ACTA UNIVERSITATIS AGRICULTURAE ET SILVICULTURAE MENDELIANAE BRUNENSIS

Volume 67 26 Number 1, 2019

DETERMINATION OF *N*-YEAR DESIGN PRECIPITATION IN THE CZECH REPUBLIC BY ANNUAL MAXIMUM SERIES METHOD

Silvie Kozlovská[1], Jakub Šácha[2], František Toman[1]

[1]Department of Applied and Landscape Ecology, Faculty of Agronomy, Mendel University in Brno, Zemědělská 1, 613 00 Brno, Czech Republic

[2]Department of Statistics and Operation Analysis, Faculty of Business and Economics, Mendel University in Brno, Zemědělská 1, 613 00 Brno, Czech Republic

To link to this article: https://doi.org/10.11118/actaun201967010275

Received: 13. 2. 2018, Accepted: 12. 11. 2018

To cite this article: KOZLOVSKÁ SILVIE, ŠÁCHA JAKUB, TOMAN FRANTIŠEK. 2019. Determination of *N*-year Design Precipitation in the Czech Republic by Annual Maximum Series Method. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, 67(1): 275–286.

Abstract

The sum of design precipitation of a selected repetition period, provided that it is evenly distributed over the river basin area, is a basic input for the calculation of the direct outflow volume by the curve number method. It is necessary to determine the design precipitation for each location using the statistical methods and the longest available data series on daily precipitation sums, or more specifically their annual maximums. This paper deals with the determination of design precipitation from data of eight stations of the Czech Hydrometeorological Institute for the period 1961–2013. From a series of annual maximum values of daily precipitation sums, N-year design precipitations were calculated using two methods (Gumbel and generalized extreme value distributions). The conformity of both models with empirical distribution of values was statistically tested to evaluate which of the models gave more accurate results. In these cases, it was more appropriate to use the generalized extreme value distribution. Finally, the newly calculated characteristics were compared with the design values used by Šamaj *et al.* (1985), where significant differences were found.

Keywords: Gumbel distribution, generalized extreme value distribution (GEV), curve number method (CN), direct outflow volume, design precipitation, annual maximum series (AMS), repeat time

INTRODUCTION

The process of water erosion occurs and takes place in connection with the outflow of water across the surface of the area and is influenced by a number of factors. According to Holý (1994), the theory of water erosion is focused mainly on the rules of surface and concentrated surface outflow and on the sediment transport processes caused by running water. Erosion phenomena caused by surface-flowing water are the result of complex natural processes. Determining their course, mathematical expression and predicting erosion processes of certain intensity and frequency of occurrence under given conditions is a complex hydrological problem.

The surface outflow is defined as a part of the precipitation water which, after deduction of the vapor losses and retention on the soil surface, flows into the soil surface and outflows over the surface of the area. In order to determine the value of the surface outflow from a slope, it is necessary to know slope parameters (angle, length), soil properties (initial soil moisture, infiltration) and intensity and time course of precipitation (Holý, 1994).

Janeček *et al.* (2012) states that the hydrological details for designing anti-erosion measures (especially technical) should be based on long-term monitored flows in the closure flow profiles. However, such data are rarely available, especially in small river basins. Even if these data are available, they may not always be reliable because there are more or less significant changes in the use of soil in the landscape, so the drainage characteristics of a river basin are constantly changing. In such cases, design values need to be determined by indirect methods based on river basin characteristics and are usually a compromise between practical simplicity and theoretical precision. The most widely used method in the Czech Republic is the US "*The Curve Number Method – CN*" (USDA-SCS, 1985) and it represents a simple precipitation outflow model with fairly easy-to-detect inputs while being sufficiently precise for agricultural river basins of up to 10 km^2. The method makes it possible to determine the volume of the direct outflow and peak flow caused by design torrential rain of chosen probability of occurrence, which are basic inputs for the design of anti-erosion measures. As stated by Janeček (2012), the method can only be used in design practice in accordance with ČSN 75 1400 "Hydrological data of surface water" for the design of technical anti-erosion measures, such as drainage paths of concentrated surface outflow (grassed valleys), ditches, protective barriers and for assessing the effect of anti-soil erosion control measures on the surface outflow. The method cannot be used to calculate the outflow of melting snow.

The basic input for the calculation of the volume of direct outflow by the CN curve method is the precipitation sum of design rain of a selected repeat period provided that it is evenly distributed across a river basin area. This design rain has to be determined for each site using the statistical methods and the longest available data series on daily precipitation sums, or more specifically their annual maxims (or values exceeding the threshold, see e.g. Fusek *et al.* 2016).

Repetition time is the number of years during which an element or characteristic is reached or exceeded on average once. Periodicity is the ratio of the number of times the hydrological element or the characteristic from a set of thresholds have been reached or exceeded to the number of years of its observation. This is actually the inverted value of the repetition time (Kohnová, 2005).

The volume of precipitation is further converted to the volume of outflow by the curve numbers – CN. Their values are then dependent on the hydrological properties of soils, the vegetation cover, the size of impermeable surfaces, interceptions and surface retention (Janeček *et al.*, 2012).

Kemel (1996) draws attention to the fact that relatively short series of observations do not allow to create a reliable empirical overrun line in areas of small and large percentages of overruns. Therefore, the theoretical overrun curve is used – the empirical line is aligned by it in its central part, and it is possible to extrapolate it to the area of extreme values that did not occur due to a short period of observation and it serves for design purposes (anti-erosion and anti-flood control measures).

The first important step in analyzing the number of phenomena is to decide on the method by which sampling data should be obtained from the quantity of measured data. Kohnová *et al.* (2005) present three possible models: Annual Maximum Series (AMS), Partial Duration Series (PDS), and Peaks-Over-Threshold (POT).

The Annual Maximum Series includes the only and the highest observed value in each year. This is the most commonly used sample compilation method, although it may lose some information – for example, the second or third largest value in a given year may be greater than the maximum value in another year. Such cases are excluded in the hydrological Peaks-Over-Threshold series (PDS, POT) because all data greater than the first set threshold are selected for subsequent statistical processing. The threshold is usually selected so low that the sample contains at least one value per year. The automated method for threshold selection is described in detail in Holešovský *et al.* (2016).

The mutual coherence between the AMS or PDS (POT) series has been investigated by several authors, e.g. Madsen *et al.* (1997), who compared methods based on AMS and POT and concluded that the POT series are generally better in a local and regional analysis of extraordinary total precipitation. On the other hand, however, the theoretical advantages of the POT series are compensated by a more complex mathematical apparatus that requires appropriate processing techniques. These methods have a problem with the correct choice of threshold, as described in detail e.g. in Holešovský, Fusek and Michálek (2015). In case of very short time series of observed values it is preferred to use the POT method supplemented by the bootstrapping methods (described e.g. in Holešovský *et al.* (2016) etc.).

Kohnová *et al.* (2005) summarize several theoretical distributions used for the analysis of extreme precipitation, from two-parameter it is e.g. exponential, Gumbel and log-normal distribution. From three-parameter distributions there is used the generalized Pareto (GP) distribution, generalized extreme value (GEV) distribution, generalized logistic (GLO) distribution, three-parameter lognormal (LN3) distribution, gamma distribution, Weibull distribution. There are also multi-parameter distributions, such as the four-parameter distribution of Kappa or the five-parameter Wakeby distribution. Distributions with at least three parameters give less biased estimates.

According to Brázdil (2007), the empirical distribution of measured values is approximated by the theoretical distribution with given parameters that determine its specific character when calculating the *N*-year precipitation (i.e. the precipitation values that are reached or exceeded on average once per *N*-years). Theoretical distributions are also called parametric because their specific attributes depend on the numerical values of their parameters. Distribution parameters can be calculated using different methods (e.g. maximum likelihood, moments, L moments) that have their strengths and weaknesses, for example as solved by Kyselý (2005). The distribution function of the theoretical distribution tells the probability of occurrence of a phenomenon. It is then possible to calculate the period of repetition *N* of the given phenomenon.

In this work, estimates of design *N*-year precipitation were made using two types of theoretical distributions – Gumbel and generalized extreme value (GEV) distribution. Gumbel's distribution was chosen for the purpose of comparing design values with previous publications (Šamaj *et al.*, 1985). The GEV distribution is then selected with respect to abundant use in practice and also because it is implemented in the software ProClimDB (Štěpánek, 2010) and EVDest 1.0 (Fusek and Holešovský, 2014).

Gumbel distribution

The probability density function (pdf) of the Gumbel distribution has the form (Castillo, 1988):

$$f(x)=\frac{\exp\left(\frac{-(x-\alpha)}{\beta}\right)\exp\left[-\exp\left(\frac{-(x-\alpha)}{\beta}\right)\right]}{\beta} \qquad (1)$$

The function has two parameters: location parameter α and scale parameter β.

The cumulative distribution function (cdf) of this distribution is calculated from the relation:

$$F(x)=\exp\left[-\exp\left(\frac{-(x-\alpha)}{\beta}\right)\right] \qquad (2)$$

for $-\infty < x < \infty$

Generalized extreme value distribution (GEV)

The generalized extreme value distribution is a combination of three theoretical distributions – Gumbel, Fréchet and Weibull. It is a generalization of the previous case and besides parameters α, β it also introduces a shape parameter *k* (Jenkinson, 1955, cited according to Brázdil, 2007).

The probability density function (pdf) of the GEV distribution has the form:

$$f(x)=\frac{1}{\beta}\left[1+k\left(\frac{x-\alpha}{\beta}\right)\right]^{-\frac{1}{k}-1}\times\exp\left\{-\left[1+k\left(\frac{x-\alpha}{\beta}\right)\right]^{-1/k}\right\}, \qquad (3)$$

for $\alpha,k\in R,\ \beta>0$ and $1+k\left(\frac{x-\alpha}{\beta}\right)>0.$

The cumulative distribution function (cdf) of the GEV distribution can be then expressed as follows:

$$F(x)=\exp\left\{-\left[1+k\left(\frac{x-\alpha}{\beta}\right)\right]^{-1/k}\right\}, \qquad (4)$$

for $\alpha,k\in R,\ \beta>0$ and $1+k\left(\frac{x-\alpha}{\beta}\right)>0.$

According to Brázdil *et al.* (2007), the advantage of the three-parameter GEV distribution is the possibility to overlay the theoretical curve of the Weibull distribution by empirical values in cases where the parameter $k<0$ and the theoretical values of the Gumbel distribution significantly underestimate the measured values. Conversely, if $k > 0$, the Gumbel distribution will produce extremely overestimated *N*-year estimates. In this situation, the theoretical curve of the Fréchet distribution can be overlaid by empirical values. In the case of the analysis of maximum values (e.g. maximum daily or monthly precipitation totals), the parameter k influences the shape of the distribution function, especially at the intervals of the values occurring with the least probability, i.e. with the longest period of repetition *N*. The cumulative distribution function has, in individual types of general distribution of extreme values, the following form:

Type I. Gumbel distribution $k = 0$ … see relation (2)

Type II. Fréchet distribution $k > 0$

$$F(x) = 0 \quad \text{for } x \le \alpha$$

$$F(x) = \exp\left[-\left(\frac{x-\alpha}{\beta}\right)^{-k}\right] \quad \text{for } x > \alpha \tag{5}$$

Type III. Weibull distribution $k < 0$

$$F(x) = \exp\left[-\left(\frac{-x-\alpha}{\beta}\right)^{-k}\right] \quad \text{for } x < \alpha$$

$$F(x) = 1 \quad \text{for } x \le \alpha \tag{6}$$

For the quantile X_N with the repetition period N, the cumulative probability is given by the relation:

$$F(X_N) = 1-(1/N) \quad \text{for } N > 1 \tag{7}$$

The repetition time N expresses the degree of significance of a design variable and it indicates the number of years during which an element or a characteristic is achieved or exceeded on average once (Kohnová, 2005).

Expressing x from equation (4), resp. (2) and substituting (7) by $F(x)$, we obtain the relation (8), resp. (9) for the determination of the maximum precipitation sums with the repetition period N years:

$$X_N = \alpha + \frac{\beta}{k}\left\{1-\left[\ln\left(1-\frac{1}{N}\right)\right]^{k}\right\} \quad \text{for } k \ne 0 \tag{8}$$

and

$$X_N = \alpha - \beta \ln\left[-\ln\left(1-\frac{1}{N}\right)\right] \quad \text{for } k = 0 \tag{9}$$

In the Czechoslovak Socialist Republic, the issue of maximum *N*-year design precipitation was for the first time more complexly solved by Šamaj *et al.* (1982; 1983; 1985) and Kulasová (1985). Šamaj *et al.* (1985) chose the Gumbel distribution and gamma (Pearson III type) distribution for the estimation of the maximum *N*-year precipitation. In practice, design precipitation from this publication, calculated according to the Gumbel distribution, is used.

Sevruk and Geiger (1981) state that a 40 to 50-year time series is usually sufficient for frequency analysis of extreme precipitation. This is confirmed by Šamaj *et al.* (1985), where the comparison of results of variously lengthy periods with a reference value of 80 years confirmed the eligibility of the 50-year period as the shortest to achieve the acceptable accuracy of calculations of maximum daily precipitation totals. However, the authors also point out that 50 years of measuring was not kept continuously in their research. The missing data were not replaced by interpolation or other methods due to the large spatial variability of maximum daily precipitation totals.

In case of very short time series it is also possible to estimate the maximum *N*-year precipitation, but using different methods (the POT method supplemented by the bootstrapping methods), as shown in Holešovský *et al.* (2016), Fusek *et al.* (2016) etc.

From other authors, for example Gaál *et al.* (2004) were devoted to the selection of an appropriate theoretical distribution of *N*-year precipitation totals. In this case, the gamma distribution was chosen for the frequency analysis of extreme sums of precipitation in Slovakia and the Gumbel distribution was rejected as problematic (GEV distribution was not among the tested).

Halásová *et al.* (2007) researched the appropriate distribution of design maximum *N*-year precipitation at several stations in the Giant Mountains and concluded that the Gumbel distribution tends to underestimate the design quantiles at the tested stations, especially in the case of long repetition periods, and that the GEV distribution seem to be the most suitable.

In the case of design *N*-year precipitation, their dependence on altitude is also tested, but it is not usually proven (e.g. Hostýnek *et al.*, 1999), especially at longer times of repetition and for summer months (in winter there is a higher proportion of frontal precipitation, which correlates more with altitude).

Papalexiou and Koutsoyiannis (2013) studied the issue of a suitable distribution, analyzing more than 15,000 data from around the world on maximum daily precipitation, and they found that long-term records are best described by the Fréchet or GEV distribution.

MATERIALS AND METHODS

Source data

Based on the analysis of the availability of the longest time series of daily precipitation totals with the lowest number of missing sections in the original measurements within the scope of the Brno Branch of the Czech Hydrometeorological Institute (CHMI), eight stations were selected. For these stations, there are available daily precipitation data from 1961 to 2013, i.e. 53 years of continuous measurement.

The following stations were selected: Brno Tuřany (BTUR), Dačice (DACI), Holešov (HOLE), Kostelní Myslová (KMYS), Kuchařovice (KUCH), Náměšť nad Oslavou (NAMO), Strážnice (STRZ) and Velké Meziříčí (VMEZ).

Analytical procedure

The task of analyzing the frequency of occurrence of extreme precipitation totals is to provide an estimate of the amount of precipitation total of a certain duration that may occur or be exceeded at a particular station or in a given territory with a given probability or repetition time. The calculated precipitation amount can then be used in the design practice, e.g. when calculating the volume of direct outflow using the CN method.

1) Creation of annual maximum series (AMS) for all the stations
2) Determination of *N*-year design precipitation for all the stations by two methods: by the GEV distribution and the Gumbel distribution.
   Specialized software was used in the calculations – ProClimDB (Štěpánek, 2010) and EVDest 1.0 (Fusek and Holešovský, 2014). EVDest software contains only estimates of GEV distribution, which is preferred in modern hydrology. The software ProClimDB uses the method of weighted moments to calculate the theoretical distribution parameters while EVDest software apply maximum likelihood (ML) method.

3) Conformity testing of model and empirical distribution of values
   The two nonparametric goodness-of-fit tests were used to evaluate the conformity of model distributions with real data– the Kolmogorov-Smirnov test (KS test) and the Anderson-Darling test (AD test). The zero hypothesis, in both cases, is that the set of empirical values (annual maximum precipitation totals) comes from the base set with a known theoretical distribution – the GEV or Gumbel distribution. The Anderson-Darling test is a modification of the K-S test, but it places more emphasis on the tails of distribution and it is more recommended in modern hydrology (e.g. Holešovský and Fusek, 2016). This test is only implemented in EVDest 1.0 software.

4) Comparison of model distributions
   The KS test p-values and information criteria were used to compare the quality of the three estimates of model distributions (GEV and Gumbel distribution were obtained using the ProClimDB software, another GEV was obtained using the EVDest software). The p-value of the KS test can be understood as a degree of consistency of the model distribution with data, and therefore we will prefer a higher value. From the information criteria, the Akaike information criterion (AIC) and the Bayes information criterion (BIC) are used in this text. These criteria are based on the likelihood function, their advantage, compared to the likelihood function, is that they consider the simplicity of the distribution in terms of the number of estimated parameters. It should be noted that for these criteria, a lower value means a better match.

5) Comparison of the calculated design values with the data of Šamaj *et al.* (1985)
   In some stations in the publication by Šamaj *et al.* (1985) the location did not match with the current state, so the results are compared with the nearest station found – Brno Husovice was selected as the closest station for Brno Tuřany, Telč for Kostelní Myslová and Znojmo for Kuchařovice.

RESULTS

Based on the set of 53 maximum annual precipitation sums for each station, ProClimDB calculated precipitation design values with a return period of 2–5–10–20–50–100 years in two ways – using the Gumbel and GEV distribution. For comparison, the design values of precipitation were also calculated in EVDest 1.0. Thus, two sets of design precipitations were created for each station, and it was necessary to evaluate which of the series shows more reliable estimates of these precipitations. It was achieved by the p-values of the KS test and the information criteria in Tab. I.

All p-values of KS and AD tests are clearly higher than the significance level 0.05. Differences between empirical data and model distributions can not be considered statistically significant for all types of distribution used and for all stations and model distributions can be considered appropriate.

Differences in results for GEV distribution (see Tab. I and II) are probably due to the use of dissimilar parameter estimation method by ProClimDB than by EVDest 1.0. These differences, however, are minimal, so this study will focus on comparing the Gumbel and GEV distributions (according to ProClimDB).

In the comparison of the GEV and Gumbel distribution, there is little controversy. For most stations, we can say that in terms of the p-value of the KS test, the GEV distribution is better. On the contrary, in terms of information criteria, Gumbel distribution is better. The reason is that the parameter k of the GEV distribution is close to zero, both distributions are similar, and the values of the likelihood function, on which the information criteria are based, are similar. Then, the number of parameters that speaks in favor of the Gumbel distribution is decisive.

Due to the low value of the k parameter in the GEV distribution, the Gumbel distribution fits better (based on all criteria) at KUCH, NAMO and STRZ stations. Both distributions practically merge. Tab. II shows that differences in design precipitation values are negligible (less than 3 mm for 100 – year precipitation).

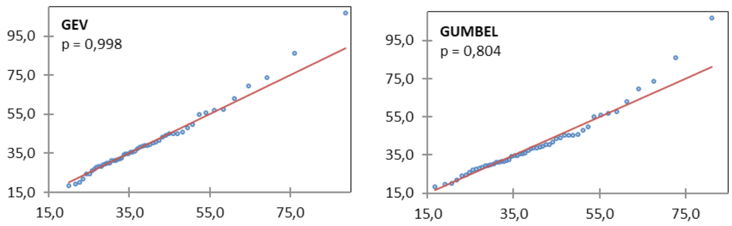
For other stations, at least for some of the criteria used, the GEV distribution is better off. Although the differences in both density functions (Fig. 2) are not so pronounced at first sight, it is noticeable in *N*-year design rainfall, especially for high *N* (50 or 100 years).

In general, it is possible to recommend a more versatile GEV distribution, which is preferred in current publications on this topic, e.g. Holešovský and Fusek (2016).

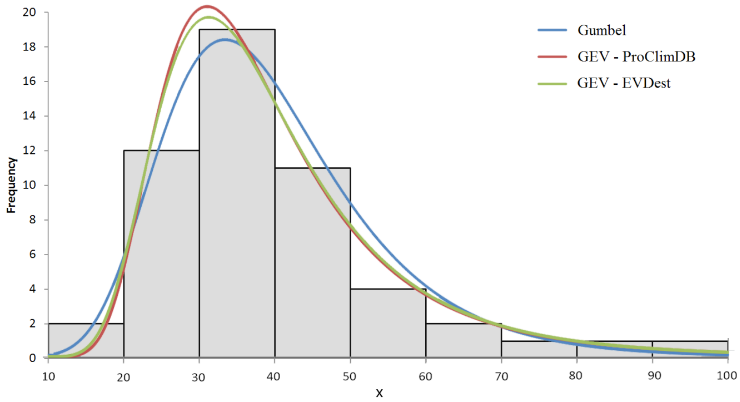
Fig. 1 shows an example of q-q plots, i.e. the comparison of the consistency of the theoretical model and the empirical values of annual maxima for Dačice station using the GEV distribution method (left graph) and Gumbel distribution (right graph). Dačice station was selected for illustration because there is a more pronounced difference between fit distributions, which is manifested in particular by the longer tail in the GEV distribution. This is also visible on the histogram supplemented of the probability density function of the estimated distributions (Fig. 2). Units in Fig. 1 are not stated, but percentages are on both axes.

I: *Comparison of the tested distributions at the tested stations*

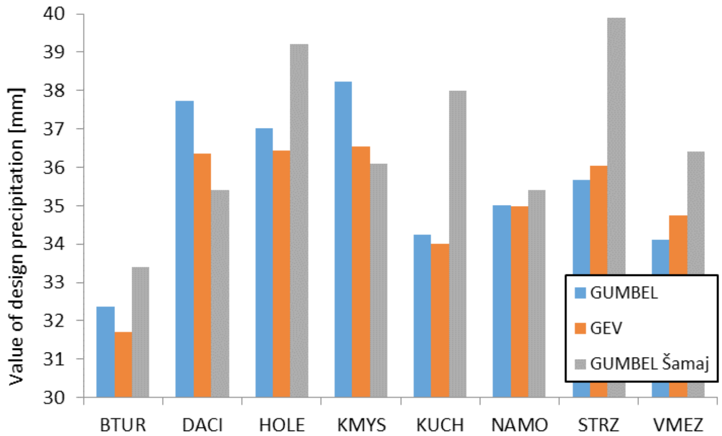
Distribution	p-value, criterion	Station							
		Brno Tuřany (BTUR)	**Dačice (DACI)**	**Holešov (HOLE)**	**Kostelní Myslová (KMYS)**	**Kuchařovice (KUCH)**	**Náměšť nad Oslavou (NAMO)**	**Strážnice (STRZ)**	**Velké Meziříčí (VMEZ)**
GEV ProClimDB	**KS test**	0.944	**0.998**	**0.968**	**0.942**	0.888	0.954	0.923	**0.607**
	AD test	–	–	–	–	–	–	–	–
	AIC	397.94	430.01	398.74	434.73	425.75	402.98	400.69	418.20
	BIC	403.67	435.75	404.48	440.47	431.48	408.72	406.42	423.94
GEV EVDest 1.0	**KS test**	**0.961**	0.989	0.895	0.933	0.813	0.848	0.861	0.599
	AD test	0,977	0.996	0.898	0.955	0.904	0.966	0.966	0.572
	AIC	397.64	**429.90**	398.18	**434.72**	425.73	402.84	400.64	418.09
	BIC	403.37	435.64	403.92	440.45	431.47	408.57	406.37	423.83
Gumbel ProClimDB	**KS test**	0.931	0.804	0.814	0.552	**0.911**	**0.959**	**0.975**	0.605
	AD test	–	–	–	–	–	–	–	–
	AIC	**396.45**	430.14	**396.38**	436.35	**423.90**	**400.97**	**399.04**	**416.30**
	BIC	**400.28**	**433.97**	**400.20**	**440.17**	**427.72**	**404.79**	**402.86**	**420.12**



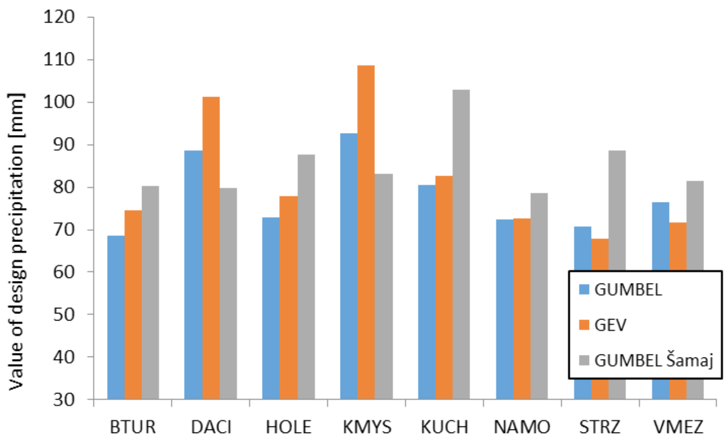
1: *Q-q plots for the station Dačice (GEV and Gumbel distribution)*



2: *Histogram of annual maxima for the station Dačice (GEV and Gumbel distribution)*



3: *Comparison of the value of two-year design precipitation calculated by two methods and according to Šamaj et al. (1985)*



4: *Comparison of the value of 100 – year design precipitation calculated by two methods and according to Šamaj et al. (1985)*

The values of design precipitation with different repetition times calculated by two methods (the Gumbel and GEV distribution) are shown in Tab. II. Part of the Tab. is the design precipitation according to Šamaj *et al.* (1985), which was created by the analysis of data from the years 1901–1980 (the length of observation was different at the stations, and in the case of missing observation years the authors did not add these values). Šamaj *et al.* (1985) calculated these design values using the Gumbel distribution. Thus, in the stated Tab. the difference in the value of design precipitation using the two different methods and older or actual input data can be compared (up

II: *Design N-year precipitation – comparison of the results according to the used methods and with data by Šamaj et al. (1985)*

Station	Method	Max. daily precipitation with the probability of repetition per N years [mm]					
		2	5	10	20	50	100
Brno Tuřany BTUR	Gumbel – ProClimDB	32.4	42.1	48.5	54.6	62.6	68.6
	GEV – ProClimDB	31.7	41.3	48.4	55.7	66.0	74.5
	GEV – EVDest	31.7	41.6	48.5	55.5	65.1	72.8
Brno Husovice	Gumbel (Šamaj *et al.*, 1985)	33.4	45.8	53.8	62.1	72.3	80.2
Dačice DACI	Gumbel – ProClimDB	37.7	51.3	60.3	69.0	80.1	88.5
	GEV – ProClimDB	36.4	49.7	59.9	71.0	87.2	101.1
	GEV – EVDest	36.5	50.0	60.1	70.8	86.4	99.4
Dačice	Gumbel (Šamaj *et al.*, 1985)	35.4	47.1	54.6	62.5	72.1	79.7
Holešov HOLE	Gumbel – ProClimDB	37.0	46.6	53.0	59.1	67.0	72.9
	GEV – ProClimDB	36.4	46.0	52.9	60.0	69.9	77.9
	GEV – EVDest	36.7	46.3	52.9	59.5	68.2	75.0
Holešov	Gumbel (Šamaj *et al.*, 1985)	39.2	52.1	60.3	68.9	79.5	87.7
Náměšť nad Oslavou NAMO	Gumbel – ProClimDB	35.0	45.0	51.6	58.0	66.2	72.4
	GEV – ProClimDB	35.0	45.0	51.6	58.0	66.3	72.6
	GEV – EVDest	34.7	44.7	51.6	58.5	67.6	74.7
Náměšť nad Oslavou	Gumbel (Šamaj *et al.*, 1985)	35.4	46.9	54.3	61.9	71.4	78.7
Strážnice STRZ	Gumbel – ProClimDB	35.7	45.0	51.2	57.2	64.9	70.7
	GEV – ProClimDB	36.0	45.4	51.2	56.6	63.2	68.0
	GEV – EVDest	36.0	45.6	51.6	57.2	63.9	68.9
Strážnice	Gumbel (Šamaj *et al.*, 1985)	39.9	52.8	61.1	69.8	80.4	88.7
Kostelní Myslová KMYS	Gumbel – ProClimDB	38.2	52.8	62.5	71.7	83.7	92.7
	GEV – ProClimDB	36.6	50.7	61.9	74.1	92.6	108.6
	GEV – EVDest	36.6	50.8	62.0	74.1	92.3	108.1
Telč	Gumbel (Šamaj *et al.*, 1985)	36.1	48.6	56.6	64.9	75.2	83.2
Velké Meziříčí VMEZ	Gumbel – ProClimDB	34.1	45.5	53.0	60.2	69.6	76.6
	GEV – ProClimDB	34.8	46.0	52.9	59.1	66.5	71.7
	GEV – EVDest	34.5	45.5	52.3	58.5	66.1	71.4
Velké Meziříčí	Gumbel (Šamaj et al, 1985)	36.4	48.4	56	64	73.9	81.5
Kuchařovice KUCH	Gumbel – ProClimDB	34.2	46.7	54.9	62.7	72.9	80.6
	GEV – ProClimDB	34.0	46.4	54.9	63.2	74.2	82.6
	GEV – EVDest	33.9	46.3	54.9	63.4	74.8	83.7
Znojmo	Gumbel (Šamaj *et al.*, 1985)	38	55.2	66.2	77.7	91.7	102.8

Explanatory notes: Šamaj *et al.* (1985) from the data for the period 1901–1980, others from the data 1961–2013 to 4 mm for 2-year design precipitation and up to 25 mm for 100-year design precipitation).

The graphical comparison of design precipitation differences for two selected repeat times (2 years and 100 years) is shown in Fig. 3 and Fig. 4. For each station, there is design precipitation in three variants (GEV, Gumbel, and Gumbel by Šamaj *et al.*, 1985). The design values calculated by the EVDest software are not included in the charts, as they almost match the ProClimDB outputs. In the two-year precipitation, it is possible to trace larger differences between the values by Šamaj and those calculated at HOLE, KUCH and STRZ stations. Interestingly, the ratio of the difference between the GEV and Gumbel method changed in most stations between two-year and 100-year precipitation. This phenomenon is most noticeable at DACI and KMYS stations. Only NAMO station almost coincide with both methods. It can be deduced that in most of the cases, the GEV distribution at low times of repetition shows lower estimates of design precipitation, and at high repetition times, higher precipitation estimates than the Gumbel distribution (except STRZ and VMEZ stations where the opposite is the case).

Recommended AMS procedure for determining *N*-year design precipitation

Based on the obtained results and the study of related expert texts, several key points were drawn up, ideally leading to the most accurate estimates of design *N*-year precipitation.

1) Determine the maximum annual precipitation for the longest possible (available) period for the required climatological station with at least 50 years of continuous precipitation measurement, use a technical dataset (modified by homogenization and data quality control). This period should be as up-to-date as possible.
2) Select a suitable method of estimating the theoretical distribution parameters (the maximum likelihood method, method of moments, L-moment method etc.).
3) Select at least two theoretical distributions of extreme values, overlay them with empirical data and test the consistency of the theoretical and empirical data distribution by statistical methods (Fusek *et al.* (2016) recommends QQ plot, Kolmogorov-Smirnov test, Anderson-Darling goodness-of-fit test, Chi-square test, etc.). Selected distributions should not be missing the GEV distribution, as it is generally considered to be very appropriate for this type of data – but it may not be the most appropriate for each location.
4) Select a model distribution best suited to empirical data and determine design precipitations for the required repetition times based on it.

DISCUSSION

The values of design precipitation according to Šamaj *et al.* (1985) were recommended for use as a relevant hydrological basis for calculations of anti-erosion measures in the previous version of the methodology by Janeček *et al.* (2007). In the currently valid methodology (Janeček *et al.*, 2012), it is stated that the value of the design rainfall for the site under investigation should be requested from the Czech Hydrometeorological Institute, but that the sum of the maximum 24-hour precipitation with the design frequency of occurrence according to Šamaj *et al.* (1985) can be used as a framework. In order to simplify and save money, these old data are used in many cases even today, and as it can be seen in the analyzes carried out, they are often quite different from those calculated from the current data. The Czech Hydrometeorological Institute uses the current data, which are processed in the ProClimDB program using the GEV method.

From a statistical point of view, the problem of estimating design precipitation is often the shortness of the time series of observed values. From this perspective, it is preferable to use the POT method supplemented by the bootstrapping methods described e.g. in Holešovský *et al.* (2016). In this article, however, we are working with a 53-year time series for which the AMS estimates can be considered as good enough.

The use of design precipitation calculated for individual isolated stations produces a certain degree of inaccuracies for the entire adjacent region because of the large spatial variability and local limitations, especially in torrential precipitation. Various authors have attempted to reduce this deficiency using statistical methods, for example, the regional frequency analysis. Kyselý and Picek (2007) dealt with the issue for the territory of the Czech Republic, and they divided the territory of the state into four relatively homogeneous regions (in terms of location but also differences in climatic conditions and synoptic phenomena causing torrential rainfall). Within the regions, they dealt with the regionalization of design precipitation estimates and tested four types of divisions. In most cases, the GEV distribution was the most appropriate. The LN3 distribution did not show better results than the GEV distribution, only the GLO distribution (Generalized Logistic Distribution) showed smaller variations for one region. The authors strongly refused to use the gamma distribution, which showed very poor conformity of the model with the measured values in the tests.

The regionalization, according to the authors, will not only allow designing of precipitation even outside the precipitation measurement site, but it also regulates the amount of design precipitation towards smaller fluctuations at the stations themselves, which, according to the authors, seems appropriate especially for 50-year and multi-year estimates, which in the calculation for some stations reach too high values.

For the comparison of the design precipitation values calculated in this work with the results of Kyselý and Picek (2007), there is a table of selected stations used in this work and tested by the cited authors. In Tab. III, there are shown repetition periods for rain of a certain amount (over 80 mm) determined by the GEV distribution method. The evaluation period of own data is 13 years longer (1961–2013) than the results in two other columns (data for the period 1961–2000) according to Kyselý and Picek (2007), which must be taken into account.

The calculation of parameters of given theoretical distribution is possible by several methods. The software ProClimDB uses the method of weighted moments, Holešovský and Fusek (2016) recommend the maximum likelihood (ML) method, Kyselý (2005) considers the L-moment method as more suitable.

The ideal would be to individually assess each location, determine the design precipitation from the latest data and the longest possible time series of daily maxima, and test at least two methods (theoretical distributions), parameters of which would be proposed by the most proven method. The further refinement would probably be achieved by regionalization (taking altitude, orography, etc. into account) with the possibility of deducting the design precipitation at any site of the processed area. The implementation of regionalization of design precipitation was not the goal of this work due to the small number of stations and their large mutual distances.

III: *Repetition times of one-day design precipitation with a sum of over 80 mm (in years)*

Station	Repeat time of rain > 80 mm (years)		
	For station (Kozlovská, GEV)	For station (Kyselý, Picek, 2007)	For regionalization (Kyselý, Picek, 2007)
Brno-Tuřany	143	119	87
Holešov	114	52	42
Kostelní Myslová	25	53	53
Kuchařovice	76	128	59
Velké Meziříčí	308	> 1,000	69

CONCLUSION

The presented analysis was carried out with the data of eight stations of the Czech Hydrometeorological Institute (CHMI). These are: Brno Tuřany (BTUR), Dačice (DACI), Holešov (HOLE), Kostelní Myslová (KMYS), Kuchařovice (KUCH), Náměšť nad Oslavou (NAMO), Strážnice and Velké Meziříčí (VMEZ). Daily precipitation data from 1961 to 2013, i.e. 53 years of continuous measurement, were tested.

The calculations of design *N*-year precipitation were performed from these data, which are also used for the calculation of the volume of direct outflow using the CN method. Two methods have been used to do this, which are used most frequently for extreme calculations. One of the methods was the statistical determination by the Gumbel distribution, which was used mainly in the past (see Šamaj *et al.*, 1985, etc.). The second method of statistical evaluation was the Generalized Extreme Value (GEV), which is currently preferred.

The conformity of both models with the empirical distribution of values was statistically tested (Kolmogorov Smirnov conformity test, Anderson Darling test, information criteria), evaluating which of the models gave more accurate results. In most cases, it was more appropriate to use the GEV distribution.

The conclusion of the analysis was the comparison of the resulting design values with the recommended and used estimates of design precipitation according to Šamaj *et al.* (1985). Significant differences were found (up to 4 mm for 2-year design precipitation and up to 25 mm for hundred-year design precipitation).

Based on these findings, it was recommended to always use the actual data obtained with thorough statistical analysis. It would be ideal to assess each site individually, to determine design precipitation from the latest data and the longest possible time series of daily maximums (at least 50 years) and to test several suitable theoretical distributions (e.g. GEV), parameters of which would be estimated by the most appropriate method (e.g. ML method, L-moment method and bootstrap-based method of estimation for parameter *k* of GEV distribution). The values of design precipitation of specific stations provided by the CHMI for financial charge should be sufficiently accurate. However, it is not possible to agree with the recommendation reported by Janeček *et al.* (2012) that the sum of the maximum 24-hour precipitation with the design frequency of occurrence according to Šamaj *et al.* (1985) could be used as a framework.

Acknowledgements

This paper was supported by the Technology Agency of the Czech Republic [no. 2014TA04020886: “New technology solutions for protection against flooding from heavy rainfall.”].

REFERENCES

BRÁZDIL, R. and KIRCHNER, K. 2007. *Selected natural extremes and their impacts in Moravia and Silesia* [in Czech: *Vybrané přírodní extrémy a jejich dopady na Moravě a ve Slezsku*]. Brno: Masaryk university.

CASTILLO, E. 1988. *Extreme value theory in engineering*. Academic Press Series in Statistical Modeling and Decision Science. San Diego, CA: Academic Press.

ÚNMZ. 2014. *Hydrological data of surface waters* [in Czech: *Hydrologické údaje povrchových vod*]. ČSN 75 1400. Praha: Úřad pro technickou normalizaci, metrologii a státní zkušebnictví.

FUSEK, M. and HOLEŠOVSKÝ, J. 2014. *EVDest; EVDest 1.0.* [Software]. Brno: UMAT FEKT VUT. Available at: http://www.umat.feec.vutbr.cz/software/evdest.html

FUSEK, M., HELLEBRAND, R. and MICHÁLEK, J. 2016. Modelling precipitation extremes in the Czech Republic. Update of intensity-duration – frequency curves. *Estonian Journal of Earth Sciences,* 65(4): 234–247.

GAÁL, L., LAPIN, M. and FAŠKO, P. 2004. Maximum multi-day precipitation totals in Slovakia [in Slovak: Maximálne viacdenné úhrny zrážok na Slovensku]. In: ROŽNOVSKÝ, J. and LITSCHMANN, T. (Eds.). *Extremes of weather and climate* [in Czech: *Extrémy počasí a podnebí].* ČHMÚ.

HALÁSOVÁ, O., HANČAROVÁ, E. and VAŠKOVÁ, I. 2007. Time and space variability of selected climatological and hydrological elements in the Giant Mountains during the period 1961-2000 [in Czech: Časová a prostorová variabilita vybraných klimatologických a hydrologických prvků na území Krkonoš za období 1961–2000]. – In: ŠTURSA, J. and KNAPÍK, R. (Eds.). *Geoecological problems of the Giant Mountains* [in Czech: *Geoekologické problémy Krkonoš*]. Int. Sci. Conf. Proc., říjen 2006, Svoboda n. Úpou.

HOLEŠOVSKÝ, J., FUSEK, M. and MICHÁLEK, J. 2015. Modelling of precipitation extremes using parametric and nonparametric methods with automated threshold selection. *International Journal of Mathematics and Computers in Simulation*, 9(1): 94–102.

HOLEŠOVSKÝ, J., FUSEK, M., BLACHUT, V. and MICHÁLEK, J. 2016. Comparison of precipitation extremes estimation using parametric and nonparametric methods. *Hydrological Sciences Journal – Journal Des Sciences Hydrologiques*, 61(13): 2376–2386.

HOLEŠOVSKÝ, J. and FUSEK, M. 2016. *Methods for extreme value analysis and their software implementation* [in Czech: *Metody analýzy extrémních hodnot a jejich softwarová implementace*]. Informační bulletin České statistické společnosti, 27(3): 1–13.

HOLÝ, M. *et al.* 1994. *Erosion and environment* [in Czech: *Eroze a životní prostředí*]. Prague: ČVUT.

HOSTÝNEK, J., LEPKA, Z. and SOSNA, V. 1999. *Processing of N-year annual and monthly maximums of daily precipitation totals in Western Bohemia* [in Czech: *Zpracování N-letých ročních a měsíčních maxim denních úhrnů srážek v západních Čechách*]. Meteorological Bulletin, 52: 73–77.

JANEČEK, M. *et al.* 2007. *Protection of agricultural land from erosion.* [in Czech: *Ochrana zemědělské půdy před erozí*]. Methodology. 1st Edition. Prague: VÚMOP.

JANEČEK, M. *et al.* 2012. *Protection of agricultural land from erosion* [in Czech: *Ochrana zemědělské půdy před erozí]. Methodology*. Prague: Czech University of Life Scienses.

JENKINSON, A. F. 1955. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society, 81(348): 158 – 171.*

Kemel, M. 1996. *Climatology, meteorology, hydrology* [in Czech: *Klimatologie, meteorologie, hydrologie*]. Prague: CTU Publishing House.

KOHNOVÁ, S., GAÁL, L., SZOLGAY, J. and HLAVČOVÁ, K. 2005. *Analysis of the maximum precipitation totals in the upper Hron basin* [in Slovak: *Analýza maximálnych úhrnov zrážok v povodí horného Hrona*]. Bratislava: STU Publishing House.

KULASOVÁ, B. *et al.* 1985. *Processing of N-year precipitation totals* [in Czech: *Zpracování N-letých srážkových úhrnů*]. Prague: CHMI.

KYSELÝ, J. 2005. Estimation of parameters of extreme values distribution by the L-moment method [in Czech: Odhad parametrů rozdělení extrémních hodnot metodou L momentů]. *Meteorological Bulletin,* 58(6): 161–167.

KYSELÝ, J. and PICEK, J. 2007. Regional growth curves and improved design value estimates of extreme precipitation events in the Czech Republic. *Climate Research*, 33: 243–255.

MADSEN, H., RASMUSSEN, P. F. and ROSBJERG, D. 1997. Comparison of Annual Maximum Series and Partialduration Series Methods for Modeling Extreme Hydrologic Events: 1. At-site Modeling. *Water resources Research,* 33(4): 747–746.

PAPALEXIOU, S. M. and KOUTSOYIANNIS, D. 2013. Battle of extreme value distributions: A global survey on extreme daily rainfall. *Water Resources Research*, 49: 187-201.

SEVRUK, B. and GEIGER, H. 1981. *Selection of Distribution Types for Extremes of Precipitation*. Operational Hydrology Report No. 15 (WMO – No. 560), Geneva: World Meteorological Organization.

ŠAMAJ, F., VALOVIČ, Š. and BRÁZDIL, R. 1983. Extreme daily totals of atmospheric precipitation in the CSSR [in Slovak: Extrémne denné úhrny atmosferických zrážok v ČSSR]. *Meteorological Bulletin*, 36(1): 14–21.

ŠAMAJ, F., VALOVIČ, Š. and BRÁZDIL, R., 1985. Daily precipitation totals with extraordinary abundance in the CSSR in the period 1901-1980 [in Slovak: Denné úhrny zrážok s mimoriadnou výdatnosťou v ČSSR v období 1901–1980]. In: *Zborník prác HMÚ 24.* Alfa. Bratislava, pp. 9–113.

ŠAMAJ, F., VALOVIČ, Š., BRÁZDIL, R. and GULČÍKOVÁ, V. 1982. Maximum daily totals of atmospheric precipitation in the CSSR [in Slovak: Maximálne denné úhrny atmosferických zrážok v ČSSR]. *Meteorological Bulletin,* 35(5): 129–135.

ŠTĚPÁNEK, P. *et al.* 2010. *ProClimDB – software for processing climatological datasets.* [Software]. Brno: ČHMI. Available at: http://www.climahom.eu/ProcData.html

USDA-SCS. 1985. *National Engineering Handbook, Section 4 - Hydrology.* Washington, D. C.: USDA-SCS.

Contact information

Silvie Kozlovská: silvie.kozlovska@mendelu.cz
Jakub Šácha: jakub.sacha@mendelu.cz
František Toman: frantisek.toman@mendelu.cz