Hydrol. Earth Syst. Sci. Discuss., 4, 2361–2401, 2007
www.hydrol-earth-syst-sci-discuss.net/4/2361/2007/

Hydrology and
Earth System
Sciences
Discussions

Region-of-influence approach to a frequency analysis of heavy precipitation in Slovakia

L. Gaál[1,3], **J. Kyselý**[2], **and J. Szolgay**[3]

[1]Radar Department, Czech Hydrometeorological Institute, Na Šabatce 17, 143 06 Prague 4, Czech Republic
[2]Institute of Atmospheric Physics, Academy of Sciences of the Czech Republic, Boční II 1401, 141 31 Prague 4, Czech Republic
[3]Department of Land and Water Resources Management, Faculty of Civil Engineering, Slovak University of Technology, Radlinského 11, 813 68 Bratislava, Slovakia

Received: 10 July 2007 – Accepted: 13 July 2007 – Published: 24 July 2007

Correspondence to: L. Gaál (ladislav.gaal@chmi.cz)

Abstract

The L-moment-based regionalization approach developed by Hosking and Wallis (1997) is a frequently used tool in regional frequency modeling of heavy precipitation events. The method consists of the delineation of homogeneous pooling groups with a fixed structure, which may, however, lead to undesirable step-like changes in growth curves and design value estimates in the case of a transition from one pooling group to another. Unlike the standard methodology, the region-of-influence (ROI) approach does not make use of groups of sites (regions) with a fixed structure; instead, each site has its own "region", i.e. a group of sites that are sufficiently similar to the site of interest. The aim of the study is to develop a version of the ROI approach, which was originally proposed in order to overcome inconsistencies involved in flood frequency analysis, for the modeling of probabilities of heavy precipitation amounts. Various settings of the distance metric and pooled weighting factors are evaluated, and a comparison with the standard regional frequency analysis over the area of Slovakia is performed. The advantages of the ROI approach are assessed by means of simulation studies. It is demonstrated that almost any setting of parameters of the ROI method yields estimates of growth curves and design values at individual sites that are superior to the standard regional and at-site estimates.

1 Introduction

Information on design values (quantiles) of heavy one-day and multi-day (2 or more, and usually up to 5 or 10-day) precipitation amounts is important in various fields of water resources engineering, e.g. the design of dams and sewer systems, flood prevention, protection against soil and vegetation loss, etc. In a traditional at-site approach to frequency analysis, the precipitation quantiles have long been estimated using a data sample at the site of interest only. This particularly holds for the area under study (Slovakia), where design values of extreme precipitation amounts were exclusively estimated by means of the Gumbel and/or Pearson type III distribution, using the at-site approach (Šamaj et al., 1985; Gaál et al., 2004). A recognized drawback of the single-site approach is related to the estimation of rare events, i.e. in practice, one often needs design values corresponding to return periods that are much larger than the lengths of available series of observations. In order to overcome the lack of at-site observations, the regional approach to frequency analysis was developed in the 1960s and “traded space for time” (Dalrymple, 1960). This approach, based on the index-flood method, has gained wider popularity since the 1980s (e.g. Wiltshire, 1986; Lettenmaier et al., 1987; Hosking and Wallis, 1993). The core idea of the regional approach is the substitution of time for space: one can obtain more reliable quantile estimates based on a multi-site analysis compared to an at-site approach. Nevertheless, in the regional approach, sites cannot be grouped in an arbitrary way; the resulting group of sites should suit the requirement of homogeneity; that is, sites pooled together exhibit similar probability distribution curves of extremes. Therefore, one of the most discussed issues of regional frequency analysis is a method for pooling groups of sites (e.g. Acreman and Sinclair, 1986; Wiltshire, 1986; Burn, 1988, 1990b).

At national meteorological offices worldwide, there have been intensive efforts to develop detailed statistical methods for design rainfall estimation during recent decades. In Great Britain, the Flood Studies Report (NERC, 1975) and later the Flood Estimation Handbook (FEH, 1999), which aimed at methods of estimation of extraordinary flood events, developed sophisticated procedures for the estimation of the design precipitation. The German KOSTRA project (Malitz, 1999; Malitz and Ertel, 2001), the Italian VAPI project (Cannarozzo et al., 1995), the HIRDS system in New Zealand (Thompson, 2002), the Australian Guide to Rainfall and Runoff (Pilgrim, 1987) and the internet-based Precipitation Frequency Atlas of the United States (Bonnin et al., 2006a, b) are other examples of complex national studies on risk assessment of heavy precipitation. Besides these projects, a number of studies in the scientific literature have dealt with regional precipitation frequency analysis. One of their common features is that the design values of heavy precipitation are estimated based on a fixed

Title Page
Abstract Introduction
Conclusions References
Tables Figures
◀◀ ▶▶
◀ ▶
Back Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

structure of the pooling groups (or regions), which are drawn either according to political (Pilon et al., 1991; Adamowski et al., 1996; Gellens, 2002; Wallis et al., 2007), geographical (Sveinsson et al., 2002; Kohnová et al., 2005), or climatological considerations (Smithers and Schulze, 2001; Fowler and Kilsby, 2003; Kyselý and Picek, 2007). Nowadays, a generally accepted guideline to regional frequency analysis is the regional L-moment algorithm (Hosking and Wallis, 1997), which combines the index flood method with the L-moments, which are generally preferred over ordinary (product) moments (e.g. Vogel and Fenessey, 1993; Sankarasubramanian and Srinivasan, 1999). An initial step of the regional L-moment algorithm is a delineation of groups of sites with a fixed (though not necessarily geographically contiguous) structure.

The region-of-influence (ROI) method introduced by Burn (1990a, b) is an alternative approach to regional frequency estimation. It was proposed for a flood frequency analysis in order to overcome possible inconsistencies that may occur on the boundaries of pooling groups (Acreman and Wiltshire, 1989). In such cases, a classical approach to regional analysis may lead to undesirable step changes of the variables and estimated quantiles. Geographical interpolation of quantiles at the boundaries of different pooling groups may be one of the possible solutions of the issue of inconsistency (Wallis et al., 2007). Nevertheless, the ROI method eliminates these deficiencies in another way. Its main idea is that there is no need to delineate fixed boundaries between the pooling groups; instead, they are defined in a flexible way. This means that each site has its own “region”; that is, a unique set of sufficiently similar stations, from which extreme precipitation information is transferred to the site of interest. Such an approach is termed focused pooling, as each site is regarded as the center of its own region, and a pooling group “is specifically tailored to a target site of interest and a given return period” (Čunderlík and Burn, 2002, 2006).

The similarity of the sites is evaluated by a properly selected set of site attributes (site characteristics and/or site statistics). Therefore, even two neighboring locations may have very different sets of stations that represent their regions of influence. Obviously, pooling groups in the ROI method do not necessarily form contiguous geographical regions.

Until now, the ROI method has been used in connection with flood frequency analysis (Burn, 1990a, b, 1997; Zrinji and Burn, 1994, 1996; Tasker et al., 1996; Provaznik and Hotchkiss, 1998; Castellarin et al., 2001; Holmes et al., 2002; Eng et al., 2005; Merz and Blöschl, 2005). The Flood Estimation Handbook (FEH, 1999) and the HIRDS system in New Zealand (Thompson, 2002) adopted the ROI concept as well. The present study focuses on the development of the ROI approach to the frequency analysis of precipitation extremes; we evaluate a number of alternative settings and present a comparison with the "standard" regional frequency analysis of Hosking and Wallis (1997).

Gaál (2006) showed that, in principle, the whole area of Slovakia may be treated as a compact homogeneous region, regardless of the duration of precipitation events (1 to 5 days) and/or seasons considered. Such a view, although it correctly fits the statistical concepts of regional frequency analysis, does not seem acceptable from a climatological point of view. Long-term experience indicates that even though the area of Slovakia is relatively small, it is unreasonable to treat the country as a single region (Faško and Lapin, 1996). Several precipitation regimes do exist in the area of the Western Carpathians, which are formed by western circulation, and Mediterranean and continental influences, and these are further differentiated by altitudinal zonality due to the rugged topography (Lapin and Tomlain, 2001). This naturally implies a need for a division of the country into sub-regions, preferably into contiguous geographical regions. The delineation of homogenous sub-regions while also conserving the major geographical units is rather difficult; moreover, it is influenced by a number of subjective judgments (Gaál, 2006). We therefore turned our attention to the ROI approach. The adoption of this methodology, particularly due to its flexibility, seemed very promising in a complex terrain like Slovakia. In the first stage, our goal was an application of Burn's original framework developed for flood frequency analysis in order to evaluate the performance of the ROI method compared to the traditional method of regionalization in Slovakia. The original concept of the ROI approach (Burn, 1990) has become the subject of criticism, particularly due to the need to choose a relatively large number of parameters according to subjective considerations (e.g. Hosking and Wallis, 1997). Zrinji and Burn (1994) revisited the ROI methodology: assignment of a site to a given pooling is judged by means of a built-in regional homogeneity test instead of subjectively selected threshold values. We decided not to use the ROI approach innovated by Zrinji and Burn (1994); since the area under study is sufficiently homogeneous (Gaál, 2006), it is anticipated that implementation of a homogeneity test into our decision processes would not contribute significantly to our analysis.

Our innovation with regard to the ROI methodology consists of alternative ways of defining the between-site similarity: the closeness of sites is determined not only according to the statistical properties of the at-site data, but also by means of the long-term characteristics of precipitation climate and geographical proximity of the stations. This paper is one of the first studies that focuses on an implementation of the ROI approach to the regional frequency analysis of heavy precipitation amounts. Data from rain gauge stations in Slovakia are used for this purpose; however, methodological findings are very likely independent of the target area and could also be useful for regional frequency models of rainfall extremes in different climatological conditions.

The paper is structured as follows: after a short description of the selected stations and their data in Sect. 2, a detailed overview of the background of the ROI method is presented in Sect. 3. Various settings of the ROI method are evaluated using the observed data in Sect. 4. A discussion and conclusions follow in Sect. 5.

2 Data

Daily precipitation amounts measured at 56 stations operated by the Slovak Hydrometeorological Institute (SHMI) were used as the input data set (Fig. 1). The altitudes of the stations range from 100 to 2635 m a.s.l., which cover the whole range of elevations in Slovakia, and the density of the selected sites is approximately one per 900 km^2.

Observations of the daily precipitation amounts without gaps over the period 1961–

2003 (in some cases since 1951) are available at 29 climatological stations. Since these sites do not cover the area of Slovakia evenly, it was necessary to extend the data set with sites having minor gaps in their daily rainfall records (breaks of several months). Each of the additional 27 sites has at least 35 complete years of observations. Figure 1 shows that the central and north parts of Western Slovakia were the main areas where the data set needed to be supplemented with other stations in order to ensure a more homogeneous spatial coverage. The basic data set at the selected 56 sites makes up 2464 station-years.

The block maxima approach to the selection of extremes has been adopted in the present study, and samples of the maximum seasonal 1 to 5-day precipitation amounts were drawn from each station record, separately in the warm (April to September) and the cold (October to March) seasons. Such an approach allows, at least in a rough approximation, for the differentiation of extreme precipitation amounts of various origins: convective events that dominate in summer and long-lasting episodes of precipitation from stratiform clouds of a frontal origin that are typical of the late autumn and winter in the area under study. All seasons with incomplete daily records at a given station were excluded from the analysis.

The entire data set (a combination of 5 durations of heavy precipitation events in 3 seasons) have been confined to 3 basic data sets representative in a frequency analysis of precipitation extremes:

- 1-D/Year – annual maxima of 1-day precipitation amounts;
- 2-D/WS – maxima of 2-day precipitation amounts in the warm season;
- 5-D/CS– maxima of 5-day precipitation amounts in the cold season.

The 1-D/Year data set has a key role in frequency modeling as it is a subject of interest of the majority of the applied studies concerning the probabilities of heavy rainfall. In Europe, impacts from heavy precipitation are mostly due to short-period rainfalls in summer and multi-day episodes in winter (Frei et al., 2006). That is why we also show results for 2-day precipitation amounts in the warm season and 5-day precipitation amounts in the cold season. Furthermore, such a selection of data sets represents short (1-day), longer (2-day) and multi-day (5-day) durations of precipitation, and both warm and cold seasons are involved in the analysis.

The data underwent standard quality checking for gross errors as well as checking in terms of a discordancy measure based on L-moments (Hosking and Wallis, 1993). A very small number of the data series were flagged as discordant; a detailed scrutiny revealed no rough errors in the data since the uniqueness of each discordant data series was induced by extraordinary local precipitation events (for further detail, see Gaál, 2006).

Due to the fact that the selected stations form a rather low density network and the precipitation extremes show a high temporal and spatial variability, it was not possible to test the homogeneity of the data series by the generally recommended "relative" methods (i.e. by comparing a site's data with a reliable and homogeneous reference series). Instead, each site's homogeneity was examined individually. Possible step-like changes in the data series were analyzed by applying four different homogeneity tests (Wijngaard et al., 2003): the standard normal homogeneity test (Alexandersson, 1986), Buishand's range test (Buishand, 1982), Pettitt's test (Pettitt, 1979) and von Neumann's ratio test (von Neumann, 1941). All but one of the data sets have been categorized as useful (according to Wijngaard's definition) at the significance level of α=1%, i.e. no clear signal of inhomogeneity is detected, and the data sets are sufficiently homogeneous for further analyses (Wijngaard et al., 2003). These findings were reaffirmed by testing for trends using the non-parametric method of Wald and Wolfowitz (1943): no significant trends were detected in the individual series (except for the highest elevated station of Lomnický štít, 2635 m a.s.l.; significant trends in the annual data and warm season), so the data sets can be regarded as stationary as well.

3 Mathematical models of the region-of-influence method

The region of influence for a given site consists of a group of sites that are sufficiently similar to the site of interest; the similarity of the sites is judged according to the site characteristics and/or site statistics (Hosking and Wallis, 1997). Site characteristics are quantities that are known a priori to the frequency analysis at a given site and include, for example, the location, the elevation and other physical-geographical properties associated with the site. Some long-term characteristics of the precipitation regime of sites (such as mean annual precipitation and long-term averages of monthly/seasonal precipitation totals) are also frequently classified here: although they are set from the at-site measurements, their values can also be estimated at ungauged sites from climatological maps with a relatively smaller degree of uncertainty compared to any extremes. Site statistics are simply the measurements or results of the statistical processing of the observed data at a given site. Hereafter, site characteristics and site statistics will generally be termed site attributes.

3.1 Distance metric

The distance metric serves to determine the proximity of sites in an attribute space. There are a number of alternative definitions of a distance metric reported mostly in connection with a cluster analysis (Cormack, 1971, classifies 10 different definitions of the distance metric). In the context of the ROI method, however, only the Euclidean distance metric is used (e.g. Burn, 1990a, b; Zrinji and Burn, 1994; Castellarin et al., 2001; Holmes et al., 2002), probably due to the fact that it is the most intuitive one. The Euclidean distance metric is defined as

$$D_{ij}=\left[\sum_{m=1}^{M} W_m \left(X_m^i - X_m^j\right)^2\right]^{\frac{1}{2}}, \tag{1}$$

where D_{ij} is the weighted Euclidean distance between sites i and j; W_m is the weight associated with the m-th site attribute; X_m^i is the value of the m-th attribute at site i; and M is the number of attributes. The distance metric matrix D is symmetrical ($D_{ij}=D_{ji}$) with zeros on its diagonal ($D_{ii}=0$).

The region of influence for a given station is formed according to the following scheme: First, the site with the lowest value from the whole set of $D_{ij},\ j=1,\ldots,N$ is added to the ROI for site i. In the very first step it is the site i itself, for which the distance metric $D_{ii}=0$ is always the one with the smallest value. Then, the next site with the second smallest value of D_{ij} is added into the ROI for site i. The sites are successively pooled into the ROI as long as a given condition (see Sect. 3.2) is satisfied.

As the site attributes X_m may have substantially different magnitudes, a transformation of the initial values before calculating D_{ij} Eq. (1) is usually applied. The simplest alternative is a standardization of the variables:

$$X \rightarrow \frac{X-\bar{X}}{\sigma_X}, \tag{2}$$

where $\bar{X}$ is the mean, and σ_X is the standard deviation of attribute X. As a result, all the site attributes X_m are of a comparable magnitude, i.e. they have a zero mean and unit variance.

The selection of the site attributes plays a key role in the ROI method: the success of the whole procedure depends on finding the right number and combination of proper site characteristics and/or statistics. In Sects. 3.1.1–3.1.3 below, three different alternatives to the selection of the site attributes are evaluated.

In the current analysis, equal (unit) weights $W_m=1,\ m=1,\ldots,M$ in Eq. (1) have been chosen for each alternative of the site attribute sets. We did not attempt to adjust the relative importance of the site attributes because (i) we did not find a-priori reasons for assigning different weights to the site attributes, and (ii) the study focuses on other, more likely important, aspects of the ROI approach.

3.1.1 Alternative #1: site statistics

The whole set of the site attributes consists exclusively of statistical characteristics that are related to the data sample examined (the annual/seasonal maxima of the k-day precipitation amounts) at each site. The following site statistics are considered:

1. coefficient of variation (c_v) – a traditional characteristics of the scale of a data sample:

   $$c_v = \frac{\sigma}{\mu}, \tag{3}$$

   where μ (σ) is the sample mean (standard deviation).

2. Pearson's second skewness coefficient (PS) – a less traditional characteristic of the skewness of a data sample (Weisstein, 2002):

   $$PS = \frac{3(\mu - m)}{\sigma}, \tag{4}$$

   where m is the median of the sample.

3. 10-year design precipitation estimated using the generalized extreme value (GEV) distribution – a characteristic of the extreme value magnitudes of a data sample.

The set of attributes in alternative #1 follows Burn's concept (Burn, 1990b). There are, however, minor differences compared to the original settings. Pearson's skewness coefficient Eq. (4) is used instead of its modification (Burn, 1990b); since attribute PS underlies standardization according to Eq. (2), there is no difference between results based on the "original" and the "modified" PS. Another difference concerns the selection of the third site attribute. Burn used a plotting position estimate of a 10-year flood event, interpolated from available annual flow series. Instead of that estimate, a GEV estimation of the 10-year value is used herein, for two main reasons: (i) According to

Lu and Stedinger (1992), "the normalized 10-year flood estimator has very little bias and approximately a normal distribution in small samples". As the sites in the present analysis have between 35 and 53 years of observations, the design precipitation with a return period of $T=10$ years is reliably estimated from the at-site data samples. (ii) The GEV is a generally acceptable distribution for the annual maxima of one-day and multi-day precipitation amounts in central Europe (Gaál, 2006; Kyselý and Picek, 2007).

The selection of the above-described site statistics ensures that the ROI for a given site consists of sites for which the characteristics of the probability distribution functions are similar to those at the target site. Using site statistics has, however, limited applicability to gauged sites – in the case of ungauged sites, the ROI method is restricted to site characteristics (Sects. 3.1.2–3.1.3) that may be estimated from maps.

3.1.2 Alternative #2: general climatological site characteristics

The second set of the site attributes consists of characteristics that describe the long-term precipitation regime of the country. Slovakia is a relatively small, landlocked country in Central Europe with an area of 49 035 km^2. Its topography is complex: rugged mountains in its central and northern parts (the Western Carpathian Mountains, encompassing the High and Low Tatras), and lowland areas in its southern parts. 60% (15%, 1%) of the area of the country are located in altitudes above 300 (800, 1500) m a.s.l. (Marečková et al., 1997).

Slovakia lies in an area where various maritime and continental influences meet. The precipitation regime is affected by different factors; the dominant ones are the effect of a) the Mediterranean area, b) the western circulation (the Atlantic Ocean), and c) the European continent. The influence of the Mediterranean area has a pronounced role in the inter-annual variability of the monthly precipitation amounts, mainly in the autumn and in southern Slovakia. In general, the annual cycle of precipitation has a maximum in June (on average 95 mm, especially in the south) and a minimum in February (on average 43 mm). However, due to cyclones moving from the area of the Ligurian Sea (the Mediterranean area), a secondary autumn maximum (in October/November) appears at the majority of stations of South Slovakia. Lapin's index of the Mediterranean effect L_M is a quantitative characteristic of the magnitude of this influence. It is defined using 3 ratios of certain monthly precipitation totals:

$$L_M = \frac{R_{\mathrm{Max}}}{R_{VII}} + \frac{R_V}{R_{VII}} + \frac{R_{\mathrm{Max2}}}{R_{\mathrm{Min2}}} - 2.5, \tag{5}$$

where the indices denote May (V), July (VII) and months with the maximum (Max), secondary maximum (Max2) and secondary minimum (Min2) precipitation amount in the annual cycle. The number 2.5 in Eq. (5) is a correction factor. For a detailed description of the index L_M, refer to Gaál (2005). At some stations in the northwestern part of Slovakia, the November/December secondary maximum in the monthly precipitation is influenced by the late-autumn/early-winter precipitation increase in the North Atlantic area.

The spatial distribution of the mean annual precipitation (MAP) exhibits strong variability. A general slight descent of the MAP from the west to the east is superimposed on by an altitudinal zonality due to the topography; therefore, the lowest values of the MAP occur in the south-west Danubian lowlands (about 500 mm), while the largest precipitation totals are observed at the highest windward slopes of the Carpathian Mountains (more than 1500 mm). The daily precipitation amounts may, in extreme cases (due to heavy convective storms), exceed 150 mm.

Considering the general precipitation climate in Slovakia, the following variables have been selected in alternative #2 of the distance metric:

1. the mean annual precipitation;
2. the ratio of the precipitation totals for the warm/cold season;
3. Lapin's index of the Mediterranean effect (Eq. 5).

Using the site characteristics of alternative #2 in the distance metric would, in principle, result in groups of sites with similar climatological conditions that may, to some extent,

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀◀ | ▶▶
◀ | ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

also be related to mechanisms generating heavy precipitation. In practice, however, there is no guarantee that the proximity of sites in the M-dimensional space of the climatological site characteristics implies a similarity in the extreme precipitation regimes.

3.1.3 Alternative #3: geographical site characteristics

The third set of site attributes consists of the following basic geographical characteristics:

1. latitude;
2. longitude;
3. elevation above sea level.

Alternative #3 yields groups of sites whose members are similar to the site of interest in a geographical sense. Nevertheless, this cannot be interpreted as a simple geographical proximity between two points; e.g. higher-elevated sites are usually grouped together and are not necessarily joined with other nearby sites in the traditional sense of latitude and longitude.

3.2 Pooling a station's ROI

When the appropriate site attributes are selected, and the distance metric matrix is calculated, two other issues need to be addressed. The first is to determine the cutoff point of the distance metric for the i-th site: only sites below a selected threshold will be included in the i-th site's ROI:

$$\mathrm{ROI}_i = \left\{ j : D_{ij} \leq \theta_i \right\}, \tag{6}$$

where ROI_i is the set of stations in the pooling group for site i, and θ_i is the threshold distance value for site i (Burn, 1990b).

The other important issue is a determination of the pooled weighting coefficients, which must reflect the relative proximity of any site in the pooling group to the site of interest. The closer a site of the ROI$_i$ to the site i according to the distance metric, the greater amount of information it provides in the pooled frequency analysis. The weight η_{ij} for site j in the ROI$_i$ is a function of the distance metric D_{ij} and several other parameters (see Sects. 2.2.1–2.2.3). Obviously, sites that are not included in the ROI$_i$ have zero weights.

Following Burn's (1990b) framework, the threshold distance θ_i and the weights η_{ij} in the current analysis are determined according to 3 different options that reflect 3 diverse concepts of pooling information from the sites of the ROI.

3.2.1 Option #1: "Fewer sites with high values of the weights"

The basic idea of option #1 is that the ROI for a given site encompasses only a limited number of stations; however, all of the selected stations are assigned weights markedly different from zero.

The threshold value θ_i (Eq. 5) is defined as follows:

$$\theta_i=\theta_L \qquad \text{if} \qquad NS_i \geq NST, \tag{7}$$

and

$$\theta_i=\theta_L+(\theta_U-\theta_L)\left(\frac{NST-NS_i}{NST}\right) \qquad \text{if} \qquad NS_i<NST, \tag{8}$$

where θ_L (θ_U) is the lower (upper) threshold value, NS_i is the number of stations in the ROI$_i$ with a threshold at θ_L, and NST is the target number of stations for the ROI.

The weighting function for this option includes two parameters (TP, n) to be determined:

$$\eta_{ij}=1-\left(\frac{D_{ij}}{TP}\right)^n . \tag{9}$$

The settings of 5 parameters to be initialized are in accordance with Burn's original concept: θ_L (θ_U, TP) is the 25 (75, 85) percentile of the distance metric distribution, $NST=15$, and $n=2.5$. For a detailed line of the reasoning concerning the parameter settings, see Burn (1990b).

3.2.2 Option #2: "More sites with different values of the weights"

In option #2, a relatively large number of sites are included in the ROI for a given site. Stations sufficiently similar to the site of interest have unit weights, while lower values of weights are assigned to those less similar.

The threshold value θ_i (Eq. 5) is constant,

$$\theta_i=\theta_U. \tag{10}$$

The weighting function for this option is defined as:

$$\eta_{ij}=1 \qquad \text{if} \qquad D_{ij} \leq \theta_L, \tag{11}$$

and

$$\eta_{ij}=1-\left(\frac{D_{ij}-\theta_L}{TN-\theta_L}\right)^n \qquad \text{if} \qquad \theta_L<D_{ij} \leq \theta_U, \tag{12}$$

where θ_L is a lower threshold value, and TN and n are the parameters of the weighting function. TN is defined using a further parameter TPP as

$$TN=\max\left[\max_{\{j\}}\left(D_{ij}\right), TPP\right]. \tag{13}$$

There are 4 parameters of the weighting function for option #2 to be initialized. θ_L (θ_U, TPP) is selected as the 25 (75, 85) percentile of the distance metric distribution, and $n=0.1$ (Burn, 1990b).

3.2.3 Option #3: "All sites with different values of the weights"

Option #3 is similar to option #2 with the only difference being that all the available stations are included in the ROI for a given site, with appropriate values of the weighting function.

The threshold value θ_i is defined as

$$\theta_i = \max_{\{j\}} \left(D_{ij}\right). \tag{14}$$

The definition of the weighting function and the parameter settings are the same as in option #2 – see Eqs. (9–12). There is no need to deal with the selection of the upper threshold θ_U; the number of the parameters to be initialized is 3 (θ_L, TPP, n) (Burn, 1990b).

3.3 Estimation of the at-site quantiles using information from the ROI

When each station's ROI and the appropriate weighting coefficients are known, it is possible to estimate the at-site precipitation quantiles using information from the ROI by means of the L-moment-based index storm procedure (Hosking and Wallis, 1997). The at-site data $X_{j,k}$, $j=1, \ldots, N$, $k=1, \ldots, n_j$ (where N stands for the number of sites, and n_j denotes the sample size of the j-th site) are rescaled by the sample mean μ_j (index storm) in order to get dimensionless data:

$$x_{j,k} = \frac{X_{j,k}}{\mu_j}, \quad k=1, \ldots, n_j. \tag{15}$$

The dimensionless values of $x_{j,k}$ at site j are then used to compute the sample L-moments $l_1^{(j)}$, $l_2^{(j)}$, ... and L-moments ratios:

$$t^{(j)} = \frac{l_2^{(j)}}{l_1^{(j)}} \tag{16}$$

and

$$t_r^{(j)} = \frac{l_r^{(j)}}{l_2^{(j)}}, \quad r=3,4,\ldots, \tag{17}$$

where $t^{(j)}$ is the sample L-coefficient of variation (L-CV) and $t_r^{(j)}$, $r=3,4,\ldots$ are the sample L-moments ratios at site j (for a definition and description of the L-moments, see Hosking, 1990; Hosking and Wallis, 1997).

The regional (pooled) L-moment ratios $t^{(i)R}$ and $t_r^{(i)R}$, $r=3,4,\ldots$, within the ROI for site i are derived from the at-site sample L-moment ratios as weighted regional averages. Two weights are applied, sample size n_j (the length of the observations) and the weighting function based on the ROI distance metric η_{ij}:

$$t^{(i)R} = \frac{\sum\limits_{j\in \mathrm{ROI}_i} t^{(j)} n_j \eta_{ij}}{\sum\limits_{j\in \mathrm{ROI}_i} n_j \eta_{ij}} \tag{18}$$

and

$$t_r^{(i)R} = \frac{\sum\limits_{j\in \mathrm{ROI}_i} t_r^{(j)} n_j \eta_{ij}}{\sum\limits_{j\in \mathrm{ROI}_i} n_j \eta_{ij}}, r=3,4,\ldots, \tag{19}$$

where ROI_i is the set of stations forming the ROI for site i (Eq. 5), for which the weighted regional L-moment ratios are calculated. The regionally weighted values $t^{(i)R}$ and $t_r^{(i)R}$, $r=3,4,\ldots$ are then used to estimate the parameters of the selected distribution function in order to get the dimensionless cumulative distribution function (growth curve). The precipitation quantiles with a return period T are obtained by multiplying the dimensionless T-year growth curve value x_i^T with the index storm μ_i:

$$X_i^T = \mu_i x_i^T. \tag{20}$$

A universal parametric model for extremes, the generalized extreme value (GEV) distribution (e.g. Coles, 2001), is applied as the pooled distribution function in the current analysis. It has been identified as a suitable model for 1-day as well multi-day precipitation extremes in central Europe, including the area of Slovakia (Gaál, 2006; Kohnová et al., 2005).

3.4 Confidence intervals for the estimated quantiles

In order to evaluate the uncertainty associated with the estimated quantiles and the performance of various ROI approaches (combinations of "alternatives" and "options"), Monte Carlo simulations are carried out. The basic idea of the Monte Carlo simulation procedure is that the unknown at-site parent distribution is assumed to be identical with the at-site sample distribution at each site (Burn, 1988).

Monte Carlo simulations consist of $NR=1000$ repetitions of the following steps:

1. Samples of annual maxima are generated at each station, having the same record lengths as their real-world counterparts. The simulated data samples at the i-th site have the GEV distribution as the parent, with parameters corresponding to the at-site L-moments $[1, t^{(i)}, t_3^{(i)}]$.

2. The at-site statistics and the distance metric matrix are calculated, and the region of influence and weighting function values are determined for each station, each alternative of the attribute sets, and each option of the pooled weight definition. Note that in case of alternatives #2 and #3, there is no need to set D_{ij}, ROI_i and η_{ij} in each repetition of the Monte Carlo simulation as the distance metric is determined from unchanged site characteristics.

3. The at-site estimates of the L-moments, the regional (pooled) L-moments within each station's ROI, and the simulated extreme precipitation quantiles for each station are determined according to the above-described alternatives and options.

From the Monte Carlo experiment, it is possible to draw confidence intervals for the estimated extreme precipitation quantiles. For each quantile x_i^T (for all the combinations of the alternatives and options), the 90% confidence interval is estimated from the 5th and 95th percentiles of the empirical distribution of the simulated quantiles (Hosking and Wallis, 1997).

3.5 Evaluation of the ROI approaches

The relative performance of the various ROI settings is evaluated using the Monte Carlo simulated extreme precipitation quantiles. The simulated quantiles, estimated by the ROI methodology, are used together with the "true" (parent) at-site quantiles to calculate the root mean square error (RMSE) and bias for each quantile at site i:

$$\mathrm{RMSE}_i^T = \left[\frac{1}{NR} \sum_{m=1}^{NR} \left(\frac{\hat{x}_{i,m}^T - x_i^T}{x_i^T} \right)^2 \right]^{\frac{1}{2}} \tag{21}$$

and

$$\mathrm{BIAS}_i^T = \frac{1}{NR} \sum_{m=1}^{NR} \left(\frac{\hat{x}_{i,m}^T - x_i^T}{x_i^T} \right). \tag{22}$$

Equations (20) and (21) are summations over repetitions of the Monte Carlo experiment (m=1 to NR); RMSE_i^T and BIAS_i^T are the root mean square error and the relative bias for the return period T at site i, respectively; x_i^T is the "true" value for the T-year event at site i (from the GEV distribution used as the parent in the simulations), and $\hat{x}_{i,m}^T$ is the simulated value of the T-year event at site i from the m-th sample of the Monte Carlo simulations. A summary characteristic describing the performance of the given model is the average root mean square error (RMSE^T) and average bias (BIAS^T), respectively, obtained by summations over all the stations:

$$\mathrm{RMSE}^T = \frac{1}{N}\sum_{i=1}^{N} \mathrm{RMSE}_i^T \tag{23}$$

and

$$\mathrm{BIAS}^T = \frac{1}{N}\sum_{i=1}^{N} \mathrm{BIAS}_i^T . \tag{24}$$

In the current study, only the statistical properties of the dimensionless growth curves are examined. We decided not to focus on an analysis of the design values since the simulated design values $\hat{X}_{i,m}^T$ (within the Monte Carlo experiments) are a product of the simulated growth curves $\hat{x}_{i,m}^T$ and simulated index storm values $\hat{\mu}_{i,m}$ (Eq. 20). The uncertainty of the design values is affected by the uncertainties of both factors, which makes the interpretation more difficult and less relevant to the aims of the study.

The performance of the different ROI models is compared (i) with the results of a regional frequency analysis using the "conventional" regionalization approach of Hosking and Wallis (1997), in which 3 homogeneous regions are delineated within Slovakia (HW3r, see Sect. 3.6); (ii) with the results of a regional frequency analysis which treats the whole country as a single homogeneous region (HW1r); and (iii) with the results of a traditional at-site (local) frequency analysis lacking a regional approach.

The 9 ROI models are labeled as aXoY, where X=1, 2, 3, and Y=1, 2, 3. The first part of the acronym denotes the selected alternative of the distance metric; the second part shows which of the three options for the transfer of the regional information is applied.

3.6 Regional frequency analysis in Slovakia using a traditional approach

Homogeneous regions for precipitation frequency analysis in Slovakia by the "traditional" approach (Hosking and Wallis, 1997) have been delineated, using the data set described in Sect. 2, in 3 steps (Gaál, 2006):

1. Cluster analysis, an objective method of pooling has been used with 5 discriminating variables: latitude, longitude, elevation, the ratio of the precipitation totals for the warm/cold season and Lapin's index of the Mediterranean effect (Sect. 3.1.2). The analysis resulted in delineation of 5 homogeneous and geographically contiguous regions.

2. A process-based (subjective) regionalization has been proposed by P. Faško, one of the best specialists on the long-term precipitation conditions of Slovakia (P. Faško, personal communication, 2006[1]). Taking into consideration the rich topography of the country and the effects of different patterns of general air-mass circulation, 4 regions of extreme precipitation have been identified.

3. The final set of Slovak regions for the regional frequency analysis of heavy precipitation is a compromise between the result of the cluster analysis (objective method) and a physically-based delineation (subjective method). It consists of 3 regions (Fig. 2). Further details on the whole process of regionalization are given in Gaál (2006).

4 Results

A summary of the results of the Monte Carlo simulations is shown in Tables 1–3 and Figs. 3–5. The tables present the point characteristics of each model in terms of the average values of the root mean square error (RMSE^T, Eq. 23) and the bias (BIAS^T, Eq. 24) of the simulated growth curves (averaged over the stations). The box plots (Figs. 3–5) offer a broader overview of the mathematical models analyzed in a more transparent form. Besides displaying the point characteristics (median), they enable a comparison of the spread of the statistical characteristics among the stations, in terms

---

[1] pavol.fasko@shmu.sk, Slovak Hydrometeorological Institute, March, 2006.

of the inter-quartile range (25%–75%) and the 5% and 95% quantiles. Even though only the box plots of RMSE_i^T, $i=1, \ldots, 56$ corresponding to the return periods of $T=10$, 20, 50 and 100 years are presented, the general conclusions are drawn according to the whole set of results.

The description of the results is organized as follows: the 9 ROI models (the 2 Hosking-Wallis regional models) are inter-compared in Sect. 4.1 (4.2); the local models are briefly evaluated in Sect. 4.3; and the performance of all the frequency models examined is compared in Sect. 4.4.

4.1 Evaluation of the ROI models

Focusing on the 9 ROI models only, a majority of them have a very small positive bias (regardless of the return period), i.e. they slightly overestimate the actual growth curve values (Tables 1–3). The models a3oX (X stands for 1, 2 or 3) for 5-D/CS, however, represent exceptions with an increased bias for the return periods $T \geq 10$ years (Table 3). This can be explained by the complex topography of the area under study. The frontal systems that bring moist air masses toward Slovakia in the cold season mostly come from the western and south-western directions (Lapin and Tomlain, 2001); therefore, different precipitation regimes on the windward and leeward sides of the mountains in Slovakia appear. A distance metric based on the geographical proximity of sites in such a terrain does not prove to be a good choice for a frequency analysis of large-scale events like cold season precipitation, as it may often pool stations with considerably different precipitation climates in the ROI of a given site. For example, the ROI for any station in the central parts of the country may, if geographical proximity is utilized in the distance metric, consist of stations from higher elevated locations (the High and Low Tatras) as well as lowland stations (South Slovakia), and sites located at windward as well as leeward mountain slopes. The performance of the a2oX models, which do not show large bias values for the same precipitation events (Table 3), since the sites of a given ROI are pooled according to climatological characteristics, supports this hypothesis.

Since the inter-model variability of the bias is relatively small and does not yield a clear figure that would support a distinction between the various ROI models, the evaluation of the models is based on RMSE in the following parts of the paper.

The root mean square error of the simulated growth curves enables a more efficient comparison of the ROI models. The average values of RMSE^T in Tables 1–3 clearly show that the a1o3 model performs best for all the durations (seasons). This dominance is evident for the return periods $T \geq 20$ years and does not depend on the duration of the extreme precipitation events. Such a pattern is also captured in the box plots (Figs. 3–5). The ROI model a1o3 always possesses the best statistical properties within alternative #1 as well as overall; only for shorter return periods and the 75% and 95% quantiles, may the models a1o1/a1o2 perform slightly better.

When comparing the various alternatives of the distance metric (a1/a2/a3), the first one, which pools regional information according to the statistical properties of the data samples at the individual stations, is superior to the other two. This is pronounced for the annual maxima of the 1-D amounts (Table 1) as well as the maxima of the 5-D amounts in the cold season (Table 3). For 2-D events in the warm season (Table 2), the performance of the a3oX models is comparable with that of the a1oX models, but a1o3 is still the best one among the models. For the other two options (o1 and o2), a3oX show lower values of RMSE^T than a1oX at high return periods. The small difference between a1 and a3 in the warm season may be due to the fact that heavy rainfall in the warm season mostly originates from local, small-scale convective events, so the distance metric based on the geographical proximity of the stations also manages to pool stations with similar properties of extreme precipitation. Nevertheless, the group of the a1oX models has the narrowest boxes and whiskers among the pooled frequency models (Figs. 3–5), for all the precipitation durations and in terms of both RMSE and bias (the latter is not shown). This means that they are generally suitable models, performing relatively poorly at a very small number of stations (compared to the other models).

Alternative #1 is obviously the best one for determining the proximity of sites; the other two alternatives yield comparable results. Alternative #3 performs slightly better (it has narrower intervals of RMSE) than #2 in the warm season (Fig. 4), which is, again, a consequence of the fact that extreme precipitation events in the warm season are rather local phenomena, so the distance metric based on geographical characteristics is able to cope with the warm season precipitation in a more efficient way than the one in alternative #2. The poorer performance of alternative #3 in the cold season (Fig. 5) has been discussed above.

From the various options for the transfer of regional information (o1/o2/o3) at a given alternative of the distance metric, it is much more difficult to pick out the best one. In the first alternative, option #3 gains from incorporating each station into a given site's ROI; however, it is nearly impossible to determine which one of the three options is the best one of the other two alternatives (a2/a3). Figures of RMSE for various X in a2oX and a3oX are very similar or even equal to each other (tiny differences occur often only at the third decimal place, Tables 1–3), so this information is insufficient for drawing more general conclusions. According to the widths of the 25%–75% and 5%–95% intervals (Figs. 3–5), it is also hard to make an unequivocal decision as to whether the best option of constructing a site's ROI is o1, o2 or o3. However, using the preferred distance metric in a1, the fact that the narrowest 5%–95% interval belongs to option #3, regardless of the duration, becomes more and more evident with increasing return period.

4.2 Evaluation of the Hosking-Wallis regional models

Of the two Hosking-Wallis models of regional frequency analysis, HW3r demonstrates a better performance than HW1r for 1-D/Year (Table 1) and 5-D/CS (Table 3), with a pronounced difference between the RMSE values in the cold season, mainly for the return periods $T \geq 50$ years. On the other hand, the quantile estimates from HW3r in the warm season demonstrate slightly less favourable statistical properties (lower values of average RMSE in Table 2) than the quantiles from HW1r. These differences are in accordance with those reported in Sect. 4.1, as they result from the varied spatial extent of the precipitation events in the warm/cold season. Extraordinary precipitation amounts in the warm season are likely to arise mostly from strong, local convective rain showers that are little dependent on topography; therefore, in principle, similar statistical properties of heavy warm-season amounts may be found at relatively distant locations. As a consequence, the delineation of smaller sub-regions does not necessarily improve the quantile estimates. On the other hand, precipitation extremes in the cold season are predominantly large-scale events that may have remarkably diverse properties at different parts of the country. As a result, the delineation of the groups of stations seems to be beneficial in the cold season, since pooling in smaller regions may capture the regional differences in the extreme precipitation in a more efficient way.

4.3 Evaluation of the local models

The RMSE statistic of the at-site estimates is either the highest one (the warm season, Table 2) or ranks among the two highest ones (the whole year and the cold season, Tables 1 and 3) for any return period T. Furthermore, box and whiskers plots of the at-site model (Figs. 3–5) are markedly different from the other ones: although both intervals are narrow, they are related to the highest values of the median, regardless of the duration of the precipitation events. The at-site model is an odd-one-out among the frequency models, since the uncertainty of the estimated quantiles stems only from the sampling variability of the L-moment estimators and not from different settings of the regional approach. The results demonstrate that the at-site approach to frequency analysis is the least suitable method for the estimation of heavy precipitation quantiles.

4.4 Comparison of the ROI, Hosking-Wallis and local models

The results of the comparison of the 12 examined models of the frequency analysis (9 ROI models, 2 HW models, and at-site model) are summarized as follows:

- The most appropriate model of the regional/pooled frequency analysis is a variant of the ROI approach a1o3, in which the between-site similarity is determined according to the statistical properties of the at-site data samples, and the regional information is pooled with appropriate weighting coefficients from all the stations under study.

- The three ROI models based on the first alternative of the distance metric (a1oX, X=1, 2 or 3) are obviously superior to both conventional Hosking-Wallis models. Clear conclusions, however, cannot be drawn for other models of the ROI approach (a2oX, a3oX). For example, the HW3r model indicates an apparently better performance than any of the models a2oX and a3oX for 5-D/CS; on the other hand, in the case of 1-D/Year and 2-D/WS, even the performance of the worst ROI model is comparable to that of the better one of the HW models.
- Local at-site models should be avoided since they lead to a large variance in the estimated growth curves.

5 Conclusions

The region-of-influence (ROI) method, which was designed in flood frequency studies to avoid inconsistencies at the boundaries of regions involved in conventional regional approaches, is shown to also be a very useful tool for the frequency modeling of heavy precipitation events. Nine different combinations of the site attributes (that enter the distance metric) and weighting functions (used to pool regional information) were evaluated using Monte Carlo simulations. Among the 9 ROI models, the best alternative of the distance metric is the one in which the proximity of the sites is determined according to the statistical characteristics of the empirical frequency distributions of the examined data. It is less obvious which option for the transfer of regional information is the most suitable one, but the approach that makes use of all the available observations (with different weights) seems to be superior to the other two. A smaller number of parameters that need to be initialized makes this option advantageous, too. Nevertheless, the superiority of the best ROI model (which uses all the sites available in the analysis) may stem from the fact that the area under study is relatively small and that the analyzed data sets of 1 to 5-day precipitation extremes are sufficiently homogeneous (Gaál, 2006). An analysis carried out using heterogeneous data sets covering a considerably larger geographical area may lead to different results.

The most important finding stems from a comparison of the two basic approaches to the regional frequency analysis: the ROI method, which makes use of flexible regions in order to pool regional information to the site of interest, is superior to the traditional approach of a regional analysis based on firmly separated groups of sites. Even though the performance of the ROI models based on climatological and geographical characteristics is comparable with that of the Hosking-Wallis models, the ROI method based on the statistical properties of data samples obviously outperforms the conventional regional frequency models.

Note that in the presented paper, we only focused on a regional analysis of extreme precipitation from sites equipped with rain gauges. It is obvious that when a regional analysis is aimed at obtaining quantiles at ungauged sites as well, alternative #1 is not feasible as data samples to determine the site statistics are not available. In such cases, much attention should be paid to finding the most appropriate climatological characteristics describing the heavy precipitation regime.

The results presented were obtained using simulation experiments based on precipitation data in a particular area in central Europe (Slovakia); however, it is likely that at least some of the methodological findings may be rather general and independent of the target area. We recommend the ROI method for frequency estimates of heavy precipitation events in different climatological conditions in other parts of the world, particularly in areas with complex orography.

Acknowledgements. The work was supported under the young-scientists' research project KJB300420601 by the Grant Agency of the Academy of Sciences of the Czech Republic (L. Gaál, and J. Kyselý), the HYDRATE project, Contract GOCE 037024 (L. Gaál and J. Szol-

gay) and the project 1/4024/07 of the Slovak VEGA Grant Agency (J. Szolgay).

References

Acreman, M. C. and Sinclair, C. D.: Classification of drainage basins according to their physical characteristics, and application for flood frequency analysis in Scotland, J. Hydrol., 84(3), 365–380, 1986.

Acreman, M. C. and Wiltshire, S. E.: The regions are dead; long live the regions. Methods of identifying and dispensing with regions for flood frequency analysis, FRIENDS in Hydrology, IAHS Pub. no. 187, 175–188, 1989.

Adamowski, K., Alila, Y., and Pilon, P. J.: Regional rainfall distribution for Canada, Atmos. Res., 42(1–4), 75–88, 1996.

Alexandersson, H.: A homogeneity test applied to precipitation data, J. Climatol., 6, 661–675, 1986.

Bonnin, G. M., Todd, D., Lin, B., Parzybok, T., Yekta, M., and Riley, D.: Precipitation-Frequency Atlas of the United States. NOAA Atlas, Vol. 1, v4.0. NOAA, NWS, Silver Spring, Maryland, 261pp., http://hdsc.nws.noaa.gov/hdsc/pfds/docs/NA14Vol1.pdf, 2006a.

Bonnin, G. M., Todd, D., Lin, B., Parzybok, T., Yekta, M., and Riley, D.: Precipitation-Frequency Atlas of the United States. NOAA Atlas, Vol. 2, v3.0. NOAA, NWS, Silver Spring, Maryland, 301pp., http://hdsc.nws.noaa.gov/hdsc/pfds/docs/NA14Vol2.pdf, 2006b.

Buishand, A.: Some methods for testing the homogeneity of rainfall records, J. Hydrol., 58(1–2), 11–27, 1982.

Burn, D. H.: Delineation of groups for regional flood frequency analysis, J. Hydrol., 104, 345–361, 1988.

Burn, D. H.: An appraisal of the "region of influence" approach to flood frequency analysis, Hydrolog. Sci. J., 35(2), 149–165, 1990a.

Burn, D. H.: Evaluation of regional flood frequency analysis with a region of influence approach, Water Resour. Res., 26(8), 2257–2265, 1990b.

Burn, D. H.: Catchment similarity for regional flood frequency analysis using seasonality measures, J. Hydrol., 202(1–4), 212–230, 1997.

Cannarozzo, M., D'Asaro, F., and Ferro, F.: Regional rainfall and flood frequency analysis for

Sicily using the two component extreme value distribution, Hydrolog. Sci. J., 40(1), 19–42, 1995.

Castellarin, A., Burn, D. H., and Brath, A.: Assessing the effectiveness of hydrological similarity measures for flood frequency analysis, J. Hydrol., 241, 270–287, 2001.

Coles, S.: An introduction to statistical modeling of extreme values. Springer-Verlag, London, 222pp., 2001.

Čunderlík, J. M. and Burn, D. H.: Analysis of the linkage between rain and flood regime and its application to regional flood frequency estimation, J. Hydrol., 261, 115–131, 2002.

Čunderlík, J. M. and Burn, D. H.: Switching the pooling similarity distances: Mahalanobis for Euclidean, Water Resour. Res., 42(3), W03409, doi:10.1029/2005WR004245, 2006.

Dalrymple, T.: Flood frequency analyses, Water Supply Paper 1543-A., U.S. Geological Survey, Reston, 1960.

Eng, K., Tanker, G. D., and Milly, P. C. D.: An analysis of region-of-influence methods for flood regionalization in the Gulf-Atlantic rolling plainss, J. Am. Water Resour. As., 41(1), 135–143, 2005.

Faško, P. and Lapin, M.: Relation of daily and monthly precipitation totals in Slovakia and atmospheric circulation types in the period 1874–1993, in: Universitas Iagellonica Acta Scientiarum Litteraumque, Proc. of Int. Conf. on Climate Dynamics and the Global Change Perspective, Cracow, 175–182, 1996.

FEH: Flood Estimation Handbook, 1–5, Institute of Hydrology, Wallingford, 1999.

Fowler, H. J. and Kilsby, C. G.: A regional frequency analysis of United Kingdom extreme rainfall from 1961 to 2000, Int. J. Climatol., 23, 1313–1334, 2003.

Frei, C., Schöll, R., Fukutome, S., Schmidli, J., and Vidale, P. L.: Future change of precipitation extremes in Europe: Intercomparison of scenarios from regional climate models, J. Geophys. Res., 111, D06105, doi:10.1029/2005JD005965, 2006.

Gaál, L., Lapin, M., and Faško, P.: Maximum several-day precipitation totals in Slovakia, in: Proc. of Extremes of the Weather and Climate, Brno, 11-03-2004, edited by: Rožnovský, J. and Litschmann, T., Czech Bioclimatological Assoc., Brno, 15pp., (CD-ROM), 2004.

Gaál, L.: Introduction of Lapin's indices into the cluster analysis of maximum k-day precipitation totals in Slovakia, Meteorological Journal, 8(2), 85–94, 2005.

Gaál, L.: Estimation methods of statistical properties of short-term to several-day design precipitation in Slovakia, (PhD thesis), Faculty of Mathematics, Physics and Informatics, Comenius University in Bratislava, 228pp., (in Slovak, with English abstract), 2006.

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

Holmes, M. G. R., Young, A. R., Gustard, A., and Grew, R.: A region of influence approach to predicting flow duration curves within ungauged catchments, Hydrol. Earth Syst. Sci., 6, 721–731, 2002, http://www.hydrol-earth-syst-sci.net/6/721/2002/.

Hosking, J. R. M. and Wallis, J. R.: Some statistics useful in regional frequency analysis, Water Resour. Res., 29, 271–281, 1993.

Hosking, J. R. M. and Wallis, J. R.: Regional frequency analysis: an approach based on L-moments. Cambridge University Press, Cambridge, 224pp., 1997.

Kohnová, S., Gaál, L., Szolgay, J., and Hlavčová, K.: Analysis of heavy precipitation amounts in the Upper Hron region. Faculty of Civil Engineering, Slovak University of Technology, Bratislava, 120pp., (in Slovak), 2005.

Kyselý, J. and Picek, J.: Regional growth curves and improved design value estimates of extreme precipitation events in the Czech Republic, Climate Res., 33, 243–255, 2007.

Lapin, M. and Tomlain, J.: General and Regional Climatology, Comenius University Edition, Bratislava, 184pp., (in Slovak), 2001.

Lettenmaier, D. P., Wallis, J. R., and Wood, E. F.: Effect of regional heterogeneity on flood frequency estimation, Water Resour. Res., 23(2), 313–323, 1987.

Lu, L.-H. and Stedinger, J. R.: Sampling variance of normalized GEV/PWM quantile estimators and a regional homogeneity test, J. Hydrol., 138, 223–245, 1992.

Malitz, G.: Starkniederschlag in Deutschland – Messergebnisse, statistische Auswertungen, Schätzungen, in: Klimastatusbericht 1999, Deutscher Wetterdienst, 35–41, http://www.dwd.de/de/FundE/Klima/KLIS/prod/KSB/ksb99/malitz.pdf, 1999.

Malitz, G. and Ertel, H.: Extremwertstatistische Auswertungen von Tageswerten der Niederschlagshöhe, in: Klimastatusbericht 2001, Deutscher Wetterdienst, 94–99, http://www.dwd.de/de/FundE/Klima/KLIS/prod/KSB/ksb01/ExtremwertstatistischeAuswertungen.pdf, 2001.

Marečková, K., Lapin, M., Minárik, B., Mojík, I., Závadský, I., Závodský, D., and Zuzula, I.: Country Study Slovakia, Final Report, U.S. edited by: Závodský, D., Country Studies Program, Slovak Ministry of the Environment, Slovak Hydrometeorological Institute, Bratislava, 108pp., 1997.

Merz, R. and Blöschl, G.: Flood frequency regionalization – spatial proximity vs. catchment attributes, J. Hydrol., 302, 283–306, 2005.

NERC: Flood Studies Report, 1–5, Natural Environment Research Council, London, 1975.

Pettitt, A. N.: A non-parametric approach to the change-point detection, Appl. Stat., 28, 126–135, 1979.

Pilgrim, D. H. (Ed.): Australian Rainfall and Runoff: A Guide to Flood Estimation, 1, Institution of Engineers, Barton, Australia, ACT, 1987.

Pilon, P. J., Adamowski, K., and Alila, Y.: Regional analysis of annual maxima precipitation using L-moments, Atmos. Res., 27, 81–92, 1991.

Provaznik, M. K. and Hotchkiss, R. H.: Analysis of gauging station flood frequency estimates in Nebraska using L-moments and region of influence methods, Transport. Res. Rec., 1647, 53–60, 1998.

Sankarasubramanian, A. and Srinivasan, K.: Investigation and comparison of sampling properties of L-moments and conventional moments, J. Hydrol., 218(1–2), 13–34, 1999.

Smithers, J. C. and Schulze, R. E.: 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, 2001.

Sveinsson, O. G. B., Salas, J. D., and Boes, D. C.: Regional frequency analysis of extreme precipitation in Northeastern Colorado and Fort Collins Flood of 1997, J. Hydrol. Eng., 7(1), 49–63, 2002.

Šamaj, F., Valovič, Š., and Brázdil, R.: Daily precipitation amounts with extraordinary intensity in ČSSR in the period 1901–1980. In Proceedings of Works of SHMI, No. 24. Alfa, Bratislava, (in Slovak, with English summary), 9–113, 1985.

Tasker, G. D., Hodge, S. A. and Barks, C. S.: Region of influence regression for estimating the 50-year flood at ungaged sites, Water Resour. Bull., 32(1), 163–170, 1996.

Thompson, C. S.: The high intensity rainfall design system: HIRDS (Abstract), in: Int. Conf. on Flood Estimation, Bern, Switzerland, 6–8 March, 273–282, 2002.

Vogel, R. M. and Fenessey, N. M.: L-moments diagrams should replace product moment diagrams, Water Resour. Res., 29(6), 1745–1752, 1993.

von Neumann, J.: Distribution of the ratio of the mean square successive difference to the variance, Ann. Math. Statist., 12(4), 367–395, 1941.

Wald, A. and Wolfowitz, J.: An exact test for randomness in the nonparametric case based on serial correlation, Ann. Math. Statist., 14(4), 378–388, 1943.

Wallis, J. R., Schaefer, M. G., Barker, B. L., and Taylor, G. H.: Regional precipitation-frequency analysis and spatial mapping for 24-hour and 2-hour durations for Washington State, Hydrol.

Earth Syst. Sci., 11, 415–442, 2007, http://www.hydrol-earth-syst-sci.net/11/415/2007/.

Weisstein, E. W.: Pearson's Skewness Coefficients. From MathWorld – A Wolfram Web Resource, http://mathworld.wolfram.com/PearsonsSkewnessCoefficients.html, 2002.

Wijngaard, J. B., Klein Tank, A. M. G., and Können, G. P.: Homogeneity of 20th century European daily temperature and precipitation series, Int. J. Climatol., 23(6), 679–692, 2003.

Wiltshire, S. E.: Identification of homogeneous regions for flood frequency analysis, J. Hydrol., 84, 287–302, 1986.

Zrinji, Z. and Burn, D. H.: Flood frequency analysis for ungauged sites using a region of influence approach, J. Hydrol., 153(1–4), 1–21, 1994.

Zrinji, Z. and Burn, D. H.: Regional flood frequency with hierarchical region of influence, J. Water Res. Pl.-ASCE, 122(4), 245–252, 1996.

Table 1. Average root mean square error (RMSE^T) and average bias (BIAS^T) of growth curves of annual maxima of 1-day precipitation amounts for return periods T. The smallest values of RMSE^T and BIAS^T (in absolute sense) are indicated in bold.

RMSE^T T[yrs]	a1o1	a1o2	a1o3	a2o1	a2o2	a2o3	a3o1	a3o2	a3o3	At-site	HW3r	HW1r
5	0.017	0.017	0.016	0.017	0.016	0.016	0.016	0.016	0.016	0.021	0.015	0.022
10	0.028	0.028	0.027	0.031	0.031	0.031	0.032	0.031	0.031	0.033	0.032	0.038
20	0.050	0.049	0.047	0.054	0.054	0.054	0.056	0.054	0.054	0.057	0.056	0.060
50	0.085	0.084	0.078	0.090	0.090	0.090	0.091	0.089	0.089	0.099	0.093	0.092
100	0.115	0.115	0.103	0.119	0.120	0.119	0.120	0.117	0.117	0.136	0.123	0.121
200	0.148	0.148	0.130	0.149	0.150	0.149	0.151	0.146	0.147	0.178	0.155	0.156

BIAS^T ($\times 10^{-2}$) T[yrs]	a1o1	a1o2	a1o3	a2o1	a2o2	a2o3	a3o1	a3o2	a3o3	At-site	HW3r	HW1r
5	0.349	0.396	0.495	0.793	0.708	0.704	0.732	0.704	0.710	0.084	0.507	0.265
10	0.299	0.461	0.441	0.855	0.747	0.704	0.853	0.800	0.774	0.133	0.398	0.655
20	0.208	0.485	0.303	0.812	0.700	0.612	0.892	0.813	0.751	0.268	0.251	0.869
50	0.099	0.531	0.068	0.708	0.607	0.453	0.919	0.804	0.687	0.657	0.077	0.774
100	0.072	0.623	–0.110	0.650	0.565	0.355	0.973	0.828	0.666	1.158	0.009	0.486
200	0.122	0.794	–0.260	0.643	0.578	0.310	1.088	0.908	0.699	1.872	0.031	0.105

Table 2. Average root mean square error (RMSE^T) and average bias (BIAS^T) of growth curves of maxima of 2-day precipitation amounts in the warm season for return periods T. The smallest values of RMSE^T and BIAS^T (in absolute sense) are indicated in bold.

RMSE^T T [yrs]	a1o1	a1o2	a1o3	a2o1	a2o2	a2o3	a3o1	a3o2	a3o3	At-site	HW3r	HW1r
5	0.020	0.019	0.019	0.019	0.019	0.019	0.018	0.018	0.018	0.021	0.020	0.020
10	0.030	0.030	0.029	0.034	0.032	0.032	0.031	0.030	0.030	0.034	0.034	0.033
20	0.050	0.049	0.046	0.055	0.053	0.053	0.051	0.050	0.051	0.058	0.057	0.054
50	0.081	0.080	0.074	0.087	0.085	0.083	0.082	0.080	0.080	0.098	0.090	0.086
100	0.109	0.108	0.098	0.112	0.110	0.108	0.107	0.104	0.105	0.134	0.117	0.111
200	0.139	0.138	0.123	0.138	0.136	0.133	0.134	0.130	0.130	0.174	0.145	0.137

BIAS^T ($\times 10^{-2}$) T[yrs]	a1o1	a1o2	a1o3	a2o1	a2o2	a2o3	a3o1	a3o2	a3o3	At-site	HW3r	HW1r
5	0.313	0.352	0.444	0.746	0.633	0.638	0.619	0.568	0.581	0.023	0.451	0.468
10	0.328	0.467	0.475	0.946	0.781	0.762	0.709	0.662	0.658	0.097	0.464	0.476
20	0.310	0.550	0.437	1.058	0.859	0.810	0.728	0.694	0.668	0.273	0.438	0.439
50	0.306	0.678	0.348	1.169	0.941	0.845	0.729	0.722	0.661	0.724	0.417	0.391
100	0.357	0.828	0.288	1.273	1.030	0.892	0.759	0.775	0.684	1.268	0.450	0.395
200	0.482	1.051	0.260	1.424	1.170	0.986	0.841	0.882	0.758	2.016	0.551	0.458

Table 3. Average root mean square error (RMSE^T) and average bias (BIAS^T) of growth curves of maxima of 5-day precipitation amounts in the cold season for return periods T. The smallest values of RMSE^T and BIAS^T (in absolute sense) are indicated in bold.

RMSE^T T[yrs]	a1o1	a1o2	a1o3	a2o1	a2o2	a2o3	a3o1	a3o2	a3o3	At-site	HW3r	HW1r
5	0.019	0.019	0.018	0.019	0.018	0.018	0.019	0.019	0.019	0.023	0.019	0.019
10	0.030	0.030	0.029	0.034	0.033	0.033	0.035	0.035	0.035	0.035	0.035	0.036
20	0.053	0.053	0.051	0.060	0.059	0.059	0.061	0.062	0.061	0.060	0.059	0.064
50	0.090	0.090	0.084	0.100	0.099	0.099	0.101	0.101	0.100	0.101	0.096	0.107
100	0.121	0.121	0.111	0.131	0.130	0.131	0.133	0.133	0.132	0.137	0.125	0.141
200	0.154	0.155	0.139	0.164	0.163	0.164	0.166	0.166	0.166	0.177	0.156	0.176
BIAS^T ($\times 10^{-2}$) T[yrs]	a1o1	a1o2	a1o3	a2o1	a2o2	a2o3	a3o1	a3o2	a3o3	At-site	HW3r	HW1r
5	0.360	0.405	0.500	0.599	0.616	0.621	0.832	0.825	0.792	–0.052	0.556	0.706
10	0.205	0.383	0.347	0.331	0.534	0.484	1.475	1.281	1.251	0.021	0.363	0.539
20	0.008	0.316	0.109	–0.014	0.382	0.275	2.061	1.669	1.646	0.212	0.102	0.297
50	–0.241	0.240	–0.258	–0.490	0.167	–0.018	2.842	2.171	2.162	0.700	–0.258	–0.027
100	–0.378	0.234	-0.535	-0.816	0.040	–0.203	3.480	2.589	2.591	1.279	–0.496	–0.222
200	–0.442	0.301	–0.788	–1.084	–0.030	–0.331	4.186	3.065	3.079	2.066	–0.681	–0.344

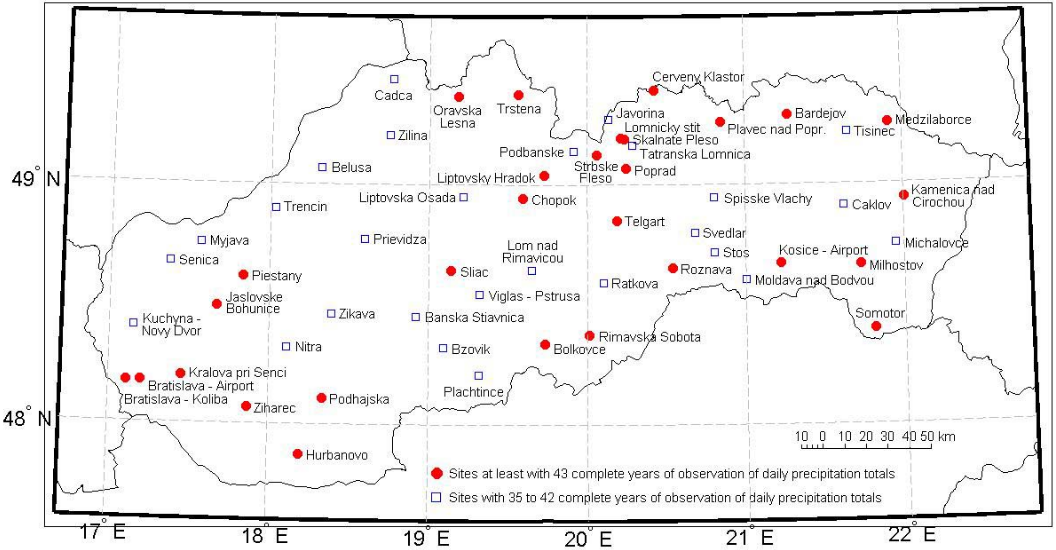


Fig. 1. 56 climatological stations in Slovakia selected for a regional frequency analysis of heavy precipitation amounts.

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

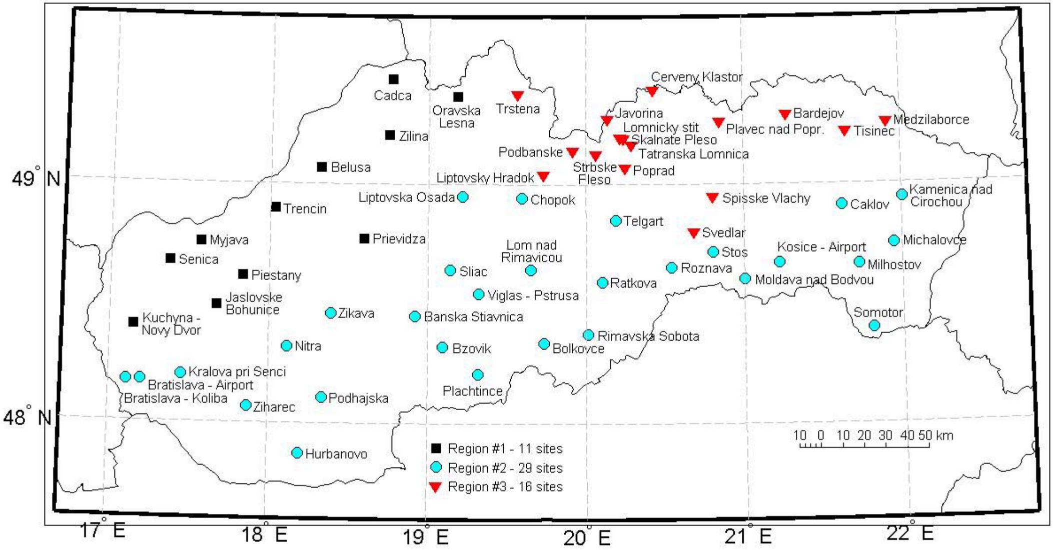


Fig. 2. Delineation of 3 homogeneous regions for frequency analysis of heavy precipitation amounts using the conventional regionalization approach of Hosking and Wallis.

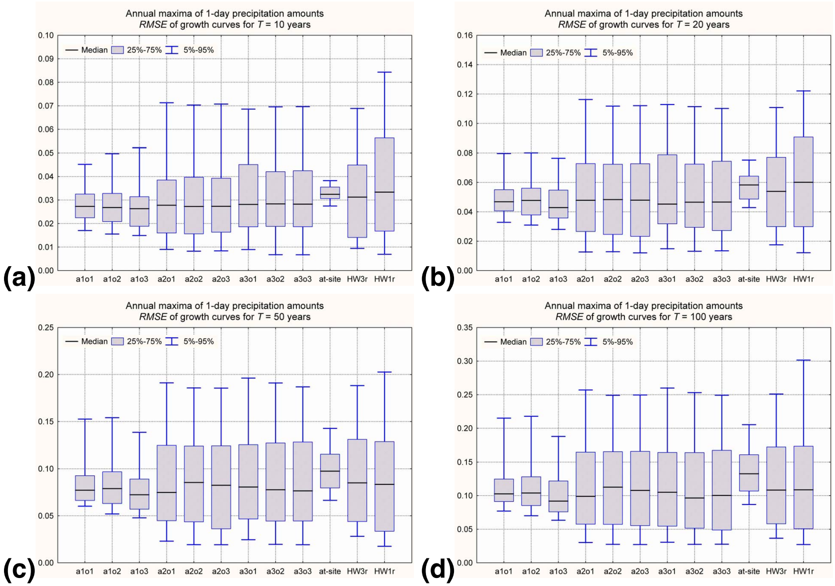


Fig. 3. Root mean square error (RMSE) of growth curves of annual maxima of 1-day precipitation amounts. T denotes return period. **(a)** $T=10$ years, **(b)** $T=20$ years, **(c)** $T=50$ years, **(d)** $T=100$ years.

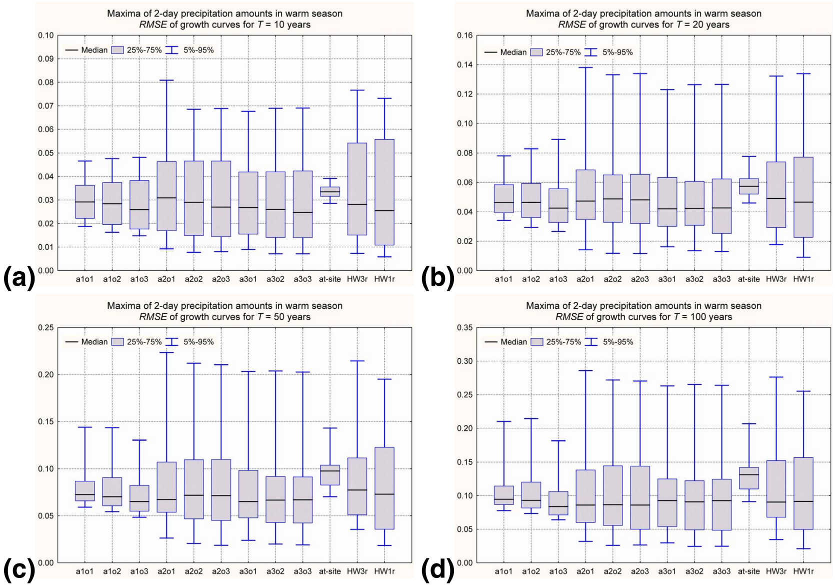
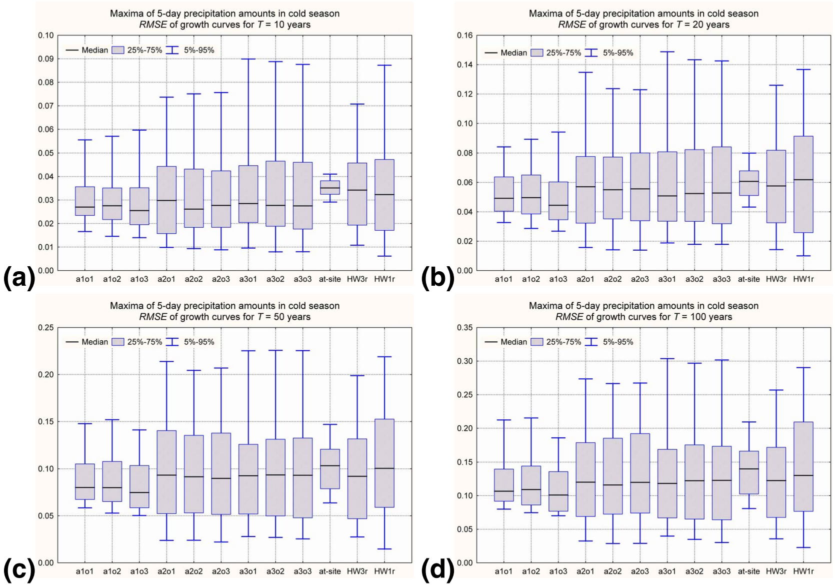


Fig. 4. Root mean square error (RMSE) of growth curves of maxima of 2-day precipitation amounts in the warm season. T denotes return period. **(a)** T=10 years, **(b)** T=20 years, **(c)** T=50 years, **(d)** T=100 years.

Title Page
Abstract
Introduction
Conclusions
References
Tables
Figures
◀◀
▶▶
◀
▶
Back
Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion



Maxima of 5-day precipitation amounts in cold season
RMSE of growth curves for *T* = 10 years

Maxima of 5-day precipitation amounts in cold season
RMSE of growth curves for *T* = 20 years

Maxima of 5-day precipitation amounts in cold season
RMSE of growth curves for *T* = 50 years

Maxima of 5-day precipitation amounts in cold season
RMSE of growth curves for *T* = 100 years

Fig. 5. Root mean square error (RMSE) of growth curves of maxima of 5-day precipitation amounts in the cold season. T denotes return period. **(a)** T=10 years, **(b)** T=20 years, **(c)** T=50 years, **(d)** T=100 years.