Return values of 60-minute extreme rainfall for Hungary

Mónika Lakatos*, Beatrix Izsák, Olivér Szentes, Lilla Hoffmann, Andrea Kircsi, and Zita Bihari

Hungarian Meteorological Service
Kitabel P. s 1, H-1024, Budapest, Hungary

*Corresponding Author e-mail: lakatos.m@met.hu

(Manuscript received in final form February 27, 2020)

Abstract—The rainfall intensity for various return periods are commonly used for hydrological design. In this study, we focus on rare, short-term, 60-minute precipitation extremes and related return values which are one of the relevant durations in the planning and operating demands of drainage and sewerage systems in Hungary. Time series of 60-minute yearly maxima were analyzed at 96 meteorological stations. To estimate the return values for a given return period, the General Extreme Value (GEV) distribution was fit to the yearly maxima. The GEV fit and also the Gumbel fit (GEV Type I.) were tested. According to the goodness of fit test results, both GEV and Gumbel distributions, are adequate choices. The return values for 2, 4, 5, 10, 20, and 50 year return periods are illustrated on maps, and together with their 95% confidence intervals, are listed in tables for selected stations. The maps of return values demonstrate that the spatial patterns of the return values are similar, although the enhancing effect of orography can be explored in the Transdanubia region and in the North Hungarian Range. As the return period is increasing, so the range of the confidence are widening as it is expected.

Key-words: extreme precipitation, rainfall intensity, GEV, return level

1. Introduction

In recent decades, an increasing number of studies have reported about the presence of significant positive trends in precipitation extremes in Europe (e.g., Alexander et al., 2006; Klein Tank and Konnen, 2003; Moberg et al., 2006; Zolina et al., 2009). The observed more frequent heavy precipitation events are also consistent with increasing amounts of water vapor in the atmosphere due to global warming (Allen and Ingram, 2002; Willett et al., 2008). The warming climate induces increasing frequency of extreme precipitation in some region. Significant trends in precipitation extremes over Europe have been found since the middle of the 20th
century in earlier regional studies for Northern Europe (Groisman et al., 2005), the UK (Fowler and Kilsby, 2003), the Mediterranean region (Pujol et al., 2007), and western and eastern parts of the Czech Republic (Kysely, 2009) and Poland (Lupikasza et al., 2010). E.J.M. van den Besselaar and co-authors (Besselaar et al., 2013) showed that despite the considerable decadal variability, 5-, 10-, and 20-year events of 1-day and 5-day precipitation amounts for the first 20 years in the period 1951-2010 became more common in the analyzed 60 years for the daily precipitation series from the European Climate Assessment and Dataset (ECA&D, http://www.ecad.eu) project (Klein Tank et al., 2002; Klok and Klein Tank, 2008).

Considering the precipitation changes in Hungary, the decrease of the yearly precipitation sum is not remarkable, it is 3% from 1901 to 2019. From the beginning of the 1990s, precipitation has been increasing both on annual and seasonal scales, however, this rise is not significant. Recent years have been dominated by extremes. The magnitude of the change in precipitation intensity (mm/day) is about 1.3 mm/day in the countrywide average. The number of days with daily sum above 20 mm increased by 1 day in the period 1901-2019.

Estimation of precipitation extremes are essential for the planning of important infrastructure, such as water control systems, reservoirs, dams, and urban runoff. The rainfall intensity for various return periods are commonly used for hydrological design (Hailegeorgis, et al., 2013). The Hungarian Meteorological Service provides climate services on return periods of short-term precipitation (10 min, 20 min, 30 min, 60 min, 180 min) to the authorities responsible for roads, railways, hydrological and urban planning. The return period is the average time between the occurrences of extremes of a specified size. The return period of extremely high values of short-term rainfall has shortened in recent years in Hungary, as it is shown in a case study for the Pécs-Pogány meteorological station (Lakatos and Hoffmann, 2019). Lakatos and Hoffmann (2019) also published the intensity-duration-frequency (IDF) curves for the Pécs-Pogány station. The IDF curves are commonly used in planning to describe the return period associated with a given rainfall event. Earlier studies of the short-term precipitation intensity covered the period 1967–1990, when the rainfall was registered by ombrographs (Váradi and Nemes, 1992). The return values (or design values) for different return period were estimated by Váradi and Nemes for 26 stations in Hungary. Gayer and Ligetvári (2006) refer to results of Váradi and Nemes as an exemplary work in the municipal water management planning in Hungary. High intensity, short-term showers, typically 1 hour or less, occasionally up to 3 hours has the greatest impact on the urban environment (Gayer and Ligetvári, 2006). Due to the recent and projected climate change, existing design criteria for infrastructure should be revised as it is pointed out in Varga et al., 2016.

In this study, we focus on rare, short-term, hourly, namely 60-minute precipitation extremes and related return values which are relevant in the planning and operating demands of drainage and sewerage systems in Hungary.
2. Data

The focus area is Hungary. Automatic weather stations replaced the ombrographs in many places in Hungary, particularly from the late 1990s. As a result of automatization, 22-year-long 10 minutes data series are available to analyze the behavior of the sub-daily precipitation for Hungary from 1998. The 10-minute sums of precipitation are stored in the digital meteorological database of the Hungarian Meteorological Service. The 10-minute data was used to derive all the 60-minute rainfall sums which are used in this study. Then the yearly maxima of the 60-minute rainfall were computed for each station to apply extreme value analysis. Time series of yearly maxima were analyzed at 96 meteorological stations in this paper (Fig. 1). The analysis performed here covers almost the complete automatic weather station network in Hungary. Stations with lot of missing data and with shorter time series were excluded from this study. Time series of the yearly maxima have been quality controlled, and their temporal homogeneity were checked using the MASH (Szentimrey, 1999) method before applying extreme value analysis. The MASH method is an internationally well-known homogenization procedure, one of the best performing methods according to the benchmarking test executed in the COST (European Cooperation in Science and Technology) Action ES060: advances in homogenization methods of climate series: an integrated approach (HOME) (Venema et al., 2012). Erroneous data and inhomogeneities may severely affect the extremes. Therefore, great care must be taken during the decision making on inclusion or exclusion of the erroneous data, as sometimes it is difficult to distinguish the extremes and outliers.

Fig. 1. Automatic weather stations used in this study. Additional statistics are shown for the labeled stations.
3. Methods

The objective of our analyses is to estimate the amount of rainfall falling at a given point for a fixed duration (60 minutes in our case) and a given return period. To estimate the return values for a given return period, the extreme value theory and relating statistical methods have to be applied. Extremes can be found in the tails of the probability distribution. An introduction to extreme value theory can be found in many publications, for example in Coles (2001) and Katz (2002).

In recent years, two statistical approaches are frequently used in modeling extreme rainfall events: the model of annual maxima (block maxima) and the peaks over threshold (POT) model (Coles, 2001). The most common approach in modeling extremes involves fitting a statistical model to the annual extremes in a time series of data. The classical POT model defines a threshold and considers all events with intensity higher than this threshold. The block maxima are usually modeled by a Gumbel or a generalized extreme value (GEV) distribution (Katz et al., 2002), while the POT is modeled by the generalized Pareto distribution (Katz et al., 2002; Coles, 2001). Despite its advantages, the POT model is much less applied in the analysis of hydrological extremes (Madsen et al., 1997). The World Meteorological Organization (WMO) report entitled “Statistical Distributions for Flood Frequency Analysis” (Cunnane, 1989) provides a review of probability distributions of extreme values and methods for estimation of their parameters. Other WMO source in this topic is the “Guidelines on Analysis of extremes in a changing climate in support of informed decisions for adaptation” (Klein Tank, 2009).

The GEV distribution was introduced by Jenkinson (Jenkinson, 1955). It describes the three types of extreme value distributions for block maxima of any variable (Coles, 2001). The distribution of the block maxima converges to a GEV distribution $G(x)$ while the record length approaches infinity. The three-parameter GEV distribution can be defined in the form

$$G(x; \mu, \sigma, \xi) = \exp \left\{ - \left[ 1 + \xi \left( \frac{x-\mu}{\sigma} \right) \right]^{-1/\xi} \right\},$$

where $\mu$ is the location, $\sigma$ is the scale, and $\xi$ is the shape parameter. Note that $\sigma$ and $1 + \xi(x-\mu)/\sigma$ must be greater than zero. Depending on $\xi$, the $G(x)$ converges to one of three types: Type I (Gumbel) ($\xi = 0$), Type II (Fréchet) ($\xi > 0$), and Type III (Weibull) ($\xi < 0$). Thereby the shape parameter determines whether the fitted distribution will have a finite lower bound, a finite upper bound, or no bound. In the unbounded case, the shape parameter has value zero, and the GEV distribution becomes the well-known Gumbel distribution (Gumbel, 1958), which have been used extensively in hydrology, meteorology, and engineering. The Gumbel/Type 1 distribution have been applied in hydrology to model floods
and extreme rainfalls (Chow et al., 1988; Stedinger et al., 1993). According to several studies, extreme 24-hour precipitations follow Type II distribution (heavy upper tail; $\xi > 0$) (Wilks, 1993; Koutsoyiannis and Baloutsos, 2000; Katz et al. 2002; Coles and Pericchi, 2003; Coles et al., 2003; Koutsoyiannis, 2004). Fréchet distribution represents the lowest risk for technological architecture, as design values are higher than for Type I (Gumbel) and Type III (Weibull).

From the fitted extreme value distribution, we can estimate the return value which is defined as a value that is expected to be equaled or exceeded on average once in every time interval $T$ (return period), with a probability of $p = 1/T$. Therefore, the return values ($z$) can be calculated from the Eq.(1) by inverting the GEV distribution in the case of a stationary climate (Matyasovszky, 2002):

$$z_{Type I.} = -\sigma \ln(-\ln p) + \mu$$
$$z_{Type II.} = \sigma (-\ln p)^{-\xi} + \mu$$
$$z_{Type III.} = -\sigma (-\ln p)^{-\xi} + \mu.$$

The POT method was applied by Lakatos and Matyasovszky for 10-minute precipitations measured at Baja meteorological station. Pareto distribution was fitted to 10-minute sums of rainfalls above a specified threshold to estimate the distribution of extreme values (Lakatos and Matyasovszky, 2004).

In accordance with the recent practice of the Hungarian Meteorological Service, the GEV distribution is fitted to the time series of annual maxima/minima. The data are stored in the digital meteorological database, and functions to estimate the different return values are implemented and operated within the digital database. These functions are based on the procedure that was published by Tibor Faragó (Faragó, 1989; Faragó et al., 1989; Faragó and Katz, 1990). The return values pertaining to predefined return periods (2, 4, 5, 10, 20, 50, 100, and 200 years) are computed and displayed for various meteorological elements, e.g., for the short-term precipitation sum (10-, 20-, 30-, 60-, and 180-minute).

In this paper GEV distribution was fit to yearly maximum series of 60-minute rainfall measured at 96 automatic weather stations (AWS) in the period of 1998–2019 first. Moreover, the Gumbel fit was tested. The second step was the computation of the return values for 2, 4, 5, 10, 20, 50 years together with their 95% confidence intervals. Maximum-likelihood method was used to estimate $\mu$ location, $\sigma$ scale, and $\xi$ shape parameters of the GEV distribution (Prescott and Walden, 1980). The distribution fit and the computation of the return values and the 95% confidence intervals were executed by applying the R statistical software (R Core Team, 2012). The return values are illustrated on maps. The results of the distribution fit and the bounds of the confidence intervals are presented in tables for 10 selected meteorological stations.
4. Results and discussion

Return value can be interpreted as the value that is expected to be equal or exceeded on average once in every return period \((T)\), or with probability \(p=1/T\) in any given year. The \(T=2, 4, 5, 10, 20,\) and 50-year return values of the 60 minutes rainfall can be seen on the maps in Figs. 2-7 for Hungary. The GEV distribution were fitted to the 60-minute yearly maxima at the 96 measuring sites, and the return values were then derived. Return values calculated for return periods frequently used in engineering practice are calculated on the maps. For instance, the intensity of a 60-minute event which would be exceeded once every 2 years, on average, is shown in Fig. 2. The maps of return values demonstrate that the longer the return period the greater the pertaining return value. The spatial patterns of the return values are similar on each map. The enhancing effect of orography can be explored in the Transdanubia region and in the North Hungarian Range, although it is not pronounced. It often occurs that hills and mountains enhance the moving air masses, and the largest intensity can be experienced afterwards under the mountains. On the maps greater values appear also in plain regions, typically in the southeast. The reason for this pattern is that short-duration high-intensity rainfalls happen mainly through local convective cells, which have similar physical properties less respective of geographical location. Occasionally, intensive rainstorms evolve during fast moving cold fronts, and sometimes the storms structured in squall-line and cause extreme rainfall in a very short time interval.

Fig. 2. 2-year return values of the 60-minute rainfall.
Fig. 3. 4-year return values of the 60-minute rainfall.

Fig. 4. 5-year return values of the 60-minute rainfall.

Fig. 5. 10-year return values of the 60-minute rainfall.
Over the years, the GEV distribution has become a widely used model in extreme value analysis. Although it is essential to draw attention to the fact that the GEV distribution is simply a model, our observational series and associated statistics do not precisely follow the theory. The GEV fit procedure resulted in the $\mu$ (location), $\sigma$ (scale), and $\xi$ (shape) parameters of the asymptotic probability distribution function. Depending on the value of the $\xi$ parameter, the extremes will converge to the Gumbel ($\xi=0$), Fréchet ($\xi > 0$), or Weibull ($\xi < 0$) types. According to our analysis, the $\xi$ parameter is under -0.1, at about one fourth (27%) of the stations resulting the upper bound, and $\xi$ is above 0.1 at 42% of the stations resulting the lower bound in the extreme value distribution. Table 1 contains these three parameters ($\mu$, $\sigma$, $\xi$) for 10 selected stations from the 96 in total, equally covering the territory of Hungary. Possibly the $\xi$ parameter varies according to the dominating precipitation systems and orographic effects. The
larger positive $\xi$ values appear at stations located chiefly in the western and southern part of Hungary which is under the influence of the Atlantic and Mediterranean climate respectively. The largest $\xi$ value appears at observation stations operating in rural areas of the capital.

Table 1. The $\mu$ (location), $\sigma$ (scale), and $\xi$ (shape) parameters of the GEV distribution and the p-values that are the outputs of the goodness of fit tests for GEV ($p_{value\_KS}$) and Gumbel ($p_{value\_C}$ and $p_{value\_A}$) distribution for selected stations

<table>
<thead>
<tr>
<th></th>
<th>$\mu$</th>
<th>$\sigma$</th>
<th>$\xi$</th>
<th>$p_{value_KS}$</th>
<th>$p_{value_C}$</th>
<th>$p_{value_A}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Szombathely</td>
<td>18.44</td>
<td>5.67</td>
<td>0.15</td>
<td>0.934</td>
<td>0.458</td>
<td>0.322</td>
</tr>
<tr>
<td>Nagykanizsa</td>
<td>19.29</td>
<td>4.08</td>
<td>0.28</td>
<td>0.988</td>
<td>0.272</td>
<td>0.198</td>
</tr>
<tr>
<td>Győr-Likócs</td>
<td>16.56</td>
<td>3.02</td>
<td>0.53</td>
<td>0.998</td>
<td>0.270</td>
<td>0.207</td>
</tr>
<tr>
<td>Siófok</td>
<td>20.33</td>
<td>6.60</td>
<td>-0.21</td>
<td>0.890</td>
<td>0.804</td>
<td>0.695</td>
</tr>
<tr>
<td>Paks</td>
<td>21.08</td>
<td>6.77</td>
<td>-0.37</td>
<td>0.934</td>
<td>0.717</td>
<td>0.524</td>
</tr>
<tr>
<td>Budapest-Lőrinc</td>
<td>19.12</td>
<td>8.00</td>
<td>0.44</td>
<td>0.944</td>
<td>0.237</td>
<td>0.166</td>
</tr>
<tr>
<td>Baja</td>
<td>21.13</td>
<td>8.53</td>
<td>0.02</td>
<td>0.756</td>
<td>0.169</td>
<td>0.146</td>
</tr>
<tr>
<td>Miskolc</td>
<td>20.04</td>
<td>7.01</td>
<td>0.06</td>
<td>0.861</td>
<td>0.754</td>
<td>0.645</td>
</tr>
<tr>
<td>Szeged</td>
<td>19.69</td>
<td>9.28</td>
<td>-0.06</td>
<td>0.931</td>
<td>0.836</td>
<td>0.718</td>
</tr>
<tr>
<td>Debrecen</td>
<td>20.22</td>
<td>6.67</td>
<td>0.14</td>
<td>0.871</td>
<td>0.697</td>
<td>0.704</td>
</tr>
</tbody>
</table>

It is necessary to check if the GEV model fits to the series. Various tests can be used for this purpose. We applied the Kolmogorov–Smirnov test which is a nonparametric test to compare our samples with the GEV probability distribution (Stephens, 1970). The Gumbel distribution is frequently used in hydrological applications to estimate the extremes. Therefore, the Type I. (Gumbel) distribution has been checked too. The Cramer-von Mises and Anderson-Darling tests for Gumbel distribution function proposed by Chen and Balakrishnan (1995) was applied to check the Gumbel fit (Anderson and Darling, 1954; Stephens, 1986; Marsaglia, 2004). The null hypothesis is that the GEV is an appropriate model in the case of Kolmogorov-Smirnov test, and the Gumbel model is appropriate in the case of Cramer-von Mises and Anderson-Darling tests. We reject the null hypothesis at level $\alpha$ if the p-value is smaller than $\alpha$ (usually 0.05 or 0.01), otherwise we fail to reject the null hypothesis at level $\alpha$. The p-values resulted in the goodness of fit tests are listed in Table 1. for 10 selected stations. The higher p-values represents greater strength of evidence in support of the null hypothesis. The $p_{value\_KS}$ (Kolmogorov-Smirnov test) describes the measure of the evidence of the GEV fit, while the $p_{value\_C}$ (Cramer-von Mises test) and the $p_{value\_A}$ (Anderson-Darling test) describe the evidence of the Gumbel fit.
Although smaller sizes of data record may hide the appropriateness of GEV distribution, in our case it fits adequately to the 22-year-long records for each station. The Gumbel model was considered suitable for all station at all reasonable significance level using Cramer-von Mises and Anderson-Darling tests.

Having the parameters of the GEV distribution, the return values and related 95% confidence intervals were computed. The 2-year, 5-year, 10-year, 20-year, and 50-year return levels together with the lower bound and upper bound of the interval where the return level lies in with $p=0.95$ probability can be found in Table 2. As the time interval is increasing, the range of the confidence are widening. For the 5-year return period (which is the usual time interval for designing the sewerage system in urban environment) the estimated return values are between 23.5 mm and 36.1 mm, but the lowest upper bound of the 95% confidence intervals is 28.7 mm (Table 2). In the case of the 50-year return period, the largest return values turned out to be at Győr-Likócs and Budapest-Lőrinc stations, where the $\xi$ is a large positive value (see Table 1.). The confidence interval is extremely expanded in the case of the 50-year return level at Budapest-Lőrinc, which is a rural station in the capital. The return levels for very long return periods tends to enlarge the error due to inaccurate estimates of the shape parameters which describe the tails of a distribution. Generally, the confidence of the return levels decreases rapidly, when the return period is about two times longer than the length of the data series (Klein Tank et al., 2009).

Table 2. The return values of the 60-minute rainfall with the lower bound and the upper bound of the 95% confidence interval (mm)

<table>
<thead>
<tr>
<th>Location</th>
<th>2-year</th>
<th>5-year</th>
<th>10-year</th>
<th>20-year</th>
<th>50-year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Szombathely</td>
<td>17.5</td>
<td>20.6</td>
<td>23.7</td>
<td>26.6</td>
<td>26.1</td>
</tr>
<tr>
<td>Nagykanizsa</td>
<td>18.5</td>
<td>20.9</td>
<td>23.2</td>
<td>24.0</td>
<td>24.5</td>
</tr>
<tr>
<td>Győr-Likócs</td>
<td>15.7</td>
<td>17.8</td>
<td>19.8</td>
<td>18.1</td>
<td>13.9</td>
</tr>
<tr>
<td>Siófok</td>
<td>19.3</td>
<td>22.7</td>
<td>26.1</td>
<td>27.8</td>
<td>29.2</td>
</tr>
<tr>
<td>Paks</td>
<td>19.7</td>
<td>23.4</td>
<td>27.1</td>
<td>28.4</td>
<td>29.3</td>
</tr>
<tr>
<td>Budapest-Lőrinc</td>
<td>17.1</td>
<td>22.3</td>
<td>27.5</td>
<td>25.4</td>
<td>19.6</td>
</tr>
<tr>
<td>Baja</td>
<td>19.8</td>
<td>24.3</td>
<td>28.7</td>
<td>31.4</td>
<td>33.7</td>
</tr>
<tr>
<td>Miskolc</td>
<td>18.5</td>
<td>22.6</td>
<td>26.8</td>
<td>28.0</td>
<td>28.4</td>
</tr>
<tr>
<td>Szeged</td>
<td>17.8</td>
<td>23.0</td>
<td>28.2</td>
<td>30.3</td>
<td>31.5</td>
</tr>
<tr>
<td>Debrecen</td>
<td>18.7</td>
<td>22.7</td>
<td>26.8</td>
<td>27.6</td>
<td>27.4</td>
</tr>
</tbody>
</table>
The return levels, or design values are often expressed in terms of rainfall intensity (mm/h) rather than rainfall depth (mm) over a certain duration in construction engineering. The latter has been chosen here, because this is what is actually measured at the meteorological stations.

5. Conclusion

This study focused on the short-term precipitation that is essential for hydrological planning. Hourly, namely 60-minute maximum precipitation series were analyzed from 96 stations operated by the Hungarian Metrological Service. The return values are widely used parameters in planning, for example in construction of drainage systems. The return value can be determined as a value that is expected to occur on average once during the return period. To give an estimation for return values, the general extreme value distribution fitting is an adequate procedure. The 60-minute yearly maxima follow the GEV distribution and also the Gumbel fit suits to the sample series from all 96 stations. The spatial patterns of the return values for 2, 4, 5, 10, 20, 50 years coming from the GEV fit are similar. The influence of the orographic effect turns up in the Transdanubia region and in the territory of the North Hungarian Range, although greater values appear also in the plain regions in the southeastern part of Hungary. The 95% confidence intervals were computed to illustrate the uncertainty of the rainfall estimates to various return periods. Naturally, as the return period is increasing the range of the confidence are widening.

The methodology introduced here can be applied in the future for the renewing of the existing design criteria for infrastructure. The uncertainty can be decreased with using longer data series for estimation of the parameters of the GEV distribution. The series of automatic measurements can be lengthened for about a dozen of stations by the measurements rescued from the ombrograph registering papers after eliminating the inhomogeneity caused by the different sampling. Estimation of areal precipitation return levels by applying regional GEV distribution (Stedinger et al., 1993) is one of the possible directions of the continuation of our examinations. The methodology for estimating extreme areal precipitation by shifting the station-based precipitation to areal precipitation from the grid-based (Dyrrdal et al., 2016) also an option to consider.

References


Gayer, J. and Ligetvári, F., 2006: Települési vízgazdálkodás csapadékvíz elhelyezés, VITUKI, Budapest, 76–90. (in Hungarian)


Hailegeorgis, T.T., Thorolfsson, S.T., and Alfredsen, K., 2013: Regional frequency analysis of extreme precipitation with consideration of uncertainties to update IDF curves for the city of Trondheim. J. Hydrol. 498, 305–318. https://doi.org/10.10115/JCLI3339.1


Matyasovszky I., 2002: Statisztikus klimatológia. ELTE Eötvös Kiadó, Budapest. (In Hungarian)


Varga, L., Buzás, K. és Honti, M., 2016: Új csapadékmáximo-függvények. *Időjárás* 96, 64–69. (in Hungarian)


