13:3760–3788
Kjeldsen TR, Rosbjerg D (2002) Comparison of regional index flood estimation procedures based on the extreme value type I distribution. Stochastic Environ Res Risk Assess 16:358–373
Kjeldsen TR, Smithers JC, Schulze RE (2002) Regional flood frequency analysis in the KwaZulu-Natal province, South Africa, using the index-flood method. J Hydrol 255:194–211
Kundzewicz ZW, Ulbrich U, Brucher T, Graczyk D and 6 others (2005) Summer floods in central Europe—climate change track? Nat Hazards 36:165–189
Květoň V, Tolasz R, Zahradníček J, Stříž M (2002) Precipitation distribution during the floods in the Czech Republic in August 2002. Meteorol Zpr 55:180–187 (in Czech with English summary)
Kyselý J, Picek J, Huth R (2005) Modelling extreme precipitation events in the Czech Republic using the regional approach. In: Michaelides SC (ed) Proc 7th Panhellenic Int Conf on Meteorol, Climatol and Atmos Physics, Vol B. University of Cyprus, Nicosia, p 549–556
Lu LH, Stedinger JR (1992) Sampling variance of normalized GEV/PWM quantile estimators and a regional homogeneity test. J Hydrol 138:223–245
Naghavi B, Yu FX (1995) Regional frequency analysis of extreme precipitation in Louisiana. J Hydraul Eng 121: 819–827
Pal JS, Giorgi F, Bi XQ (2004) Consistency of recent European summer precipitation trends and extremes with future regional climate projections. Geoph Res Lett 31:L13202, doi:10-1029/2004GL019836
Park JS, Jung HS, Kim RS, Oh JH (2001) Modelling summer extreme rainfall over the Korean peninsula using Wakeby distribution. Int J Climatol 21:1371–1384
Pilon PJ, Adamowski K (1992) The value of regional flood frequency analysis using the method of L-moments. Can J Civ Engin 19:137–147
Rezáčová D, Kašpar M, Müller M, Sokol Z, Kakos V, Hanslian D, Pešice P (2005) A comparison of flood precipitation in August 2002 with historical extreme precipitation events from the Czech territory. Atmos Res 77:354–366
Šamaj F, Valovič Š, Brázdil R, Gulčíková V (1982) Maximum daily atmospheric precipitation totals in Czechoslovakia. Meteorol Zpr 35:129–135 (in Slovak with English summary)
Šamaj F, Valovič Š, Brázdil R (1983) Extreme daily precipitation totals in Czechoslovakia. Meteorol Zpr 36:14–21 (in Slovak with English summary)
Semmler T, Jacob D (2004) Modeling extreme precipitation events—a climate change simulation for Europe. Global Planet Change 44:119–127
Smithers JC, Schulze RE (2001) A methodology for the estimation of short duration design storms in South Africa using a regional approach based on L-moments. J Hydrol 241:42–52
Štekl J, Brázdil R, Kakos V, Jež J, Tolasz R, Sokol Z (2001) Extreme daily precipitation on the territory of the Czech Republic in the period 1879–2000 and their synoptic causes. National Climatic Programme of the Czech Republic, 31, Czech Hydrometeorological Institute, Prague (in Czech with English summary)
Sveinsson OGB, Boes DC, Salas JD (2001) Population index flood method for regional frequency analysis. Water Resour Res 37:2733–2748
Tolasz R (ed) (2007) Climate atlas of the Czech Republic. Czech Hydrometeorological Institute and Palack University, Prague and Olomouc
Ulrych TJ, Velis DR, Woodbury AD, Sacchi MD (2000) L-moments and C-moments. Stochastic Environ Res Risk Assess 14:50–68
Werick WJ, Willeke GE, Guttman NB, Hosking JRM, Wallis JR (1993) National drought atlas developed. Geophys News: 8–10

Appendix 1. Nonparametric test on the tail index

The model assumes that distribution function F satisfies $1 - F(x) = x^{1/m}L(x)$, i.e. F has a heavy tail, where $m < 0$ is the parameter of interest (tail index) and $L(x)$ is a function, slowly varying at infinity:

$$\lim_{x\to\infty} \frac{L(tx)}{L(x)} = 1 \text{ for each } t > 0$$

The test of the one-sided hypothesis H_0: $m \le m_0$ against the one-sided alternative H_1: $m > m_0$ is based on the maxima of N sub-samples of equal size n from a given dataset.

The data are partitioned into N samples $X_j = (X_{j1}, \dots, X_{jn})$ of fixed size n, $j = 1, \dots, N$. Denote $X^{(1)}, \dots, X^{(N)}$ the respective sample maxima. Let $F_N^\star$ be the empirical distribution function of the sample maxima, i.e.

$$F_N^\star = \frac{1}{N}\sum_{j=1}^{N} I\left[X^{(j)} \le x\right]$$

where I is the indicator function

$$I[u \le x] = \begin{cases} 1, & \text{if } u \le x \\ 0, & \text{if } u > x \end{cases}$$

For any fixed $m_0 < 0$, put $a_{N,n} = (nN^{1-\delta})^{-m_0}$ where $0 < \delta < 0.5$ is a chosen constant.

The test rejects the hypothesis H_0 at the asymptotic significance level α provided

either $\quad 1 - F_N^\star(a_{N,m_0}) = 0$

or $\quad 1 - F_N^\star(a_{N,m_0}) > 0$ and

$$T_m = N^{\delta/2}\left\{-\log\left[1 - F_N^\star(a_{N,m_0})\right] - (1-\delta)\log(N)\right\} \ge \Phi^{-1}(1-\alpha)$$

where Φ stands for the standard normal distribution function (Jurečková & Picek 2001, 2004)

Editorial responsibility: Bryson Bates, Wembley, Western Australia, Australia

Submitted: April 2, 2006; Accepted: February 2, 2007
Proofs received from author(s): April 3, 2007