

Automatic threshold and run parameter selection: a climatology for extreme hourly precipitation in Switzerland

S. Fukutome · M. A. Liniger · M. Süveges

Received: 22 July 2013 / Accepted: 19 May 2014 / Published online: 29 May 2014
© The Author(s) 2014. This article is published with open access at Springerlink.com

Abstract Extreme value analyses of a large number of relatively short time series are in increasing demand in environmental sciences and design. Here, we present an automated procedure for the peaks-over-threshold (POT) approach to extreme value theory and use it to provide a climatology of extreme hourly precipitation in Switzerland. The POT approach fits the generalized Pareto distribution (GPD) to independent exceedances above some high threshold. To guarantee independence, the time series is pruned: exceedances separated by less than a fixed interval called the run parameter are considered a cluster, and all but the cluster maxima are discarded. We propose the automation of an existing graphical method for joint selection of threshold and run parameter. Hourly precipitation is analyzed at 59 stations of the MeteoSwiss observational network over the period 1981–2010. The four seasons are considered separately. When necessary, a simple detrending is applied. Results suggest that unnecessarily large run parameters have adverse effects on the estimation of the GPD parameters. The proposed method yields mean cluster sizes that reflect the seasonal and geographical variation of lag dependence of hourly precipitation. The climatology, as represented by the return level maps and Alpine cross-section, mirror known aspects of the Swiss climate. Unlike for daily precipitation, summer thunderstorm tracks are visible in the seasonal frequency of events, rather than in the amplitude of rare events.

1 Introduction

Peaks-over-threshold (POT) analysis (Davison and Smith 1990) is an approach to extreme value theory that, as its name indicates, uses the observations exceeding a high threshold to elicit information regarding the behavior of extremes. Since it takes into account more than one value per year, it lends itself well to the analysis of time series covering only a short period. Unfortunately, the selection of the threshold requires expert judgment, and stands in the way of an automatic analysis. In the present paper, we propose a “blind” selection procedure that allows the analysis to be performed automatically at a large number of stations. We apply it to hourly precipitation in Switzerland.

Due to its potentially dramatic consequences, intense hourly precipitation is of great importance for engineering and design. In the Alpine region, it is associated with flash floods, mudslides and landslides, and debris flow (Guzzetti et al. 2007; Borga et al. 2010; Toreti et al. 2013). In the form of snow, it can help trigger avalanches. Engineers have long exploited empirical relationships between precipitation amplitudes of different duration to derive the information they need for design (e.g., Geiger et al. 1991; Koutsoyiannis et al. 1998). Yet systematic analyses of extreme hourly precipitation have been hindered by the lack of long records of data at high temporal resolution. Since the 1980s, MeteoSwiss has at its disposal a network of automatic stations that record precipitation at 10-min intervals, thus providing a collection of relatively long time series of precipitation at sub-daily resolution.

The idea of using as extremes the values of a distribution exceeding a high threshold for estimating the amplitude of rare events first saw the light in the context of hydrology (e.g., Todorovic and Zelenhasic 1970). Later on, it was integrated into the context of extreme value

S. Fukutome (✉) · M. A. Liniger
Federal Office of Meteorology and Climatology, MeteoSwiss,
Operation Center 1, P.O. Box 257, CH-8058
Zürich-Flughafen, Switzerland
e-mail: sophie.fukutome@meteoswiss.ch

M. Süveges
Department of Astronomy, ISDC Data Centre for Astrophysics,
University of Geneva, Chemin d’Ecogia 16, CH-1290 Versoix,
Switzerland

theory (Balkema and de Haan 1974; Pickands 1975), according to which the distribution of excesses over a high threshold can be approximated by the generalized Pareto distribution (GPD), provided the largest observations over fixed blocks of time converge towards a non-degenerate distribution. In practice, estimating the GPD parameters implies choosing a threshold large enough for this approximation to be justified.

An important assumption for the approximation with a GPD is the independence of excesses. Environmental observations are manifestations of physical processes with their own time scales and cannot, a priori, be assumed to be independent. As it turns out, threshold exceedances do not necessarily occur in isolation, but are almost always “clustered” together. Clustering is an indication of dependence in the time series, although dependence does not necessarily imply clustering. For dependent series, the distribution of the tail remains of the same family, but applies to the cluster maxima rather than individual exceedances (Davison and Smith 1990).

In practice, the “true” clusters are not known and one must resort to defining them according to some reasonable but artificial criteria. The most common method, used in this paper, is runs declustering, which assumes independent exceedances to be separated by a minimum number of non-exceedances. This minimum “distance” is called the run parameter (Coles 2001). Should two exceedances be separated by a number of non-exceedances smaller than the run parameter, they are considered to form a cluster. Ultimately, all but the largest exceedance within a cluster are discarded.

Generally, expert knowledge of the variable under consideration is used to select the run parameter. This subjective choice can be avoided, however, with the procedure proposed by Ferro and Segers (2003): the run parameter is determined by the mean cluster size, which is estimated from the data. Fawcett and Walshaw (2007) advise against declustering altogether, because their simulations indicate that it leads to biased estimates of parameters and return levels.

Thus, the peaks-over-threshold method presents a particular difficulty: an appropriate choice must be made not only for the run parameter, but also for the threshold. Choosing too high a threshold results in small samples and high uncertainties, while with too low a threshold, the variance is small, but the bias is potentially large. A widely used graphical selection for the threshold, based on stability properties, was proposed by Davison and Smith (1990). The method developed later by Dupuis (1998) to guide threshold selection, while not purely graphical, cannot be used blindly because it requires careful judgment by the practitioner.

Clearly, the value selected for the threshold can be expected to modify the dependence structure of the

resulting time series, and thereby affect the minimum distance necessary to separate dependent exceedances. Thus, as Walshaw (1994) points out, the threshold and run parameter should be chosen in combination. A high threshold allows for a smaller run parameter and vice versa (Palutikof et al. 1999). This issue is addressed by Süveges and Davison (2010), who devise a test that ultimately allows joint graphical selection of threshold and run parameter. In the context of climate research, however, it is not uncommon to be confronted with a huge number of time series. Thus, graphical selection is not practicable and there is a need for automatic selection. For heavy-tailed distributions, Frigessi et al. (2002) devise a method that can be fully automated. A further approach that can be automated was proposed by Wadsworth and Tawn (2012), but implies choosing the run parameter separately.

In this paper, we propose a simple automation of the method developed by Süveges and Davison (2010), which tests all pairs of thresholds and run parameters for misspecification of the model for inter-exceedance times. As will be explained below, the automatic selection simply consists in choosing the pair yielding the largest number of observations within a subset with particularly low misspecification. This procedure allows us both to automate our analysis, and to jointly select threshold and run parameter.

We apply this method to time series of hourly precipitation extending from 1981 to 2010 at 59 stations in Switzerland, and present a climatology of extreme hourly precipitation. Dependence at extreme levels is examined independently with the help of dependence measures derived from extreme value theory (Coles et al. 1999), in order to shed light on its seasonal and regional characteristics.

Our paper is organized as follows: The statistical methods applied in this study and the data used for the analysis are presented in Section 2. The results in terms of method and climatic characteristics of hourly extreme precipitation are given in Section 3, and discussed in Section 4.

2 Methods and data

Extreme value statistics offers three approaches to the analysis of rare events. The first, Block Maxima, approximates the parent distribution’s tail with a distribution for the maxima over time blocks of equal size (Fisher and Tippett 1928; Gnedenko 1943). The second, peaks-over-threshold (POT), approximates the behavior of extremes with a distribution for the values over a high threshold. The point process model unifies the first two approaches. It describes occurrences in time or space, and can be used to model threshold excesses as occurrences in time with a given amplitude. In the present paper, we use the POT

approach. The use of POT requires the selection of a threshold to satisfy the asymptotic conditions. In addition, in the approach used here, a minimum separation interval between exceedances, called the run parameter, is selected to guarantee independence. In a cluster consisting of exceedances separated by less than the run parameter, only the maximum exceedance is retained. Thus, each combination of values for the threshold and run parameter leads to a distinct series of observations and intervals. The time intervals between consecutive exceedances are called inter-exceedance times; those that separate clusters of exceedances are named inter-cluster times; finally, the intervals between exceedances within a cluster will be referred to as intra-cluster times.

The automatic joint selection of threshold and run parameter proposed in this paper hinges on the time intervals between exceedances. In theory, inter-cluster times must follow an exponential distribution, while intra-cluster times tend to zero. This assumption is verified separately for a set of pairs of threshold and run parameter, among which one pair is selected.

2.1 Theory

The peaks-over-threshold approach to extreme value theory formulates a limiting distribution for the excesses over a high threshold. Let X_1, \dots, X_n be a strictly stationary sequence with marginal distribution F , such that the dependence at extreme levels decays asymptotically. The excesses $Y = X - u$ over a threshold u , conditional on $X > u$, converge towards a limiting distribution H called the generalized Pareto distribution (GPD):

$$H(y) = \begin{cases} 1 - \left(1 + \frac{\xi y}{\sigma}\right)^{-\frac{1}{\xi}}, & \xi \neq 0, \\ 1 - \exp\left(-\frac{y}{\sigma}\right), & \xi = 0, \end{cases}$$

where $y > 0$, and H is defined on $1 + \xi y / \sigma > 0$ (Coles 2001; Beirlant et al. 2004).

The parameters of the GPD are the scale σ , which is a measure for the spread of the distribution, and the shape parameter ξ describing the behavior in the tail. If $\xi = 0$ ($\xi > 0$), the distribution is called light-tailed (heavy-tailed). If $\xi < 0$, the distribution is bounded, i.e., the excesses have an upper bound.

One important application of extreme value statistics is to estimate the amplitude of rare events expected to be exceeded on average once every T years. These amplitudes are referred to as return levels for the return period T . For a declustered sequence, the quantity of interest is the rate at which clusters occur (Coles 2001). In this study, we use the formulation by

Palutikof et al. (1999), in which the number of exceedances per year is modeled with a Poisson distribution with expected value λ . Let $1/\theta$ denote the mean cluster size. Then, the T year return level is given by

$$x_T = u + \frac{\hat{\sigma}}{\hat{\xi}} \left[\left(T \hat{\lambda} \hat{\theta} \right)^{\hat{\xi}} - 1 \right].$$

For stationary sequences and in the limit of large n , Hsing (1987) shows that the extremal process can be interpreted as a 2-dimensional process with dimensions time and threshold excess. The time intervals are normalized by the number of observations, so that the entire process takes place between 0 and 1. As n becomes large, intra-cluster times collapse to zero, and the clusters, rather than the individual exceedances, are independent. Projected on the time axis, the cluster occurrence follows a Poisson process, while in the other dimension, the largest excess in each cluster, follows a GPD (Davison and Smith 1990). This approach can be exploited to derive the asymptotic distribution of inter-exceedance times, which is at the heart of the automatic selection of threshold and run parameter presented in this paper.

2.2 Modeling of inter-exceedance times and misspecification test

Ferro and Segers (2003) show that for very high thresholds and in the limit of large n , the inter-exceedance times converge to a mixture distribution with parameter θ , named the extremal index: intra-cluster times tend to zero and occur with a probability $(1 - \theta)$, while inter-cluster times converge to an exponential distribution and occur with probability θ . The mean of the exponential distribution is $1/\theta$, which turns out to be the mean cluster size mentioned above. Thus, θ plays a double role: it is both the proportion of inter-cluster times, and the reciprocal of the mean inter-cluster time. Süveges and Davison (2010) apply the information matrix test (IMT) developed by White (1982) to the likelihood of the limit law of the inter-exceedance times (details of the likelihood function can be found in the Appendix).

In order to use the likelihood function for the distribution of inter-exceedance times in practical applications, Süveges and Davison (2010) truncate intervals that exceed the run parameter in length. The resulting limit law for inter-exceedance times takes the same form as in Ferro and Segers (2003), but the intra-cluster inter-exceedance times, which have length 0, can be accounted for in the likelihood function, allowing for an estimation of the mean cluster size that is not biased towards 1.

For a formal treatment of the IMT in this particular context, the reader is referred to Süveges and Davison (2010). Essentially, the IMT rests on the fact that for a well-specified model,

Fisher's information matrix $I = E \left\{ \frac{\partial^2 \mathcal{L}}{\partial \theta^2} \right\}$ equals the variance of the score vector $J = \text{Var} \left\{ \frac{\partial \mathcal{L}}{\partial \theta} \right\}$, where \mathcal{L} denotes the log-likelihood and E the expected value (see also Davison 2008, chapter 4). The null hypothesis H_0 is that the model is well specified, in which case the difference $D = I - J$ should vanish. The IMT statistic can then be constructed as D divided by its asymptotic variance, and is χ_1^2 -distributed for large samples. Thus, H_0 can be rejected at the 5 % level for $IMT > 3.84$. The formula for the IMT can be found in the Appendix.

2.3 Automatic selection of threshold and run parameter

The IMT provides a quantitative assessment of the compatibility of the threshold–run parameter pair with the two-dimensional extremal process. It is used to try, one by one, all combinations of the two parameters in a plausible range, and results in a list of pairs that are not rejected at the 5 % confidence level, i.e., with $IMT < 3.84$. The automated procedure proposed here is pragmatic: it takes a subset thereof for which the IMT is close to zero—and hence, misspecification is liable to be small—and selects the pair leading to the largest number of observations after declustering. Here, we take $IMT < 0.05$ (corresponding to a p value of 0.82) as a convenient upper limit for this subset of “non-rejected” IMT values. When the threshold–run parameter pairs lead to a number of exceedances smaller than 80, they are discarded because simulations revealed that there is not enough data to determine whether the pair should be rejected (Süveges and Davison 2010).

Suppose N observations from the stationary sequence X_1, \dots, X_n exceed the threshold u . The probability of exceedance of the threshold is then N/n . Let the indices $\{j_i : X_{j_i} > u\}$ denote the locations of the exceedances, and $T_i = j_{i+1} - j_i$ ($i = 1, \dots, N-1$) the inter-exceedance times. Let K denote the run parameter, and $c_i^{(u,K)} = (N/n) \max\{T_i - K, 0\}$ be the inter-exceedance times truncated by K and normalized by the probability of exceedance. In effect, K splits the sequence of inter-exceedance times T_i into clusters, separated by inter-cluster intervals. Consecutive exceedances separated by an interval equal to or shorter than K are within the same cluster, and $c_i^{(u,K)} = 0$. Between clusters, the intervals are simply shortened by K . Let N_c denote the number of clusters, and θ the extremal index, i.e., the parameter of the asymptotic distribution for inter-exceedances times.

The automated procedure is done as follows:

1. For each (u, K) pair, compute $c_i^{(u,K)}$.
2. For each (u, K) pair, determine N_c . Compute the IMT (see Appendix), and estimate θ , the extremal index.
3. Determine the (u, K) pairs for which $IMT < 0.05$.
4. Select the (u, K) pair for which N_c is largest.

The range chosen for u is between the 90th and the 99.5th percentile of non-zero values. Note, however, that the zero values were retained in the computation of inter-exceedance times. The values for K extend from 1 to 120 h. In a previous analysis, a threshold equal to the 90th percentile of non-zero values (corresponding to $u = 0.90$) was selected by applying the graphical approach of Davison and Smith (1990) at a subset of stations representing different climatic regimes in Switzerland (Begert 2008). This threshold was combined with a run parameter of 5 days, thus safely guaranteeing independence of the observations. We shall regard this threshold–run parameter pair ($u = 0.90, K = 120$) as a reference with which to compare our results. For simplicity, we will refer to the automated procedure as the “IMT selection”, and the selected pair as the “IMT pair.” The pair ($u = 0.90, K = 120$) will be called “reference pair,” and the use of the reference pair regardless of the season “reference method” or “reference selection.”

2.4 Inference

In this study, the GPD parameters are estimated by maximum likelihood, using the log-likelihood in Davison and Smith (1990) with an added term for the Poisson distribution of the number of exceedances per year.

As we have seen above, the return level x_T depends not only on the GPD parameters, but also on the reciprocal mean cluster size θ . Thus, estimation of its confidence intervals requires knowledge of the dependence between θ and the GPD parameters, which is unknown, since they are estimated separately. Confidence intervals for the return levels can nevertheless be determined by a bootstrapping procedure inspired from Ferro and Segers (2003), in which the inter-exceedance times are first categorized into inter-cluster times (between clusters) on the one hand, and intra-cluster times (within clusters) on the other.

The clusters, each consisting of a sequence of exceedances separated by intra-cluster times, are resampled with replacement a sufficiently large number of times. Separately, the inter-cluster times are also resampled with replacement. A new, artificial, time series is then reconstructed by alternating a cluster with an inter-cluster time. The time series is truncated when the original number of exceedances is reached. This procedure preserves the structure of the inter-exceedance times, including the probability of exceedance and the sequence of exceedances within a cluster. The GPD parameters and return levels are then estimated from the resulting artificial time series for the threshold and run parameter selected on the basis of the original time series. The procedure is repeated 5,000 times and the 95 % confidence intervals for the return levels evaluated.

As each (u, K) pair selection leads to a different time series, neither common criteria for model selection, nor goodness-of-fit tests are appropriate for a quantitative comparison of the

quality of the fits based on declustered series resulting from IMT or reference selection. Here, we opt for a quantitative summary of the QQ plot, which can be seen as a visual guide to goodness of fit. Let N_c denote the number of clusters, and $z_{(1)} \leq \dots \leq z_{(k)} \leq \dots \leq z_{(N_c)}$ the ordered cluster maxima. We use the quantile normalized root mean square error

$$qnrmsc = \sqrt{\frac{1}{N_c} \sum_{k=1}^{N_c} \left(\frac{q_t(p_k) - q_e(p_k)}{q_t(p_k)} \right)^2},$$

where $p_k = (k-1/2)/N_c$, q_t is the quantile of the estimated GPD, and q_e is the empirical quantile. The computation of the empirical quantile is distribution free. It is based on the modal position (see definition 7 of Hyndman and Fan (1996)).

2.5 Extremal dependence measure

Hourly precipitation can be expected to exhibit dependence, even between its less frequently occurring extreme values. Independent information on the duration of this dependence at a particular location can be elicited by means of a pair of dependence measures denoted by $(\chi, \bar{\chi})$, provided by bivariate extreme value theory. The pair is designed specifically to express dependence at extreme levels (Coles et al. 1999; Coles 2001). Let (X, Y) be a two-dimensional random vector, such that the marginal distributions of X and Y are identical. The dependence measure χ provides information on asymptotic dependence, and it is defined as the limit, as the threshold u rises, of the probability that Y also exceeds u if X exceeds u . It can take values between 0 and 1, and vanishes if X and Y are asymptotically independent.

The second dependence measure $\bar{\chi}$ describes the dependence between asymptotically independent random variables.

It takes values between -1 and 1 . For asymptotically dependent variables, $\bar{\chi} = 1$, and for independent variables, $\bar{\chi} = 0$; the sign of $\bar{\chi}$ is positive (negative) when an increase (decrease) in X tends to correspond to an increase in Y . For asymptotically independent variables, $\bar{\chi}$ increases with the strength of dependence at moderately extreme levels. These dependence measures must be used in combination. If $\chi = 0$, X and Y are asymptotically independent, and $\bar{\chi}$ must be used to evaluate dependence at moderately extreme levels.

In order to evaluate the dependence between extreme occurrences of hourly precipitation, χ and $\bar{\chi}$ are estimated empirically—the data is transformed to the uniform distribution with the empirical distribution function—for lagged exceedances over a range of thresholds. We estimate $\bar{\chi}$ at different lags of hourly precipitation for the 99th percentile of the full data set. This corresponds to quantiles of non-zero values between 92 and 93 %, depending on the season, and offers an indication of extremal dependence at the thresholds used for the IMT statistic. As a crude estimate of the lag at which dependence between exceedances can be expected to die out, we will use the first lag at which the lower confidence bound of $\bar{\chi}$ intersects the $\bar{\chi} = 0$ line, which we will call the maximum dependent lag. The confidence intervals are estimated with the delta method, and rely on assumptions such as the independence of the observations. They are therefore likely to be much too narrow, and the results should be interpreted with caution.

2.6 Data

The data used in the present study consist of hourly precipitation observations at 59 stations of the meteorological network of the Swiss Federal Office of Meteorology and Climatology (MeteoSwiss). Their geographical and cross-Alpine altitudinal distributions are shown in Fig. 1.

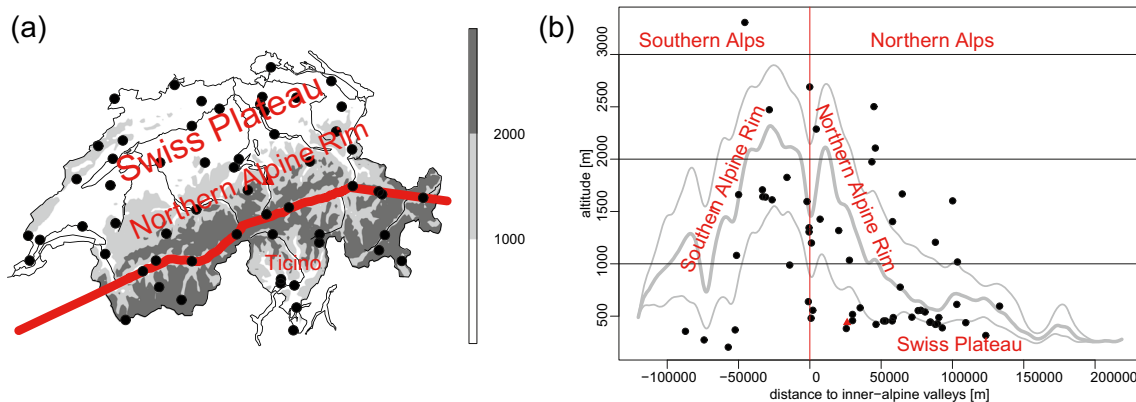


Fig. 1 Location of the stations considered in the present study (black dots; red triangle: station Altdorf) on a map of Switzerland (a) and in a vertical cross-section across the Alpine ridge (b). The latter is defined as the distance to the inner-Alpine valleys (thick red line in a). The profile of the Alps is computed with the USGS-GTOPO30 (<http://eros.usgs.gov>)

digital elevation model, and is the minimum air–line distance of each grid-point center to the inner-Alpine valleys. The thick (thin) gray line(s) represent(s) the smoothed median (10 and 90 % quantiles) of the distances in 100-m bins

Stationarity of the parent time series is an essential prerequisite for valid application of the GPD. Thus, only stations that suffered neither a change in instrument type nor a displacement in the period of interest were considered. Note that the data underwent a standard quality check, but were not homogenized. At a small number of stations, trends in the 90 % quantile of non-zero values were detected with a seasonally applied Mann-Kendall test (Mann 1945; Kendall 1948). In such cases, the linear trend was estimated with the Theil-Sen estimator, and subsequently eliminated (Theil 1950; Sen 1968).

The length of the data sets is variable, between 20 and 30 consecutive years, over the period 1981 and 2010. The analysis presented here concentrates on subsets of the original data corresponding to the four seasons. Winter is defined as December–January–February (DJF), spring as March–April–May (MAM), summer as June–July–August (JJA), and autumn as September–October–November (SON).

3 Results

3.1 IMT selection and extremal dependence

IMT selection is applied to hourly precipitation at the 59 selected stations, but is first presented here at one station in detail. The plausibility of the selection in terms of mean cluster size—which can be seen as a summary quantity for threshold and run parameter—is then examined at all stations. As it represents the average length of events, we compare it to the maximum dependent lag (see Section 2.5). Finally, the ensuing GPD estimates are compared with those obtained with the reference method (see Section 2.3).

The IMT selection is illustrated in Fig. 2 for winter and summer hourly precipitation at the station Altdorf, located in Central Switzerland in a valley of the northern Alpine rim. The IMT results for all pairs are represented as a two-dimensional surface for DJF and JJA in Fig. 2a, b. The red dots are the (u, K) pairs selected by the algorithm. The “top left” corner of the surface corresponds to the reference pair.

In winter, pairs with low u and low K lead to rejection at the 5 % confidence level. The IMT appears to favor either low u and high K , or high u and low K (Fig. 2a). This particularity is typical for the winter season (not shown). In summer, a pair with a low value for K is selected, while the combination of low u and high K yield inter-cluster times incompatible with the assumptions of a point process (Fig. 2b). All stations but a few have a similar IMT surface in summer (not shown).

Figure 2c, d show the 50-year return levels computed for all (u, K) pairs at station Altdorf in winter and summer. Clearly, the return levels of all pairs are within a narrow range, especially in winter. In the subset of pairs with $IMT < 0.05$ with a

final number of clusters greater than 80 (red segments in Fig. 2c, d; black dots in Fig. 2a, b), all return levels are within the interval between the highest lower confidence bound and the lowest upper confidence bound (blue triangles). In fact, there are only a few pairs, and only in winter, for which the return levels are outside of this interval, and therefore differ significantly from those in the subset. Thus, the estimates are stable, and the selection has little influence on the return levels. In both seasons, the reference pair ($u=0.90, K=120$) yields return levels that are among the lowest of all pairs, but still within the confidence bounds.

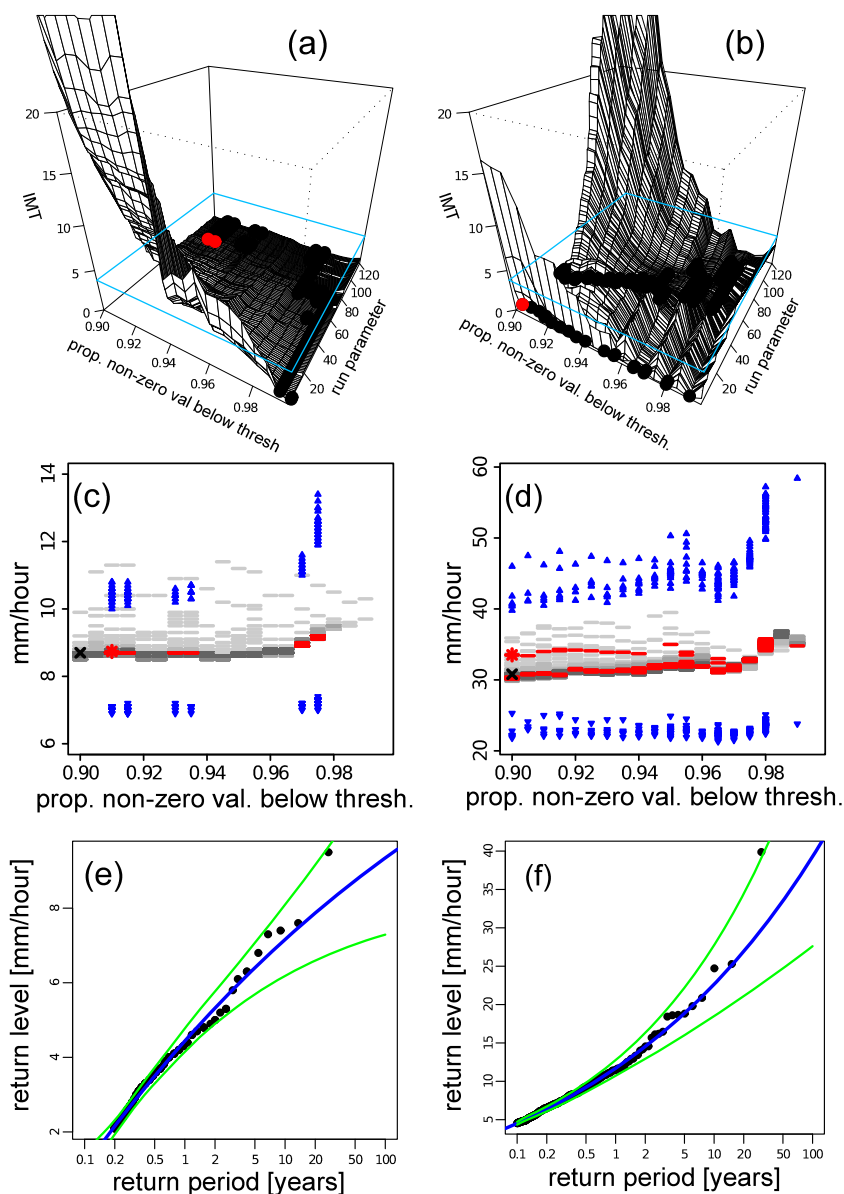
The return level plots for the selected (u, K) pair are displayed in Fig. 2e, f. The estimated distribution is bounded in winter and heavy-tailed in summer. This illustrates the difference between winter and summer at stations in the northern Alps, where all stations but one (75 % of stations) have a significantly negative (positive) shape parameter in winter (summer) (not shown).

For all stations, the mean cluster size is represented for the four seasons in Fig. 3a. It is smallest in summer and largest in winter, and consistently larger to the south than to the north of the inner-Alpine valleys. Likewise, the station to station variability increases from summer to winter with intermediate values in spring and autumn; it is also generally greater in the southern Alps. Both the seasonal variation, and the north–south differences are mirrored in the maximum dependent lag—a measure for the longest lag at which dependence might still be expected—derived from the extremal dependence measure (Fig. 3b).

At each station, the IMT selection yields a different threshold–run parameter pair (u, K) . These vary from season to season, and differ between northern and southern Alps. In winter, thresholds vary between 2–4 mm, with nearly all southern stations below 2 mm, while in summer values go from over 2 to nearly 10 mm/h (not shown). Run parameters are approximately 60–80 (20–40)h in the northern (southern) Alps in winter, and about 10–15 (10–20)h in the northern (southern) Alps in summer (not shown).

The performance in terms of $qnrmsc$ of the IMT selection vs. the reference selection is shown in Fig. 4a. In all seasons, the IMT pair generally leads to a better fit of the GPD estimates than the reference pair. Since the IMT selection allows for lower run parameters than the reference selection, this suggests that unnecessarily large run parameters may be harmful for the subsequent estimation of the GPD parameters. Even in winter, when the difference seems moderate, the $qnrmsc$ is smaller for the IMT pair than for the reference pair at 70 % of the stations. As can be seen in Fig. 4b, the reference method leads to IMT values that would strongly suggest rejection at the 5 % level at a majority of stations in all seasons except in winter. In other words, the reference method

Fig. 2 Analysis of hourly precipitation at station Altdorf (altitude 438 m; distance from inner-Alpine valleys=26 km) for winter (left) and summer (right). *Top* IMT statistic. *x-axis* threshold u as the proportion of non-zero values that are below it. *y-axis* run parameter K in hours. *Blue line* 5 % critical value $\chi^2_1(0.95)=3.84$. *Black dots* (u, K) pairs for which $IMT < 0.05$ (p value=0.82). *Red dots* IMT pair, i.e., (u, K) pair with the largest number of clusters. *Middle* 50-year return levels in millimeter per hour vs. u for all (u, K) pairs yielding more than 80 clusters (*gray segments*). The *darker shades* indicate higher K . *Red segments* pairs for which the IMT value is below 0.05. *Red star* selected (u, K) pair. *Black cross* reference (u, K) pair. *Blue point up (down) triangles* upper (lower) confidence bounds of pairs with $IMT < 0.05$. *Bottom* Return level in millimeter/hour vs. return period. *Blue* best estimate. *Green* 95 % confidence intervals. *Black dots* observations (ordinate) vs. empirical return period (abscissa)



leads to a series of exceedances with a configuration of clusters and inter-cluster times that can only poorly be represented by the distribution required by the two-dimensional process limit to the extremal process.

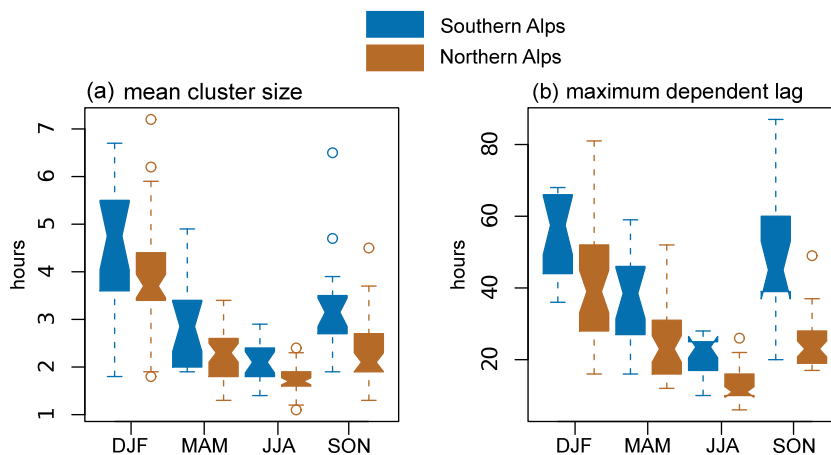
The 100-year return levels are generally higher for the IMT pair than for the reference pair, especially in summer (Fig. 4c). The negligible differences in winter can be attributed to the fact that the selected run parameters are rather large, and thus close to the reference run parameter. On average, however, the difference is rather small, although it may be substantial at individual stations. Particularly in July, it can reach 10 to 30 % at half of the stations, but the differences rarely exceed the range of the confidence intervals (not shown). Note that—as a quantitative summary of the QQ plot—the $qnrms$ is a goodness-of-fit measure that is entirely unrelated to the IMT

selection of threshold and run parameter, and therefore constitutes an independent assessment of the automation method proposed.

3.2 Climatology of extreme hourly precipitation

In this section, we consider the results from a climatological point of view, and examine the seasonal cycle and geographical patterns of extreme hourly precipitation in terms of its severity, seasonal frequency, and duration. The severity is best described by the return levels for a given return period, while the frequency can be gleaned from the mean number of clusters per season. For information on event duration, we turn to the dependence measure.

Fig. 3 Boxplots of the mean cluster size (left) and maximum dependent lag (right), both in hours, in the four seasons, for the southern (blue) and northern (brown) Alps (see Fig. 1 for definition)



The 50-year return levels represented for the four seasons in Fig. 5 show that intense hourly precipitation in Switzerland experiences its minimum in winter (a) and its maximum in summer (c). Winter is characterized by a rather narrow range of low return levels without distinctive spatial pattern. The inner-Alpine valleys, including the Inn valley, remain at low levels all year round. In the northern Alps, the Plateau witnesses an increase in the number of stations with higher return levels in spring (b). In summer, high return levels cover the entire Plateau, as well as the northern Alpine rim with slightly weaker values. Thus, there is a slight increase in the return values towards the Plateau with the distance to the Alpine ridge. In autumn (d), the severity of events wanes to a level hardly higher than in winter. In the southern Alps, the Ticino stands out with comparatively high return levels from March to November, especially in autumn, when the contrast with the north is most pronounced.

The vertical cross-section of the 50-year return levels across the Alpine ridge is shown in Fig. 6 for winter and summer. In the northern Alps, return levels increase with

height in winter. In summer, they increase with the distance to the inner-Alpine valleys. No altitudinal effect can be detected in the southern Alps in winter. In summer, although the return levels at the five stations in the southern part of the Ticino are nearly twice the return levels at the other southern stations, there are not enough stations to draw any conclusions about the altitudinal effect.

The seasonal frequency of events (see Fig. 7), i.e., the number of clusters per season, is low from September to May, generally not exceeding eight events in the season. In summer, it increases all over Switzerland, and the highest frequencies (exceeding 14 events per season) are found along the northern Alpine rim. It is noteworthy that the southern Alps, but especially the Ticino, displays lower frequencies than in the north in all seasons.

The average duration of individual events and the long-term dependence of intense or heavy hourly precipitation are shown in Fig. 3a, b. Extreme hourly precipitation turns out to be asymptotically independent, even at lags of 1 h, i.e., $\hat{\chi} = 0$ (not shown). Thus, the quantity displayed here (b) is derived

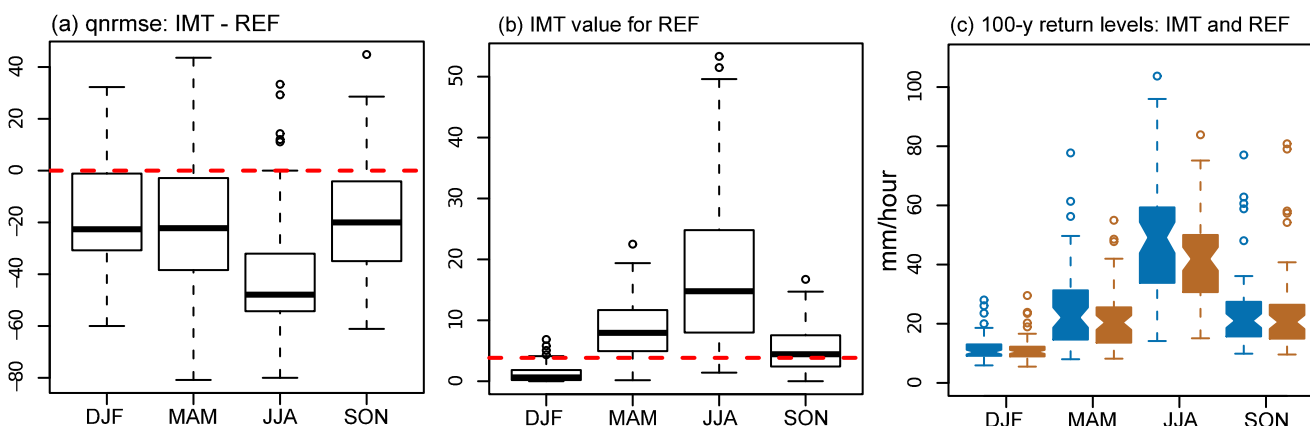


Fig. 4 Boxplots in winter, spring, summer, and autumn for all stations of a: the difference in quantile normalized root mean square error (qnrmse) between IMT and reference method (negative values indicate better

performance for the IMT method); b: the IMT values for the reference pair; c: the 100-year return levels of the IMT (blue) and reference (brown) methods

from $\widehat{\chi}$, the dependence at subasymptotic levels. Events last approximately 4 h in winter and less than 2 h in summer, and about 2–3 h in spring and autumn (a). Dependence between exceedances at different lags subsists for more than 40 h in winter, less than 30 h in spring and autumn, and only about 15 h in summer (b). To the south of the inner-Alpine valleys, the dependence lasts consistently longer than in the northern Alps, a fact reflected in the mean cluster size (a). Particularly in summer and autumn, the difference is considerable, with dependence lasting 10–15 h (20 to 30 h) in the north (south), and about 25 h (50 h) in the north (south), respectively.

4 Discussion

4.1 IMT selection and GPD estimates

The comparison of the IMT selection with the reference selection (Fig. 4) highlights the fact that it may be harmful to select an unnecessarily large run parameter, as shown by Fawcett and Walshaw (2007). Hourly precipitation would be most strongly affected in winter, when pairs with low run parameters tend to be rejected. The IMT selection yields better GPD estimates than the reference method because it looks for pairs leading to the largest possible number of clusters. As this number

increases roughly from the “upper right” corner ($u=0.995, K=120$) to the “lower left” corner ($u=0.90, K=1$), the IMT selection will automatically select a pair with a smaller run parameter than the reference pair. The disadvantage of this method is the tendency to choose rather low thresholds, thus introducing the possibility of a bias in the estimates. This might not be of great consequence in winter, when the physical processes involved are probably the same, regardless of the amplitude of the excesses. In the other seasons, on the contrary, the most extreme events and the moderate ones are likely to originate from different processes.

Extreme value theory assumes that the parent distribution is stationary. Like most climatic variables, however, precipitation undergoes an annual cycle. In the present work, seasonality is taken into account by dividing the year in 3-month bins, and considering them separately, rather than modeling it explicitly, as done in several recent studies (Katz et al. 2002; Maraun et al. 2009; Rust et al. 2009; Umbricht et al. 2013). No attempt was made to account for the daily cycle of hourly precipitation. While its amplitude is negligible in winter, the diurnal cycle experiences a maximum in the late afternoon, followed by a gentle decrease over the next 15 h (Wüest et al. 2010). Analysis of the empirical quantiles at the stations used here confirmed this behavior.

Fig. 5 Fifty-year return levels for winter (a), spring (b), summer (c), and autumn (d). The color scale codes the return levels in millimeter/hour, and is the same for all seasons. The largest (smallest) dots correspond to the largest (smallest) return levels in the respective season

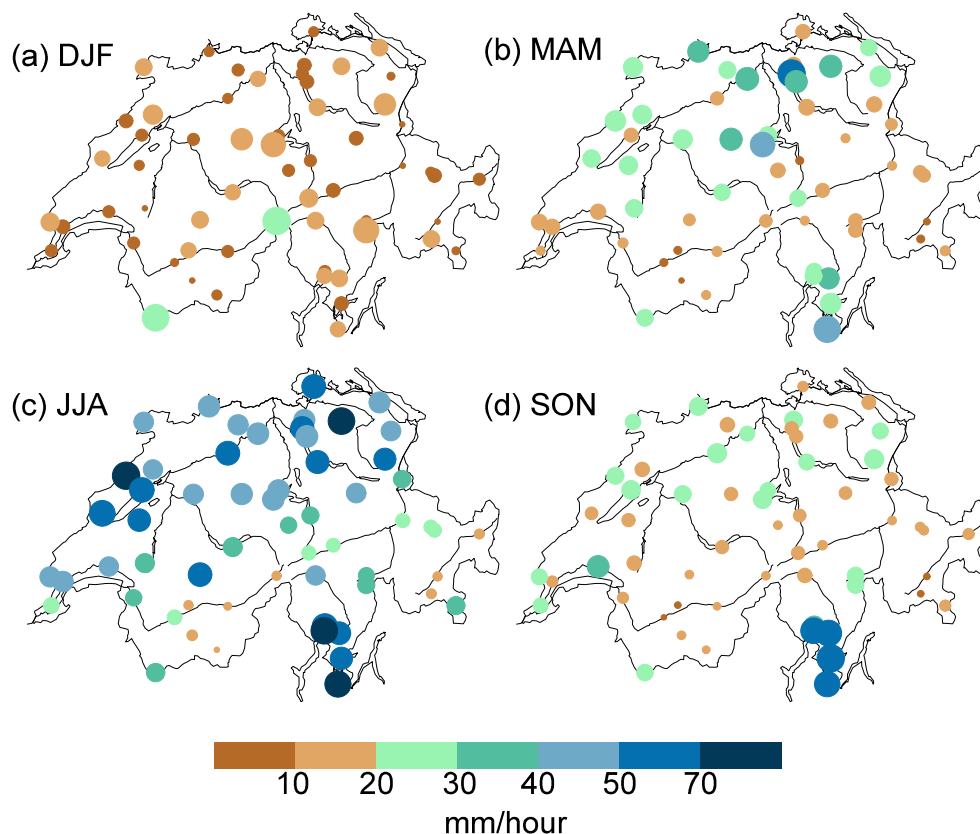
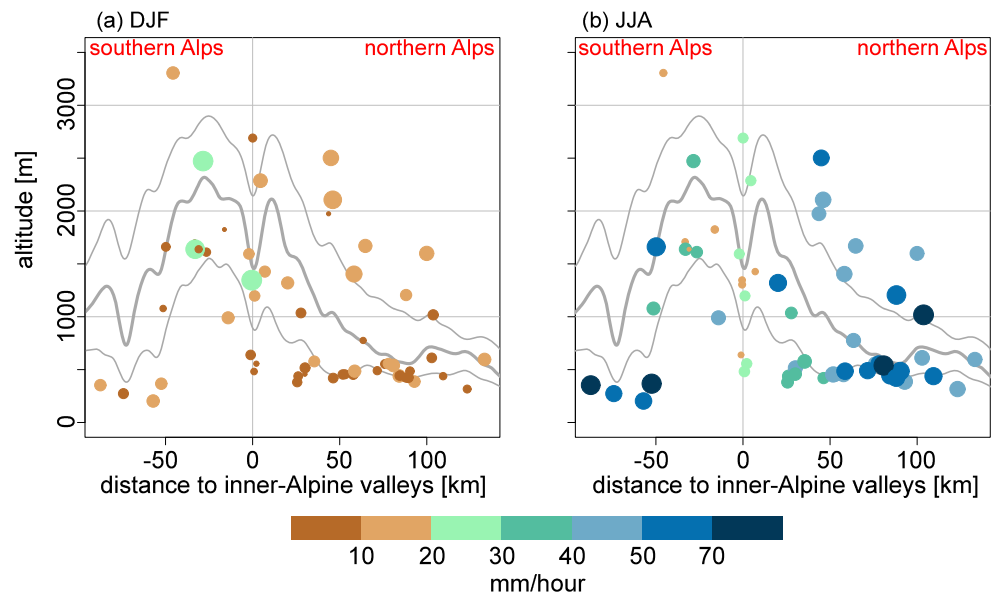


Fig. 6 Vertical cross-section of the 50-year return levels across the Alpine ridge for winter (a) and summer (b). The *color scale* codes the return levels in millimeter/hour, and is the same for both seasons. The *largest (smallest) dots* correspond to the largest (smallest) return levels in the respective season

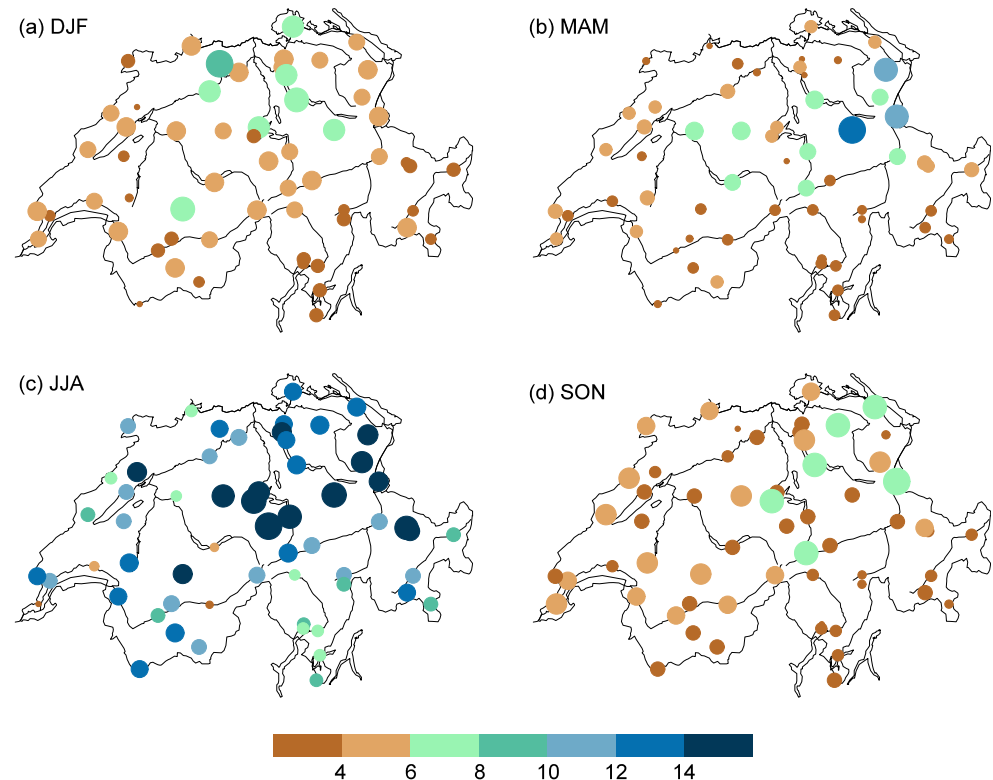


For precipitation, the form of the tail appears to vary with duration (Buishand 1991; Pearson and Henderson 1998), geographical location (Revfeim 1982; Buishand and Demaré 1990; Pearson and Henderson 1998; Friederichs 2010; Toreti et al. 2010; Maraun et al. 2011), and altitude (Pearson and Henderson 1998; Cooley et al. 2007; Gardes and Girard 2010). For daily

precipitation, the shape parameter appears to be mostly light or heavy-tailed.

It turns out, however, that the shape parameter of extreme hourly precipitation is significantly positive (negative) in the northern Alps in summer (winter) (not shown). A variation of the shape parameter with altitude could not be detected. In summer, it increases from the northern Alpine rim towards the

Fig. 7 Map of the mean number of clusters per season in winter (a), spring (b), summer (c), and autumn (d). The *color scale* represents the number of clusters per season, and is the same for all seasons. The *largest (smallest) dots* correspond to the largest (smallest) number of clusters in the respective season



Plateau, as do the return levels (not shown). This is consistent with the higher seasonal frequencies along the northern Alpine rim than in the plain.

The tendency towards more negative shape parameters in winter may, to some extent, be explained by the microphysical processes taking place. Winter precipitation is stratiform, and the precipitation particles form essentially through vapor deposition. The upward motion must remain weak to allow the droplets to fall (Houze 1997). This sets an intrinsic upper limit to the hourly precipitation rate in winter. As a result, observations of winter hourly precipitation in Switzerland are confined to a narrow interval, and the highest values rarely stray far from the body of the distribution.

4.2 Climatology

The signature of the Alps can be seen in the annual mean precipitation (Frei and Schar 1998; Isotta et al. 2013). A marked wet anomaly covers the northern Alpine rim, and most of the southern rim, while the inner-Alpine valleys, in particular the Rhone valley in the southwest and the Inn Valley in the Grisons in the southeast of Switzerland, are dry. The Ticino, on the steep southern rim, is host to the heaviest precipitation (Frei and Schmidli 2006; Isotta et al. 2013).

Some of these features are reflected in the findings of the present study. Winter return levels increase somewhat with height, as might be expected, given the role of orographic precipitation in the Alpine region. The inner-Alpine valleys witness few events of intense hourly precipitation, even in summer. In addition, the return levels there generally remain very low. The low frequency of summer thunderstorms in these deep valleys is attributed to inadequate moisture flux convergence, and lack of a lifting source (van Delden 2001).

The increased number of events in summer along the northern Alpine rim mirrors the thunderstorm path as represented by the lightning climatology (MeteoSwiss, personal communication). It is also reminiscent of the frequency of observations exceeding 10 mm/h in the analysis of reconstructed hourly precipitation by Wüest et al. (2010). Note, however, that these quantities are not directly comparable, since a cluster contains several observations, and the thresholds differ from station to station.

Finally, the Ticino displays comparatively high return levels from March to November, but stands out particularly in autumn with return levels twice to three times as large as in the rest of Switzerland. This can be attributed to southerly flow impinging on the Alps as the midlatitude cyclones reach further south with the approach of winter. These can lead to violent precipitation as the warm humid Mediterranean air is forced upwards over a short distance, rapidly reaching the level of free convection (Gheusi and Davies 2004, and references therein).

In contrast to the annual mean, the pattern of return levels of heavy hourly precipitation in summer does not disclose the structure of the Alps. In fact, the return levels experience a slight increase from the northern Alpine rim towards the Swiss Plateau. This may be explained by the thunderstorm formation and propagation. Thunderstorms form in squall lines ahead of cold fronts (Haase-Straub et al. 1994), or in response to thermally driven topographic flow (Langhans et al. 2013). Linder et al. (1999) identify the Jura and the northern Alpine rim as regions of genesis for convective cells. However, several studies show that these drift towards adjacent flat areas (Finke and Hauf 1996; Bertram and Mayr 2004), such as, in this case, the Swiss Plateau.

Given the fact that hourly precipitation results from processes of varying time scales, we can expect the observations to be dependent over a certain time interval. While synoptic systems take a few days to sweep over Switzerland, convective cells have a lifetime extending from 1 h or less for single cells up to 12 h for supercells (Bertram and Mayr 2004). Of course, they are generally not stationary, and a single station may be affected only over a much shorter time. In the present study, dependence was examined in all seasons north and south of the inner-Alpine valleys, and these climatic characteristics were found to be reflected in the seasonal variation both of the maximum dependent lag and the mean cluster size. The values for dependence in the northern Alps are in accordance with the study by Huser and Davison (2013), who detected dependence of hourly precipitation at extreme levels of the order of 10–15 h in summer at a selection of stations in the Jura and the Swiss Plateau. Both dependence and average duration of events exhibit strong spatial variability in winter (see Fig. 3), despite the fact that winter events are dictated by large-scale midlatitude cyclones. It is noteworthy that dependence itself is consistently larger in the southern Alps. The associated variability from station to station is larger from May to November on the southern side of the Alps, especially in autumn, pointing perhaps to different dominant processes at different stations.

5 Summary and conclusions

In the present paper, we propose an automated procedure for the selection of threshold and run parameter in the peaks-over-threshold (POT) approach to extreme value analysis based on the graphical method developed by Süveges and Davison (2010). The automated procedure sets aside a subset of non-rejectable threshold–run parameter pairs and, in this subset, selects the pair that generates the largest number of clusters. We apply it to hourly precipitation in Switzerland in the period 1981–2010.

The tendency of extreme events to cluster indicates underlying dependence in the data. In particular, dependence

between high exceedances should be reflected in the mean cluster size. In order to put our findings into context, lag dependence of hourly precipitation at extreme levels was computed with the help of the dependence measures by Coles et al. (1999).

Applied to hourly precipitation in Switzerland, the misspecification test brings to light typical seasonal structures in the inter-exceedance times. In winter, combinations of low thresholds and low run parameters, leading to relatively small mean cluster sizes, are rejected. In summer, strong misspecification arises for combinations of low thresholds and high run parameters, corresponding to the largest mean cluster sizes. In this context, the automatic selection picks threshold–run parameter pairs that yield mean cluster sizes in accordance with the seasonal characteristics of the separately estimated dependence at extreme levels.

The GPD estimates based on peaks-over-threshold analysis of the cluster maxima resulting from the automated selection highlight many known features regarding precipitation in the Alpine region. It is noteworthy that in summer, the signal due to thunderstorm activity is visible in the seasonal frequency of events, rather than in their severity, as represented by return levels for high return periods.

The present study exemplifies the argument by Fawcett and Walshaw (2007) that unnecessarily large run parameters have adverse consequences on subsequent estimation of the GPD parameters. Compared to a reference selection with fixed threshold and run parameter, the IMT selection allows for lower run parameters. This leads to higher return levels, and in practice to more stringent design measures, a positive outcome considering the uncertainty associated with planning long-term structures based on only 30 years of data. Finally, the patterns of extreme hourly precipitation suggest that requirements for the observational network may be different for hourly than for daily precipitation.

Acknowledgments We are indebted to Juliette Blanchet for her help with the dependence measures and the validation procedure, and to Anthony Davison for his very helpful advice. We are grateful to Pier Luigi Vidale for his detailed comments on the manuscript. We thank Stephan Bader and Thomas Schlegel for sharing their expertise on Alpine climatology, and Christoph Frei for his helpful discussions and insightful questions that greatly contributed to improving this paper. Finally, we thank Anne Schindler for her critical view and invaluable advice, regarding both form and content, which helped shape the paper into its final form. We also thank Christoph Frei for putting the R-package `gevXgpd` at our disposal for estimation of the GPD parameters. The dependence measures were estimated with `chplot` by Jan Heffernan and Alec Stephenson, which is part of the R-package `evd`.

Appendix

The equations appearing here can be found in the appendix of Süveges and Davison (2010). The derivation of the original

paper has been corrected for errata, and the notations made conform to those in Section 2.3. We assume a stationary sequence X_1, \dots, X_n with N observations exceeding the threshold u , and therefore $N - 1$ inter-exceedance times. Let K denote the run parameter, N_c the number of clusters, θ the extremal index, and $c_i^{(u,K)}$ ($i = 1, \dots, N - 1$) the inter-exceedance times.

Then, the log-likelihood function is given by:

$$\ell_{(u,K)}(\theta; c_i^{(u,K)}) = ((N-1) - N_c) \log(1-\theta) + 2N_c \log \theta - \theta \sum_{i=1}^{N-1} c_i^{(u,K)}.$$

Let ℓ_i, j_i, i_i, d_i denote for a single observation i : the log-likelihood, the score function, the expected information, and the difference between score function and expected information, respectively. Let the derivative with respect to θ be denoted by a prime. Let $\mathbb{I}(A)$ be the indicator function for the set A . Then, for a given (u, K) pair,

$$\ell'_i(\theta) = -\frac{\mathbb{I}(c_i^{(u,K)} = 0)}{(1-\theta)} + \frac{2\mathbb{I}(c_i^{(u,K)} > 0)}{\theta} - c_i^{(u,K)},$$

$$j_i(\theta) = \frac{\mathbb{I}(c_i^{(u,K)} = 0)}{(1-\theta)^2} + \frac{4\mathbb{I}(c_i^{(u,K)} > 0)}{\theta^2} + (c_i^{(u,K)})^2 - \frac{4c_i^{(u,K)}}{\theta},$$

$$i_i(\theta) = \frac{\mathbb{I}(c_i^{(u,K)} = 0)}{(1-\theta)^2} + \frac{2\mathbb{I}(c_i^{(u,K)} > 0)}{\theta^2},$$

$$d_i(\theta) = \frac{2\mathbb{I}(c_i^{(u,K)} > 0)}{\theta^2} + (c_i^{(u,K)})^2 - \frac{4c_i^{(u,K)}}{\theta},$$

$$d'_i(\theta) = -\frac{4\mathbb{I}(c_i^{(u,K)} > 0)}{\theta^3} + \frac{4c_i^{(u,K)}}{\theta^2}.$$

Let $D(\theta) = (N-1)^{-1} \sum_{k=1}^{N-1} d_k(\theta)$ and $I(\theta) = (N-1)^{-1} \sum_{k=1}^{N-1} i_k(\theta)$ denote the sample means of d_i and i_i . The sample variance of $D(\theta)$ is:

$$V(\theta) = (N-1)^{-1} \sum_{k=1}^{N-1} \left\{ d_k(\theta) - D(\theta) I(\theta)^{-1} \ell'_k(\theta) \right\}^2.$$

The Information Matrix Test (IMT) Statistic is then:

$$IMT(\hat{\theta}) = nD(\hat{\theta})^2 V(\hat{\theta})^{-1},$$

where θ has been replaced by the estimated value of $\hat{\theta}$.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Balkema AA, de Haan L (1974) Residual life time at great age. *Ann Probab* 2:792–804
- Begert M (2008) Die Repräsentativität der Stationen im Swiss National Basic Climatological Network (Swiss NBCN). 217:40
- Beirlant J, Goegebeur Y, Segers J, Teugels J (2004) Statistics of extremes: theory and applications. John Wiley, John Wiley & Sons, Chichester
- Bertram I, Mayr GJ (2004) Lightning in the eastern Alps 1993 – 1999, part I: Thunderstorm tracks. 4:501–511
- Borga M, Anagnostou EN, Blöschl G, Creutin J-D (2010) Flash floods: observations and analysis of hydro-meteorological controls. *J Hydrol* 394:1–3. doi:10.1016/j.jhydrol.2010.07.048
- Buishand TA (1991) Extreme rainfall estimation by combining data from several sites. *Hydrol Sci J* 36:345–365
- Buishand TA, Demaré GR (1990) Estimation of the annual maximum distribution from samples of maxima in separate seasons. *Stoch Hydrol Hydraul* 4:89–103
- Coles S (2001) An introduction to statistical modeling of extreme values. Springer-Verlag, London, p 210
- Coles S, Heffernan J, Tawn J (1999) Dependence measures for extreme value analyses. *Extremes* 2:339–365
- Cooley D, Nychka D, Naveau P (2007) Bayesian spatial modeling of extreme precipitation return levels. *J Am Stat Assoc* 102:824–840. doi:10.1198/016214506000000780
- Davison AC (2008) Statistical models. Cambridge University Press, Cambridge, p 738
- Davison AC, Smith RL (1990) Models for exceedances over high thresholds. *J R Stat Soc Ser B* 52:393–442
- Dupuis D (1998) Exceedances over high thresholds: a guide to threshold selection. *Extremes* 1:251–261
- Fawcett L, Walshaw D (2007) Improved estimation for temporally clustered extremes. *Environmetrics* 18:173–188. doi:10.1002/env.810
- Ferro CAT, Segers J (2003) Inference for clusters of extreme values. *J R Stat Soc B Stat Methodol* 65:545–556
- Finke U, Hauf T (1996) The characteristics of lightning occurrence in southern Germany. *Beitr Phys Atmosph* 69:361–374
- Fisher RA, Tippett LHC (1928) Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proc Camb Philos Soc* 24:180–190
- Frei C, Schar C (1998) A precipitation climatology of the Alps from high-resolution rain-gauge observations. *Int J Climatol* 18:873–900
- Frei C, Schmidli J (2006) Das Niederschlagsklima der Alpen: Wo sich Extreme nahe kommen. *promet* 32:61–67
- Friederichs P (2010) Statistical downscaling of extreme precipitation events using extreme value theory. *Extremes* 13:109–132. doi:10.1007/s10687-010-0107-5
- Frigessi A, Haug O, Rue H (2002) A dynamic mixture model for unsupervised tail estimation without threshold selection. *Extremes* 5:219–235
- Gardes L, Girard S (2010) Conditional extremes from heavy-tailed distributions: an application to the estimation of extreme rainfall return levels. *Extremes* 13:177–204. doi:10.1007/s10687-010-0100-z
- Geiger H, Zeller J, Röhrlisberger G (1991) Starkniederschläge des schweizerischen Alpen- und Alpenrandgebietes. Einführung, Methoden, Spezialstudien. Band 7:334
- Gheusi F, Davies HC (2004) Autumnal precipitation distribution on the southern flank of the Alps: A numerical-model study of the mechanisms. *Q J R Meteorol Soc* 130:2125–2152. doi:10.1256/qj.03.06
- Gnedenko B (1943) Sur la distribution limite du terme maximum d'une série aléatoire. *Ann Math* 44:423–453
- Guzzetti F, Peruccacci S, Rossi M, Stark CP (2007) The rainfall intensity-duration control of shallow landslides and debris flows: an update. *Landslides* 5:3–17. doi:10.1007/s10346-007-0112-1
- Haase-Straub SP, Heimann D, Hauf T, Smith ERK (1994) The squall line of 21 July 1992 in Switzerland and southern Germany—a documentation. DLR Forschungsbericht 94–18, DLR
- Houze RAJ (1997) Stratiform precipitation in regions of convection: a meteorological paradox? *Bull Am Meteorol Soc* 78:2179–2196
- Hsing T (1987) On the characterization of certain point processes. *Stoch Process Appl* 26:297–316
- Huser R, Davison AC (2013) Space-time modelling of extreme events. *J R Stat Soc - Ser B* in press
- Hyndman RJ, Fan Y (1996) Sample quantiles in Statistical Packages. *Am Stat* 50:361–365
- Isotta FA, Frei C, Weigluni V, et al. (2013) The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data. *Int J Climatol* in press
- Katz RW, Parlange MB, Naveau P (2002) Statistics of extremes in hydrology. *Adv Water Resour* 25:1287–1304. doi:10.1016/S0309-1708(02)00056-8
- Kendall MG (1948) Rank correlation methods. Charles Griffin, London
- Koutsoyiannis D, Kozonis D, Manetas A (1998) A mathematical framework for studying rainfall intensity-duration-frequency relationships. *J Hydrol* 206:118–135
- Langhans W, Schmidli J, Fuhrer O et al (2013) Long-term simulations of thermally driven flows and orographic convection at convection-parameterizing and cloud-resolving resolutions. *J Appl Meteorol Climatol* 52:1490–1510. doi:10.1175/JAMC-D-12-0167.1
- Linder W, Schmid W, Schiesser H-H (1999) Surface winds and development of thunderstorms along southwest-northeast oriented mountain chains. *Weather Forecast* 14:758–770
- Mann HB (1945) Nonparametric tests against trend. *Econometrica* 13: 245–259
- Maraun D, Rust HW, Osborn TJ (2009) The annual cycle of heavy precipitation across the United Kingdom: a model based on extreme value statistics. *Int J Climatol* 1744:1731–1744. doi:10.1002/joc.1811
- Maraun D, Osborn TJ, Rust HW (2011) The influence of synoptic airflow on UK daily precipitation extremes. Part I: Observed spatio-temporal relationships. *Clim Dyn* 36:261–275. doi:10.1007/s00382-009-0710-9
- Palutikof J, Brabson B, Lister D, Adcock S (1999) A review of methods to calculate extreme wind speeds. *Meteorological* 6: 119–132
- Pearson CP, Henderson RD (1998) Frequency distributions of annual storm rainfalls in New Zealand. *J Hydrol* 37:19–33
- Pickands J (1975) Statistical inference using extreme order statistics. *Ann Stat* 3:119–131
- Revfeim K (1982) Seasonal patterns in extreme 1-hour rainfalls. *Water Resour Res* 18:1741–1744
- Rust HW, Maraun D, Osborn TJ (2009) Modelling seasonality in extreme precipitation. A UK case study. *Eur Phys J Spec Top* 174:99–111. doi:10.1140/epjst/e2009-01093-7

- Sen PK (1968) Estimation of the regression coefficient based on Kendall's tau. *J Am Stat Assoc* 63:1379–1389
- Süveges M, Davison AC (2010) Model misspecification in peaks over threshold analysis. *Ann Appl Stat* 4:203–221. doi:[10.1214/09-AOAS292](https://doi.org/10.1214/09-AOAS292)
- Theil H (1950) A rank-invariant method of linear and polynomial regression analysis. I, II, III. *Ned Akad van Wet* 53:386–392, 521–525, 1397–1412
- Todorovic P, Zelenhasic E (1970) A stochastic model for flood analysis. *Water Resour Res* 6:1641–1648
- Toreti A, Xoplaki E, Maraun D et al (2010) Characterisation of extreme winter precipitation in Mediterranean coastal sites and associated anomalous atmospheric circulation patterns. *Nat Hazards Earth Syst Sci* 10:1037–1050. doi:[10.5194/nhess-10-1037-2010](https://doi.org/10.5194/nhess-10-1037-2010)
- Toreti A, Schneuwly-Bollschweiler M, Stoffel M, Luterbacher J (2013) Atmospheric forcing of debris flows in the southern Swiss Alps. *J Appl Meteorol Climatol* 52:1554–1560. doi:[10.1175/JAMC-D-13-077.1](https://doi.org/10.1175/JAMC-D-13-077.1)
- Umbrecht A, Fukutome S, Liniger MA et al (2013) Seasonal variation of daily extreme precipitation in Switzerland. *Sci Rep MeteoSwiss* 97:124pp
- Van Delden A (2001) The synoptic setting of thunderstorms in western Europe. *Atmos Res* 56:89–110. doi:[10.1016/S0169-8095\(00\)00092-2](https://doi.org/10.1016/S0169-8095(00)00092-2)
- Wadsworth JL, Tawn JA (2012) Likelihood-based procedures for threshold diagnostics and uncertainty in extreme value modelling. *J R Stat Soc B Stat Methodol* 74:543–567. doi:[10.1111/j.1467-9868.2011.01017.x](https://doi.org/10.1111/j.1467-9868.2011.01017.x)
- Walshaw D (1994) Getting the most from your extreme wind data: a step by step guide. *J Res Natl Inst Stand Technol* 99:399–411
- White H (1982) Maximum likelihood estimation of misspecified models. *Econom J Econom Soc* 50:1–25
- Wüest M, Frei C, Altenhoff A et al (2010) A gridded hourly precipitation dataset for Switzerland using rain-gauge analysis and radar-based disaggregation. *Int J Climatol* 30:1764–1775. doi:[10.1002/joc.2025](https://doi.org/10.1002/joc.2025)