Advances in Water Resources 99 (2017) 53–59

Contents lists available at ScienceDirect

Advances in Water Resources

journal homepage: www.elsevier.com/locate/advwatres

Confidence intervals for return levels for the peaks-over-threshold approach

Thomas Schendel[a,*], Rossukon Thongwichian[b]

[a] Oranienburger Straße 33, 16775 Gransee, Germany
[b] Thammasat University, Department of Biotechnology, Faculty of Science and Technology, Pathumthani, 12121 Thailand

ARTICLE INFO

Article history:
Received 5 April 2016
Revised 3 August 2016
Accepted 23 November 2016
Available online 24 November 2016

Keywords:
Flood frequency analysis
Confidence interval
Peaks-over-threshold
Test inversion bootstrap
Profile likelihood

ABSTRACT

The peaks-over-threshold (POT) approach is an important alternative to the annual block maxima (ABM) method in flood frequency analysis. POT requires the mathematical description of both, the number of exceedances over the threshold as well as the values of those exceedances. Regardless the method, estimates of extreme flood events are typically associated with a large range of uncertainty, which is usually showcased by appropriate confidence intervals (CIs). However, existing methods to estimate CIs for return levels for the POT approach have mostly neglected its dual-domain character and focused on the distribution of the magnitudes only. We present here a customization of two methods, the Profile Likelihood (PL) and test inversion bootstrap (TIB), which account for the dual-domain structure of POT. Both, PL and TIB, are in the framework of ABM already successfully employed for estimating CIs of extreme flood events. A comparison of the performance of the estimated CIs (in terms of coverage error) of the PL, TIB, and percentile bootstrap is done. As result, it is seen that both the lower and upper boundary of the CIs are strongly underestimated for the percentile bootstrap approach. A similar effect (although in a much less pronounced way) can be observed for PL. The performance of the TIB is usually superior to the percentile bootstrap and PL and yielded reasonable estimates for the CIs for large return periods.

1. Introduction

Flood frequency analysis is of vital importance for water resource design and management. The most common method to estimate extreme floods with defined return period is the annual block maxima (ABM) approach (see e.g. Katz et al., 2002), where for each year the maximal streamflow is taken into account and a distribution to this series of maximal floods is fitted. An alternative to this procedure is the use of the peaks-over-threshold (POT) approach (a good overview is given by Lang et al., 1999). POT considers all peaks over a defined threshold. Therefore, both, the number of exceedances over the threshold as well as their magnitudes need to be mathematically described. In this sense, POT requires a dual-domain modeling.

The main advantage of POT over ABM is that on the one hand POT allows considering more flood events than ABM, while on the other hand POT avoids taking minor flood events into account, which nonetheless are part of the ABM flood series. Another problem that comes with the ABM approach is that even major flood events are disregarded if a more extreme event happened in the same year. In contrast, the POT approach still includes such events (as long as they can be regarded as independent from each other).

Nevertheless, irrespective of the approach chosen, estimates of extreme flood events come typically with a large range of uncertainty. Confidence intervals (CIs) are usually used to illustrate this uncertainty. In the framework of ABM, many methods have been proposed to estimate such CIs (Burn, 2003; Kyselý, 2007; Obeysekera and Salas, 2014; Rust et al., 2011; Schendel and Thongwichian, 2015). Especially successful attempts are the Profile Likelihood (Obeysekera and Salas, 2014) and the test inversion bootstrap (mainly advocated by Carpenter, 1999 and customized to the case of return levels of extreme flood events by Schendel and Thongwichian, 2015).

However, attempts to estimate CIs of return levels in the framework of the POT approach are rare. Moreover, the overwhelming majority of these approaches ignore the dual-domain character of POT and focus only on the distribution of magnitudes over the threshold. For example, Hosking and Wallis (1987) used therefore the delta method. Coles (2001) included in this approach the variability of the number of exceedances over the threshold (this is to our best knowledge the only attempts that accounts for the dual-domain character of POT). However, this method assumes that the

* Corresponding author.
E-mail address: thomas.schendel@web.de (T. Schendel).

http://dx.doi.org/10.1016/j.advwatres.2016.11.011
0309-1708/

appropriate estimator of a specific return level is normally distributed and that the CI is symmetrically placed regarding the estimated value. While this may be valid for asymptotic sample sizes, it was found that in case of the ABM approach and realistic sample sizes, the CIs are not symmetric at all (see e.g. Obeysekera and Salas, 2014). Coles (2001) also presents a customization of Profile Likelihood to POT, but which only accounts for the variability of the distribution of magnitudes over the threshold.

Our aim in this paper is to study the performance of different methods to estimate confidence intervals for return levels for the complete POT approach. We choose the Poisson distribution for the number of exceedances and the generalized Pareto distribution (GPD) for the distribution of the exceedances over the threshold. Therefore, the Profile Likelihood as well as the TIB was customized to fit this dual-domain situation. Both methods have been proven useful in the ABM framework (see e.g. Obeysekera and Salas, 2014 for the Profile Likelihood and Schendel and Thongwichian, 2015 for the TIB). As result, the performance of three approaches (parametric percentile bootstrap, Profile Likelihood, and TIB) regarding the estimation of CIs are compared and their usefulness evaluated.

2. Methods

2.1. Generalized Pareto distribution

The Pickands–Balkema–de-Haan theorem (Balkema and de Haan, 1974; Pickands, 1975) states that for a broad class of functions the values above a sufficient large threshold c follow an approximate generalized Pareto distribution (GPD). The cumulative distribution function $F(x)$ of the GPD is given by:

$$F(x) = 1 - \left(1 + \frac{a}{b}(x - c)\right)^{-\frac{1}{a}}; \ a \neq 0$$
$$F(x) = 1 - e^{-\frac{x-c}{b}}; \ a = 0.$$

For the density function $f(x)$ we yield correspondingly:

$$f(x) = \frac{1}{b}\left(1 + \frac{a}{b}(x - c)\right)^{-\frac{1}{a}-1}; \ a \neq 0$$
$$f(x) = \frac{1}{b}e^{-\frac{x-c}{b}}; \ a = 0.$$

a denotes the shape parameter, b the scale parameter, and c the threshold. The T-year return level x_T can be calculated as follows (T^{obs} denotes the time period of the observation and N the number of exceedances over the threshold):

$$F(x_T) = 1 - \frac{T^{obs}}{N}\frac{1}{T}$$
$$\rightarrow x_T = c + \frac{b}{a}\left[\left(\frac{T^{obs}}{N}\frac{1}{T}\right)^{-a} - 1\right]; \ a \neq 0$$
$$\rightarrow x_T = c - b \cdot \ln\left(\frac{T^{obs}}{N}\frac{1}{T}\right); \ a = 0. \quad (1)$$

One important property of the GPD is self-similarity. Even if the threshold of a GPD is changed, the resulting distribution is still a GPD, but with a different value for the mean number of exceedances. Let us consider an initial threshold c_1 with a GPD with parameters a_1, b_1, and mean number of exceedances N_1. If the threshold is raised to $c_2 = c_1 + \Delta c$, the values of the parameters of the new GPD (a_2, b_2) and mean number of exceedances N_2 change as follows (see e.g. Coles, 2001 and Scarrott and MacDonald, 2012) :

$$a_2 = a_1 = a$$
$$b_2 = b_1 + a\Delta c \quad (2)$$

$$N_2 = N_1\left(1 + \frac{a}{b_1}\Delta c\right)^{-\frac{1}{a}}. \quad (3)$$

In turn, if the mean number of exceedances is changed from N_1 to N_2, the threshold changes by:

$$\Delta c = \frac{b_1}{a} \cdot \left[\left(\frac{N_1}{N_2}\right)^{a} - 1\right]. \quad (4)$$

2.2. Bootstrap – general outline and percentile bootstrap

The overall aim is to sample N values (same sample size as the observed sample) from the GPD (we only use parametric bootstrap approaches throughout this article) with the parameters ($a = a_k, b = 1, c = 0$) (which can be easily extended to all possible b and c) that belongs to the Poisson distribution with mean value $\theta = N$. In principle the sampling needs to be done for both distributions (first from the Poisson distribution in order to sample the number of exceedances over the threshold and then the values itself from the GPD). However, another valid approach (which has numerical advantages later on needed) would be to keep the sample size N constant and vary the threshold. To achieve this, we transform the problem as follows: First, we choose a value M, $M > N$ such that in practice basically all samples drawn from the Poisson distribution with mean value M have more than N values. Second, in order to correspond to the initial problem (mean number of exceedances N with corresponding GPD with parameters ($a = a_k, b = 1, c = 0$)), the parameters b and c are transformed according to Eqs. (4) and (2). This yields the set of parameters:

$$a = a_k$$
$$b = \left(\frac{N}{M}\right)^{a_k}$$
$$c = -\frac{1}{a_k}\left[1 - \left(\frac{N}{M}\right)^{a_k}\right] \quad (5)$$

Finally, we sample first from the Poisson distribution (with mean value M) and the GPD with the parameters given in Eq. (5) above. Then we select only the N largest values. The resulting sample still corresponds to the parent GPD with parameters ($a = a_k, b = 1, c = 0$). We estimated the parameters of each sample using maximum likelihood. While Hosking and Wallis (1987) have pointed out that methods like probability weighted moments or method of moments often lead to smaller bias and RMSE, such approaches can lead to inconsistent results (see Ashkar and Tatsambon, 2007). These inconsistencies occur if values of the observed sample are outside of the range of the estimated distribution. After determination of the parameters, the appropriate return levels are determined for each of the B samples. For the percentile bootstrap approach, the $(1-\alpha) \cdot 100\%$ double sided CI is defined by the $\alpha/2 \cdot (B+1)th$ and $(1-\alpha/2) \cdot (B+1)th$ smallest value of return levels (Carpenter and Bithell, 2000) to account for the respective empirical quantiles (Makkonen, 2008). In this work, $B = 9999$ simulations were used.

2.3. Likelihood function of the peaks-over-threshold-process

The test inversion bootstrap (TIB) as well as the Profile Likelihood requires the likelihood function of the whole POT process, and not only the one for the GPD. The likelihood function for POT problem can be expressed as follows:

$$\mathcal{L} = \mathcal{L}^{Pois} \cdot \mathcal{L}^{GPD},$$

where $\mathcal{L}^{Pois}$ and $\mathcal{L}^{GPD}$ are the likelihood functions of the Poisson distribution and the GPD, respectively. For the log-likelihood function $L = \ln(\mathcal{L})$, we get:

$$L = L^{Pois} + L^{GPD}. \quad (6)$$