SFB 823

Comparison between classical annual maxima and peak over threshold approach concerning robustness

Svenja Fischer, Andreas Schumann

Discussion Paper

Nr. 26/2014

SFB 823

Comparison between classical annual maxima and Peak over Threshold approach concerning robustness

Svenja Fischer Andreas Schumann

Ruhr-Universität Bochum, Fakultät für Bau- und Umweltingenieurwissenschaften, 44780 Bochum

svenja.fischer@rub.de andreas.schumann@rub.de

Abstract

Partial duration series (peak over threshold) form a considerable alternative to the classical annual maximum approach since they enlarge the information spectrum. The classical POT approach is based on a Poisson distribution for the annual number of exceedances although this is can be questionable in some cases. Therefore two different distributions (Binomial and Gumbel-Schelling (Gumbel and Schelling (1950))) are considered. The results show that they do rarely make a difference to the Poisson distribution. In a second step we investigate the robustness in the sense of stability against the occurrence of extreme events of the POT compared to annual maxima and show that in the case of extreme events the POT behaves much more robust and fits very good to the data.

KEYWORDS: Partial duration; annual maxima; robustness; Poisson process

1 Introduction

In the context of the climate change discussion and some extreme floods occurring in the last decade the question of a robust estimation of flood probabilities and the connected flood protection is more and more foregrounded.

Since the recorded series of floods (monthly or annual) are very limited and seldom longer than 100 years the needed information about extreme floods is often not or not well enough represented in the data sample. Therefore the fitting of a suitable distribution and the extrapolation to domains lacking of data is the key question to gain information about events with very low exceedance probabilities. The decision of the underlying statistical model is crucial in this context.

It is known that partial duration series (Peak over Threshold (POT) approach) form a considerable alternative to the classical annual maximum flood series in performing flood frequency analysis. It has been argued that the POT approach uses more data and thus more information about the floods whereas the annual maximum approach only uses one event per year and so maybe ignores further floods in the same year, which are higher than annual maxima in other years. So it is not unlikely that the POT has a more robust behavior.

Robustness can become an important point in the estimation of the underlying statistical model when an extreme event with low exceedance probability which is significantly smaller than $\frac{1}{n}$ occurs in a time series of n years. Also robustness does not only mean robustness against outliers but also against model misspecifications or errors in the data.

In literature (Langbein (1949), Stedinger (1993)) there is a comparison between the statistics of time periods between flood events derived from partial duration series and annual series. The return periods based on annual series are integers of years (annualities), the return periods of POT are time spans, expressed as real numbers of years. They show that for return periods higher than 10 years it does not make a difference which approach is used. Nevertheless we will see that it makes a difference, if we calculate the *annual* return period based on annual maxima or based on the partial series.

Additionally Cunnane (1973), Madsen *et al.* (1997) and Rosbjerg (1985) compare the efficiency of the estimation of annual return periods by the annual maxima and partial duration series respectively for independent and dependent peaks and different estimators. Also Tavares and Da Silva (1983) show the differences in the estimation variance of both cases.

Rasmussen and Rosbjerg (1991) considered seasonal approaches for the partial duration series.

In this article we want to compare the classical statistic of annual maxima with the annual maximum statistic resulting of the POT in respect of its robustness. In particular we want to estimate extreme quantiles (99% and 99.5%) of different gauges within the drainage basin of the river Mulde in Saxony. In this area several extreme events happened in the last 12 years so that we can find large outliers in our data samples. Robustness in this context means stability against the occurrence of extreme events.

As the distribution of the annual maxima based on the POT is specified by a combination of the distribution of the magnitude of exceedances with the discrete distribution for the annual number of exceedances (for details see Section 2.2 POT) we investigate the influence of the choice of the discrete distribution in the POT approach. The often used application of the Poisson distribution is based on the assumption of an underlying Poisson process, which is not the most fitting distribution in this case, and extending the study of Önöz and Bayazit (2001) we replace it by other discrete distributions, especially the one proposed by Gumbel and Schelling (1950).

We first give a detailed overview of the different models and methods and in a second step apply them on our data.

2 Methods

Let us consider a data set $(X_1^{(d)},\ldots,X_n^{(d)})$, that is for example a $n=100$ -year series of $d=12$ monthly maximum discharges in a year. Of course this model can also be used for describing seasonal models, for example for winter floods, by choosing d as the times of seasonality (for winter e.g. $d=6$). It is important to consider that not every maximum specifies a flood event, which has to be defined by a threshold.

2.1 Annual Maxima

Regarding annual maxima one has to deal with the time series $\left(\max_{1\le k\le d}(X_1^{(k)}),\ldots,\max_{1\le k\le d}(X_n^{(k)})\right)$.

It seems naturally to fit the Generalized Extreme Value (GEV) Distribution to annual maxima having in mind that the Fisher-Tippet-Gnedenko Theorem says that for independent, identically distributed random variables $Y_1,\ldots,Y_n$ the maximum $max(Y_1,\ldots,Y_n)$ (properly normed) converges in distribution to a GEV-distribution (under some technical conditions) for $n™ \infty$. In hydrology this limit distribution is widely known and often used because of its flexibility reasoned by its three parameter distribution function:

$$G(x)=exp\left(-\left(1+\xi\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}}\right), \text{for} 1+\xi(x-\mu)/\sigma>0,$$

where $\xi\in\Box$ is the shape parameter, $\sigma<0$ the scale parameter and $\mu\in\Box$ the location parameter. The special case $\xi=0$ (Gumbel-distribution) is given by

$$G(x)=exp\left(-exp\left(-\frac{x-\mu}{\sigma}\right)\right).$$

The estimation of the parameters can be done by the Probability Weighted Moments (PWM) (Hosking, J. R. M. *et al.* (1985)), since they are in comparison to Maximum-Likelihood or Moment estimators a bit less efficient but a bit more robust (Kumar *et al.* (1994)).

2.2 POT

The POT approach in comparison enlarges the information used for the fitting by considering not only the annual maxima but every (monthly) maximum above a threshold specifying a flood. In particular this means that we involve every flood peak above a certain threshold x_0 (in our case this will be $x_0 = \min_{1\le i\le n}\left(\max_{1\le k\le d}(X_i^{(k)})\right)$, the minimum annual maximum of all years). This is the so called partial duration series. The independence between the single events has to be ensured, which can be done by a consideration of the date of each peak.

Again we use a result of the extreme value theory to find a suitable distribution: by the Balkema-de Haan-Pickands Theorem the conditional excess distribution of $Y_1,\ldots,Y_n$ above a certain threshold x_0 converges to a Generalized Pareto Distribution (GPD) with distribution function:

$F(x) = 1-\left(1+\kappa\frac{x-x_0}{\beta}\right)^{-\frac{1}{\kappa}}$, for $x > x_0$ if the shape parameter $\kappa \ge 0$ and $x_0 \le x \le x_0 - \beta/\kappa$ otherwise and the scale parameter $\beta > 0$.

We want to mention that in literature there are many other distributions used for the conditional excess distribution, especially the special case $\kappa = 0$ of the GPD, the shifted Exponential distribution given by $F(x) = 1-exp\left(-\frac{x-x_0}{\beta}\right),\ x > x_0$.

Rosbjerg *et al.* (1992) showed that this distribution is preferable in the case where $\kappa < 0.1$ in the sense that it gives a better approximation to the data. Nevertheless we want to ensure a great flexibility and therefore consider the three parameter case, having in mind that we can reduce it to the special case of the Exponential distribution.

To estimate the parameters of the GPD we have a wide range of possibilities. Often the Probability Weighted Moments (Wang (1991)) are used, since they are rather efficient . Nevertheless it is also worth considering more robust estimators to counter uncertainties in the model selection and also outliers in the data. We will become more specific in this point later on. The threshold parameter x_0 does not have to be estimated and is given by definition. But it is very important to mention that the choice of it plays an important role in the behavior of the estimates (cf. Begueria (2005)).

Still one problem remains since we are interested in annual return periods. Therefore we have to transform the results we get for the distribution of the magnitude of excesses, which is the GPD, to get results for annual maxima. An often used relationship between annual maxima and the partial duration series is the following: using the total probability theorem we get for the distribution function of the annual maxima F_a (cf. Shane and Lynn (1964)):

$$F_a(x) = \sum_{k=0}^{\infty} \mathrm{P}(l=k)\left(F(x)\right)^k, \qquad (1)$$

where $\mathrm{P}(l=k)$ is the probability that the annual number of exceedances of x_0 equals k.

In the following we will compare three discrete probability distributions which could represent P properly.

2.2.1 The Poisson distribution

The most used discrete distribution for describing the occurrence of rare events is the Poisson distribution

$P_P(l=k)=\frac{\lambda^k}{k!}e^{-\lambda}$, based on the assumption that the underlying process is a Poisson process.

It seems therefore naturally to use this distribution also in the case of the annual number of exceedances and actually this done in most cases (e.g. Stedinger (1993), Cunnane (1973)). The parameter λ represents both the mean and the variance of the distribution and is estimated by the mean of the annual number of exceedances $\hat{\lambda}=\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{d}1_{[X_i^{(j)}>x_0]}$.

Nevertheless there are also some disadvantages which question the applicability of the Poisson distribution in the case of partial duration series. One important point is that because of equation (1) there will be probability mass for every $k=0,1,\ldots$, where in reality it is not possible that k could equal more than d. Having the example of $d=12$ monthly maxima every year in mind it is not possible that x_0 is exceeded more than 12 times, that is every month. It was also mentioned in literature before that the assumption of equal mean and variance , which is the base of the Poisson distribution, does not always hold (Taesombut and Yevjevich (1978), Cunnane (1979), Önöz and Bayazit (2001)).

This was the reason for us to consider different distributions and to compare the results.

Combing the Poisson distribution with the GPD in equation (1) we gain for the distribution function of annual maxima:

$$F_{a,P}(x)=\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}e^{-\lambda}\left(1-\left(1+\kappa\frac{x-x_0}{\beta}\right)-\frac{1}{\kappa}\right)^k=e^{-\lambda}exp\left(\lambda\left(1-\left(1+\kappa\frac{x-x_0}{\beta}\right)^{-\frac{1}{\kappa}}\right)\right)$$

$$=exp\left(-\lambda\left(1+\kappa\frac{x-x_0}{\beta}\right)^{-\frac{1}{\kappa}}\right)$$

This is indeed the GEV distribution with parameters $\xi=\kappa$, $\sigma=\beta\lambda^{\kappa}$ and $\mu=x_0-\frac{\beta\left(1-\lambda^{\kappa}\right)}{\kappa}$.

2.2.2 The Binomial distribution

Mathematically the Binomial distribution seems to be another distribution worth considering since it is a typical distribution for describing the number of successes (in our case exceedances) in a sample. Therefore it is not surprising that this distribution was also considered before (c.f. Önöz and Bayazit (2001)). Its probability mass function is given by

$$P_B(l=k)=\binom{r}{k}p^k(1-p)^{r-k},$$

where p is the probability of an exceedance of the threshold and therefore estimated by $\hat{p} = \frac{1}{nd}\sum_{i=1}^{n}\sum_{j=1}^{d} 1_{[X_i^{(d)} > x_0]}$. By choosing $r = d$ we do not have the problem of having probability mass for $k > d$, since then the Binomial distribution equals 0. For $n \to \infty$ and $p \to 0$ with $np \to \lambda$ the Binomial distribution converges against the Poisson distribution.

Although the using of this distribution seems to be great advantage against the Poisson distribution, Önöz and Bayazit (2001) showed that it does not make a difference.

The annual maxima distribution is then

$$F_{a,B}(x) = \sum_{k=0}^{\infty} \binom{d}{k} p^k (1-p)^k \left(1 - \left(1 + \kappa \frac{x - x_0}{\beta}\right) - \frac{1}{\kappa}\right)^k = (1-p)^d \left(1 + \frac{\alpha}{1-\alpha}\left(1 - \left(1 + \kappa \frac{x - x_0}{\beta}\right) - \frac{1}{\kappa}\right)\right)^d$$

2.2.3 The Gumbel-Schelling distribution

As third and last distribution we want to consider is the distribution proposed by Gumbel and Schelling (1950)

$$\mathrm{P}_G(l = k) = \frac{\binom{nd}{m} m \binom{d}{k}}{d(1+n)\binom{d(n+1)-1}{m+k-1}},$$

where m is the rank of x_0 in the sample $(X_1^{(d)}, \dots, X_n^{(d)})$ in decreasing order.

If $n, d \to \infty$ and $\frac{n}{d} \to 1$, Gumbel and Schelling (1950) showed that this distribution converges asymptotically to a normal distribution, otherwise if $n, d \to \infty$ and m and k remain small it converges to the Poisson distribution.

This distribution has the great advantage that we do not have to estimate any parameters and therefore have no uncertainties in this point. The only assumption needed is continuity of the data.

For the distribution of the annual maxima we get

$$F_{a,G}(x) = \sum_{k=0}^{\infty} \frac{\binom{nd}{m} m \binom{d}{k}}{d(1+n)\binom{d(n+1)-1}{m+k-1}} \left(1 - \left(1 + \kappa \frac{x - x_0}{\beta}\right)^{-\frac{1}{\kappa}}\right)^k .$$

3 Data Application

We analyzed data from 17 gauges in Saxony within the Mulde river basin (Figure 1) with a length of at most 75 years and floods recorded up to year 2012. As mentioned above these data series have the special property of having several extreme events where the most extreme occurred in the year 2002. An example is shown in Figure 2 where we can see the gauge Nossen (gauge number 10).

We are aware that there could be some dependence in the time series when considering monthly maximum discharges. Nevertheless we recognized in our calculations that the used estimators are due to their robustness able to cope with slight dependencies, that are dependencies of monthly maxima. However the influence of dependence has to be studied further since the model is conceived for independent data and an underlying process of dependence could lead to errors.

As threshold we chose $x_0 = \min_{1 \le i \le n} \left(\max_{1 \le k \le d} (X_i^{(k)}) \right)$, that is the minimum annual maximum of the whole data sample.

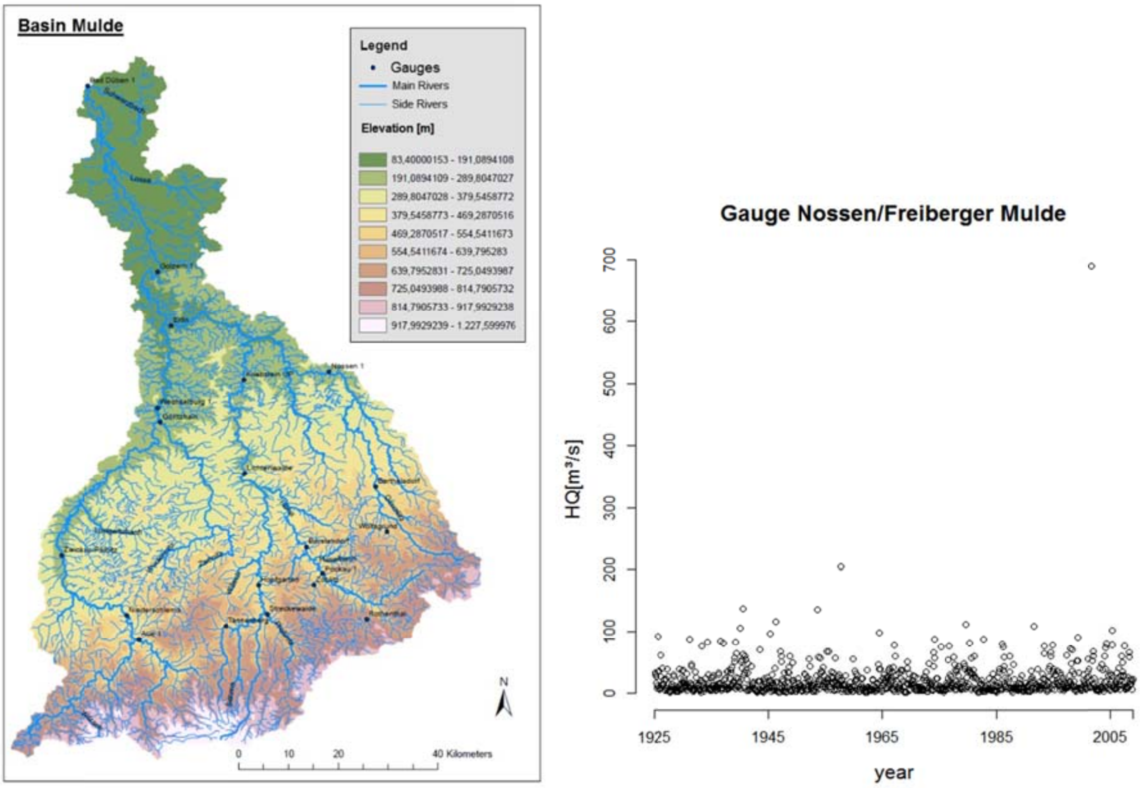

Figure 1: Drainage basin of the river Mulde in Saxony (left) and monthly maximum discharges of the gauge Nossen at the river Freiberger Mulde (right).

At first we want to compare the different POT models. For this we calculated the 99.5% and the 99.9% quantile for all three cases that is the 200 and 1000 years flood respectively. Since there does not exist an explicit representation of the quantile in the Gumbel-Schelling case we used numerical approaches to estimate the quantiles. For a better comparison of the results between gauges we use the ratio of the Poisson- and the Binomial-model quantiles ($R_{P,B} = \frac{x_P(q)}{x_B(q)}$ for $q = 0.995, 0.999$) as well as of the Poisson- and Gumbel-Schelling-model quantiles ($R_{P,G}$). If they are nearly identically this ratio of course has to be close to 1.

We calculated these ratios for all 20 gauges with all three models using the PWM estimation for the GPD distribution.

Although the Poisson distribution has some really undesirable properties, as we stated above, we cannot find a distinct difference in the estimation of the difference models even if we consider very extreme quantiles. Comparing Poisson and Binomial distribution in the POT model they are rather identical. This goes along with the results of Önöz and Bayazit (2001). But even the Gumbel-Schelling distribution does not make a real difference in estimation, though the estimation tends to be smaller.

The ratios are all very close to one and differ only in the fourth decimal place, which is reasoned numerically.

A possible reason could be that the Poisson distribution is the limit distribution of both the Binomial distribution and the Gumbel-Schelling distribution.

Since the threshold has a great influence on the estimation we also tested the threshold $T = 2.5MQ$, that is 2.5 times the average discharge. The tendency of the results remained the same and the results are therefore omitted here.

Since it seems to make no difference, if we use the supposed to be unfitting Poisson distribution or one of the other discrete ones, in the following we will continue to use the Poisson distribution, since the model has a much easier representation than in the other cases, having an underlying GEV distribution.

In a second step we compare the POT approach (from now on model 2.1.2) with the classical annual maximum approach. We fit a GEV distribution to the sample $\left(\max_{1\le k\le d}(X_1^{(k)}),\dots,\max_{1\le k\le d}(X_n^{(k)})\right)$ using the PWM estimation and the POT model to the whole sample $(X_1^{(d)},\dots,X_n^{(d)})$ the same way as before. We also considered a robust method to estimate the parameters of the GPD by using a Minimum Distance (MD) estimator with Cramer-von-Mises-distance (cf. Dietrich, D. and Hüsler, J. (1996) in the GEV case).

That is, the parameters are estimated by solving

$$\left(\hat{\kappa}_{MD}, \hat{\beta}_{MD}\right) = \arg\min_{\kappa,\beta} \int_{-\infty}^{\infty} \left(F_n(x) - F(x;x_0)\right)^2 dx,$$

where F_n is the empirical distribution function of all $X_i^{(j)} > x_0$ and $F(x;x_0)$ is the Generalized Pareto distribution with given threshold x_0. Contrary to the ML estimator the MD estimator uses the minimum distance to gain estimates instead of the maximum probability. That means that the MD estimator declares an estimation even as optimal, if one value does not fit but all the others do perfectly. Using the ML estimator this is not possible, since such an outlier would be declared as non-probable and therefore the estimation as non-fitting. Therefore the MD estimator has robust behavior, which can be also shown by its bounded influence function (Dietrich, D. and Hüsler, J. (1996))

We will refer to this as the robust POT approach.

First of all we wanted to know the reliability of the three different fitting approaches. Therefore we used a criterion proposed by Renard *et al.* (2013). We calculated the reliability index

$$pval_i(l) = \hat{F}_l\left(\max_{1\le k\le d}(X_i^{(k)}(l))\right)$$

where $l = 1,\dots,17$ is the gauge number, $\hat{F}_l$ the fitted distribution function for gauge l (based on the first 50 years of records) and $\max_{1\le k\le d}(X_i^{(k)}(l))$ the annual maximum discharge of gauge l in year i of the records. By application of the inverse distribution function one can easily see that under the hypothesis of an reliable estimation ($\hat{F}_l = F_l \;\forall l$) $(pval_i(l))_{i=1,\dots,n_l}$ is uniformly distributed in the interval *[0,1]* for every gauge l with record length n_l. Therefore a graphical tool to investigate the goodness of fit of our approaches is a QQ-Plot of *pval* for every fitting approach (**Figure 2**). Often also a PP-Plot is used but since in a QQ-Plot especially the extreme domains are illustrated in a better way we choose this presentation.

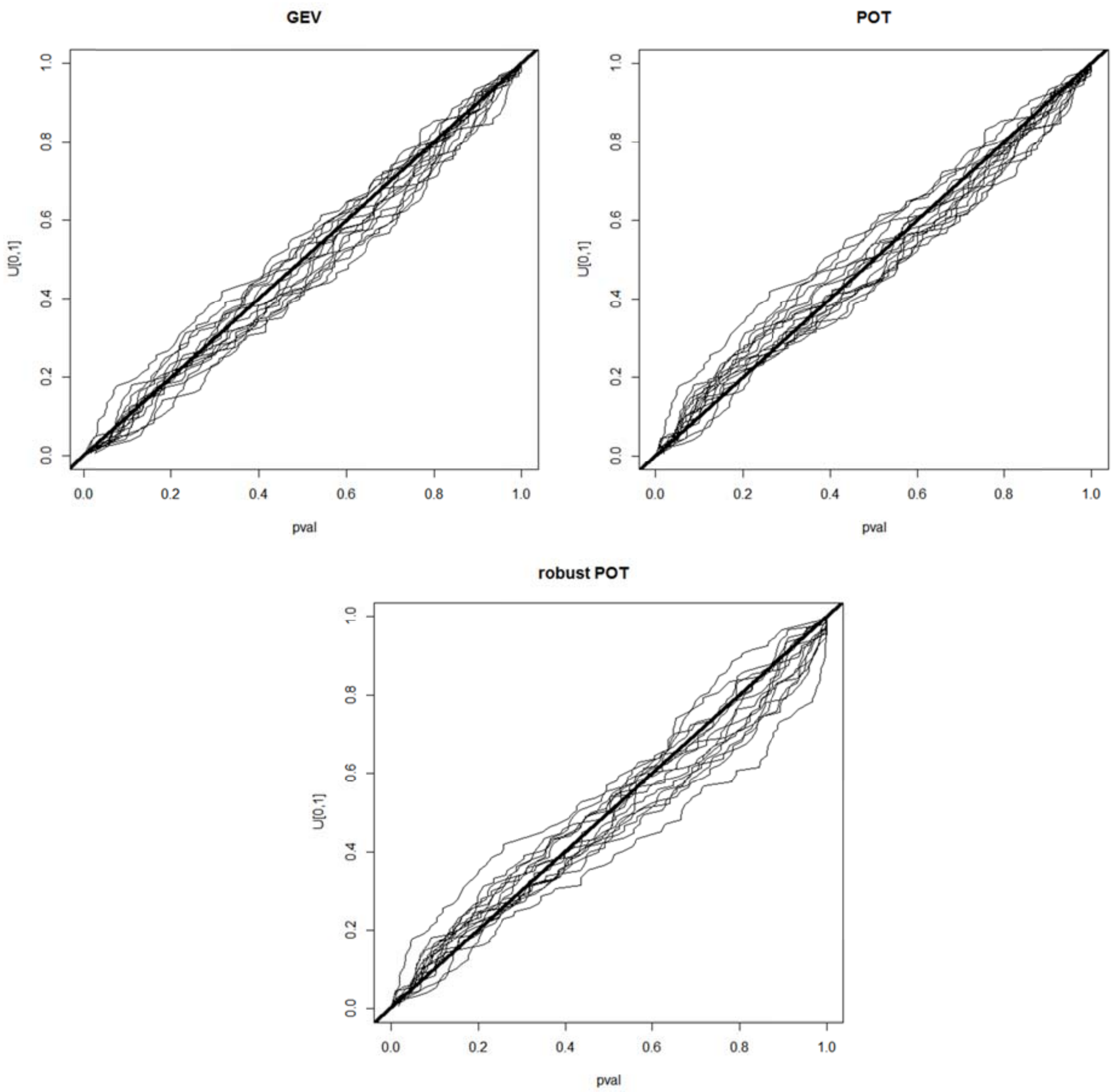

Figure 2: pval QQ-Plot for annual maximum (upper left), POT (upper right) and robust POT (bottom) approach for gauges $l = 1, \dots, 17$

We can see that all three approaches give a good estimation of the distribution of the whole sample. In the upper quantiles the POT seems to give the best goodness of fit and all in all seems not to differ very much from the classical annual maximum statistic. For the robust POT we can see a greater deviation for some gauges. Nevertheless the deviation of all three approaches does not seem so large that we have to reject one of the approaches.

We are aware that it is possible by this usage of the *pval* that for some *i* we plug in a value on which the fitting of $\hat{F}_l$ is based and therefore could possibly have an alleged good fitting. To consider this we also applied cross validation, that is

$$pval_i'(l) = \hat{F}_{l;(i)}\left(\max_{1 \le k \le d}(X_i^{(k)}(l))\right),$$

where $\hat{F}_{l;(i)}$ is the fitted distribution function for gauge l based on all data except of $\max_{1 \le k \le d}(X_i^{(k)}(l))$. The results we got from this were very similar to the ones of the originally *pval* so we came to the conclusion that the abovementioned problem did not influence our results.

To investigate the fitting in the higher quantiles, were a difference in the fitting seems to be biggest, we want to have a close look on the gauge Nossen (see above) by using a QQ-Plot (**Figure 3**).

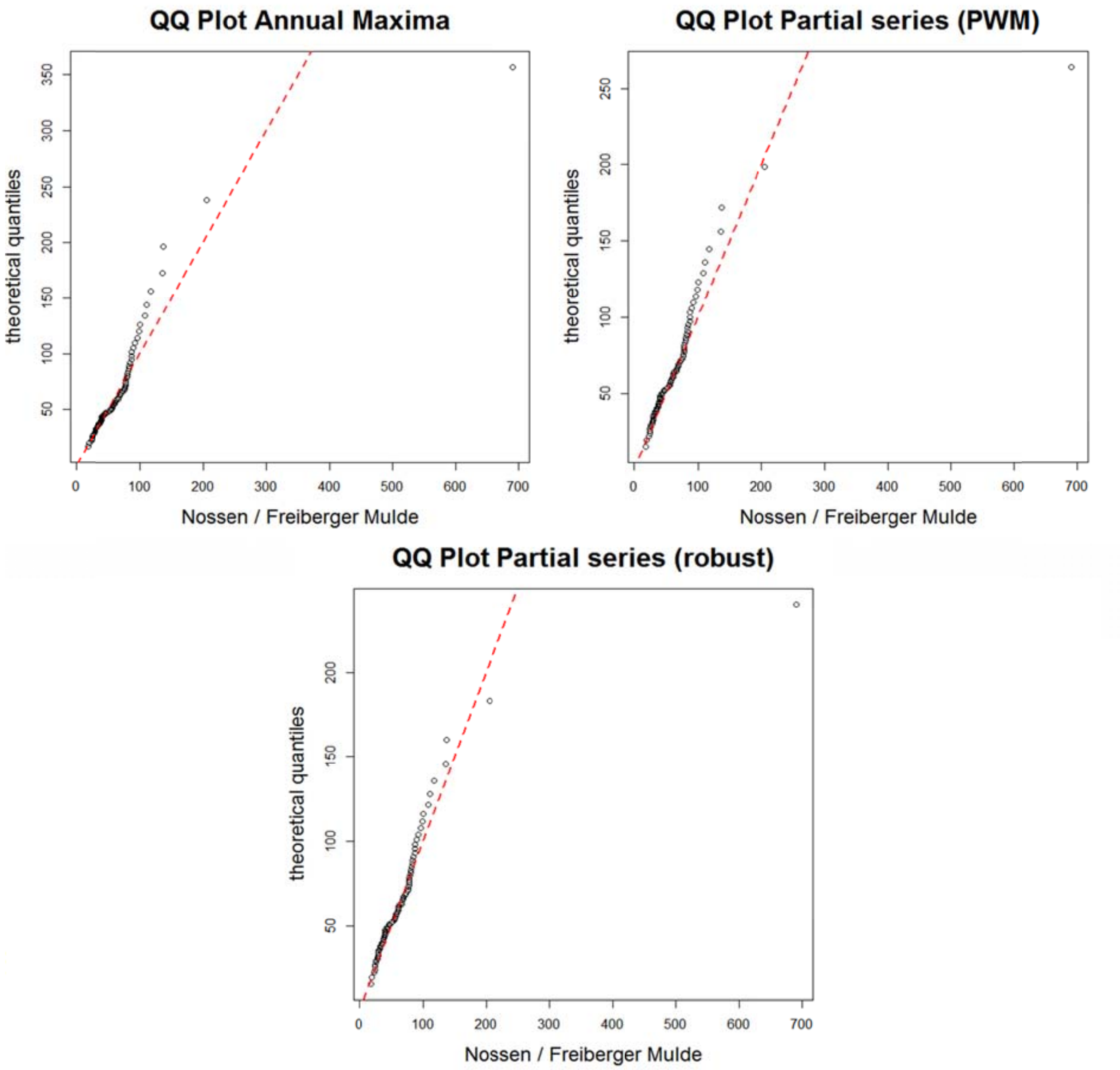

Figure 3: QQ Plots of the gauge Nossen and the estimates using the Annual Maximum (top left), the POT with PWM (top right) and the robust POT (bottom) model respectively.

We can see how much the classical annual maximum model is influenced by the single extreme event leading to a worse fitting to the middle quantiles. The model based on partial series using additional data values, which mostly lie in the lower domain, has a better fitting in the middle quantiles. Thus the greater information spectrum leads even in the estimation of annual maximum discharges to a more robust fitting. If this is robustified by the application of the MD estimator once more we can see a good fitting of all quantiles except of the single extreme event.

Another criterion for a fitting is its stability. We show the stability of the estimation of the annual return period generic for the gauges Nossen and Wechselburg (number 2). As tool we used the absolute deviation of the annual return period of an event estimated on the basis of all recorded events up to this point (we call this predicted return period) and the annual return period of the same event estimated on the basis of all recorded events up to this point including this event (occurring return period). This is done for an increasing sample length, starting with ten years (**Tables 1 and 2**).

We can see that after a short time of stabilization, which is needed for an adequate calibration, the POT and especially the robust POT approach give much more stable estimations of the annual return period. Whereas the classical annual maximum approach is highly influenced by the occurring of extreme events the other two approaches are not so much. Where the difference is ∞ the model was not able to fit a suitable distribution with the given data. We can also see that the flood in the years 2002 in Nossen is so extreme that every model estimates return periods much higher than 10000.

Table 1: absolute deviation of predicted and occurring annual return period for the gauge Wechselburg with the three different fitting approaches. Only differences higher than 5 years in one case are shown.

year	GEV	POT	robust POT
1923	4	35	∞
1924	67	57925	∞
1926	10	69	∞
1932	92	76	127
1944	354	445	131
1975	19	16	6
2002	173	124	85

Table 2: absolute deviation of predicted and occurring annual return period for the gauge Nossen with the three different fitting approaches. Only differences higher than 5 years in one case are shown.

year	GEV	POT	robust POT
1931	∞	30	3
1956	18	6	5
1958	750	65	26
2002	19528768	2712055	10213

The stability was also investigated by using the robustness criteria $SPAN_T$ proposed by Garavaglia *et al.* (2010) and often used to compare the robustness of fitting methods (Kochanek *et al.* (2013), Renard *et al.* (2013)). The *SPAN* for the annual return period *T* at gauge number *l* is calculated by the following formula

$$SPAN_T(l) = \frac{\max\limits_{1\le s\le b}\{\hat{q}_T(s)\} - \min\limits_{1\le s\le b}\{\hat{q}_T(s)\}}{\frac{1}{m}\sum_{s=1}^{b}\hat{q}_T(s)}, \qquad (2)$$

where $\hat{q}_T(s)$ is the estimated quantile related to annual return period T for a subperiod s of b non-overlapping subperiods. It is a kind of rank function, where the optimal value of $SPAN_T$ indicating a robust behavior of the model is 0.

Since in our case the sample length is very limited and the estimators need a certain quantity of data we reduce the quantity of subperiods to two, having one with length 50. Equation (2) therefore reduces to

$$SPAN_T(l) = \frac{\left|\hat{q}_T(s_1) - \hat{q}_T(s_2)\right|}{\frac{1}{2}\left(\hat{q}_T(s_1) + \hat{q}_T(s_2)\right)}.$$

To give a graphical evaluation, $SPAN_T$ is calculated for all gauges and annual return periods of 100, 200 and 1000 years and finally the empirical distribution of these values is plotted (**Figure 4**).

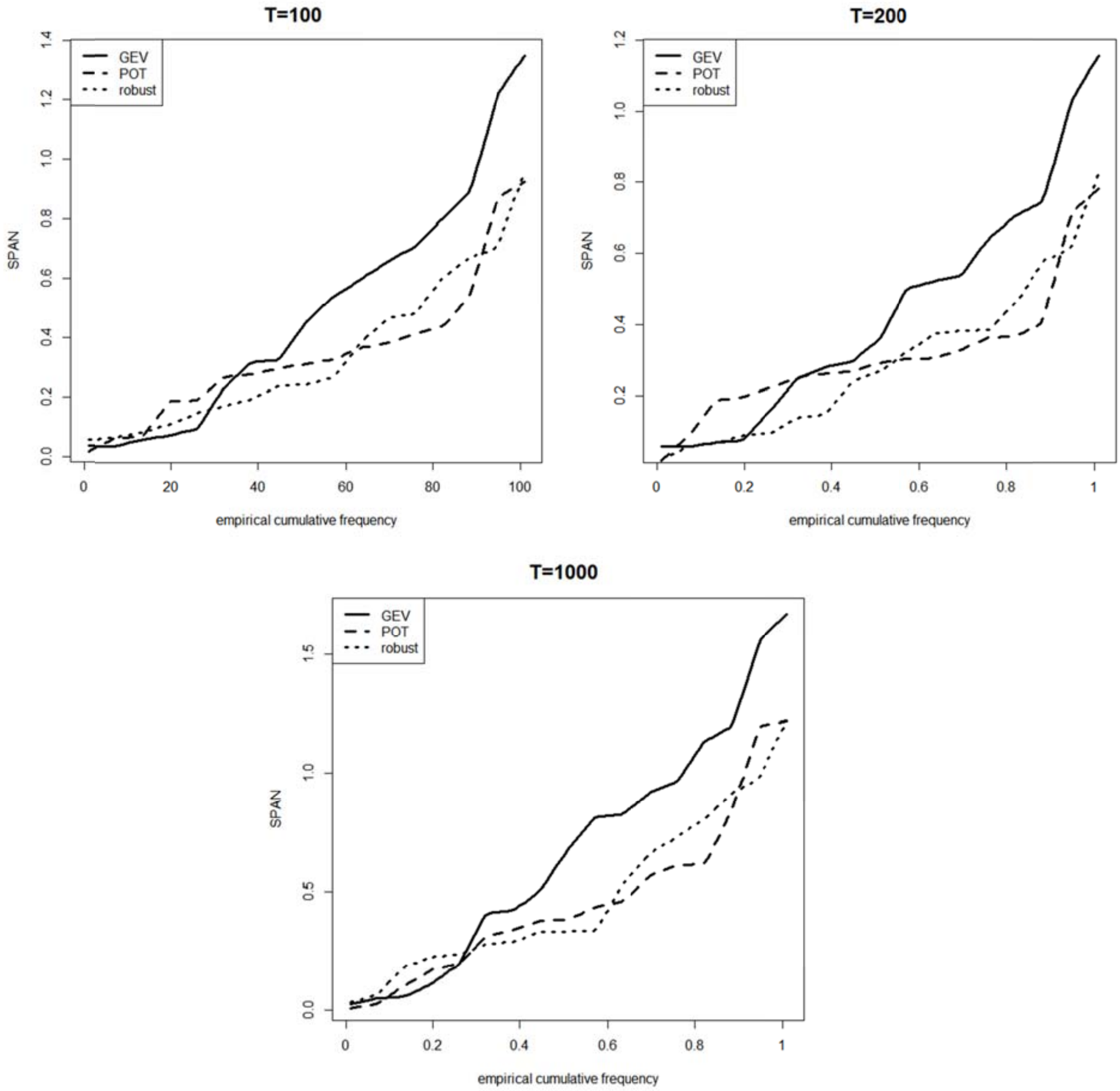

Figure 4: Empirical distribution of $SPAN_T$ for annual return periods of T=100,200 and 1000 years and all gauges.

The GEV approach is most sensitive to changes in the data having clearly higher *SPAN* scores for every three different annual return periods. In general the maximum value of *SPAN* increases with increasing *T*, indicating that the estimates become more sensitive for higher quantiles. This is reasoned by the small quantity of data in these high domains. POT and robust POT behave similar in the high quantiles, whereas the robust POT approach seems to have advantages in the lower quantiles. That is the main part of the calculated *SPAN* were small and therefore the robust POT approach can be declared as the most robust approach of these three.

The behavior observed in the investigations above leads us to the assumption that the POT approach with its robust property gives us an overtime stable estimation of high quantiles which is less influenced by single high floods. To underline this we consider the 99%-quantile of a gauge where we use an increasing sample length. That is we start with the first 10 years of recorded monthly maximum discharges (respectively annual maximum discharges for the classical approach) and estimated the 99%-quantile of this sample. Then the sample length is increased step by step by one year and the 99%-quantile is estimated again. This is done up to the point where all recorded floods are included, in our case the year 2012. Plotting these results the typical shape would be a “sawtooth” curve, since every large event occurring in the considered series causes a jump in the quantile estimation in the moment of its occurrence which then decreases slowly until the next large event occurs. All in all we normally have a slight increase in the quantile-value with growing sample length.

An example of such a moving quantile estimation is shown in **Figure 5** for the gauge Wechselburg and the gauge Nossen. For the classical annual maximum approach (GEV) we see the typical “sawtooth” curve as mentioned above. Every large event occurring in the sample leads to an abrupt increase. This behavior is strongly impaired when the POT approach is used. Although there are still jumps in the estimation their height is much smaller and all in all we have a smaller variability of the quantiles and a smoother increase. When a robust parameter estimation is used additionally, we can see that the estimation is no longer affected by any jumps. Nevertheless the variability is very high caused by the low efficiency of the MD estimator, a typical property of robust estimators.

One can also see that all three approaches need a sample length of nearly 30 years to deliver homogenous results, analogous to the results in **Fehler! Verweisquelle konnte nicht gefunden werden.**.

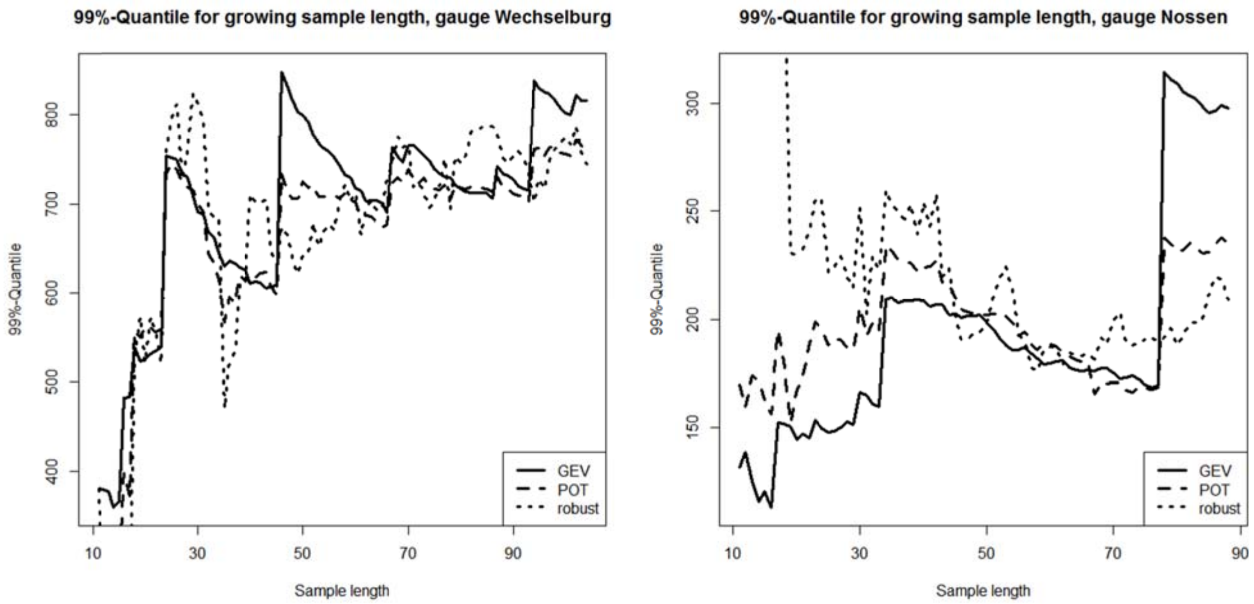

Figure 5: Estimation of the 99%-Quantile of the gauges Wechselburg (left) and Nossen (right) for growing sample length with three different approaches.

Therefore we investigated all 17 gauges and calculated the coefficient of variation of the series of quantile estimations for growing sample length starting with 30 years (**Figure 6**). As expected we can see that the coefficient of variation for the POT approach is mostly smaller than that for the classical approach. The coefficient of variation for the POT with robust estimators is often very high what comes along with what we mentioned before. When the minimum length of the given series is increased the variation for the robust POT approach decreases averagely faster than that of the other approaches.

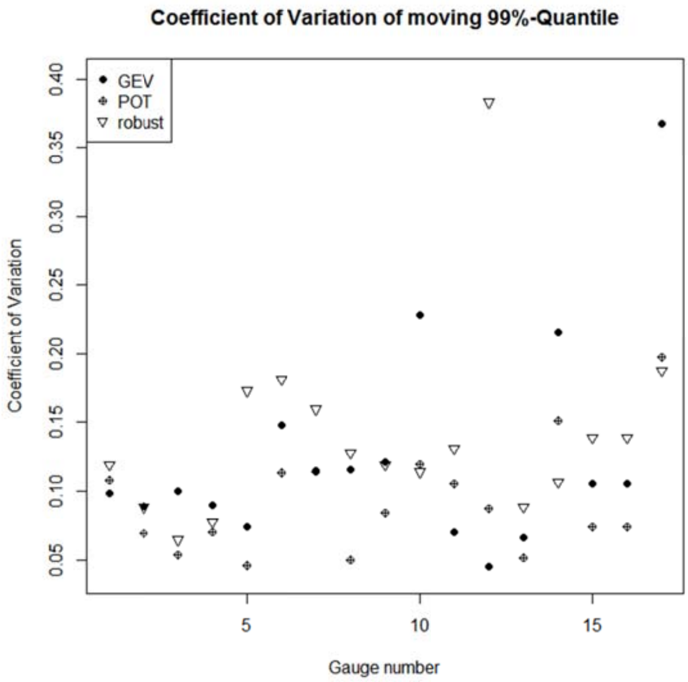

Figure 6: Coefficient of Variation of the 99%-quantile estimation for increasing sample length with three different approaches for 17 gauges of the Mulde basin.

These observations lead of course to similar differences in the estimation of the 99.5%-quantile of the whole sample. For a better comparison we again used the ratio of the estimated 200-year flood for classical annual maxima and POT or robust POT respectively and additionally calculated the ratio also for the 5-year flood to show differences in the lower quantiles. The results can be found in **Figure 7.**

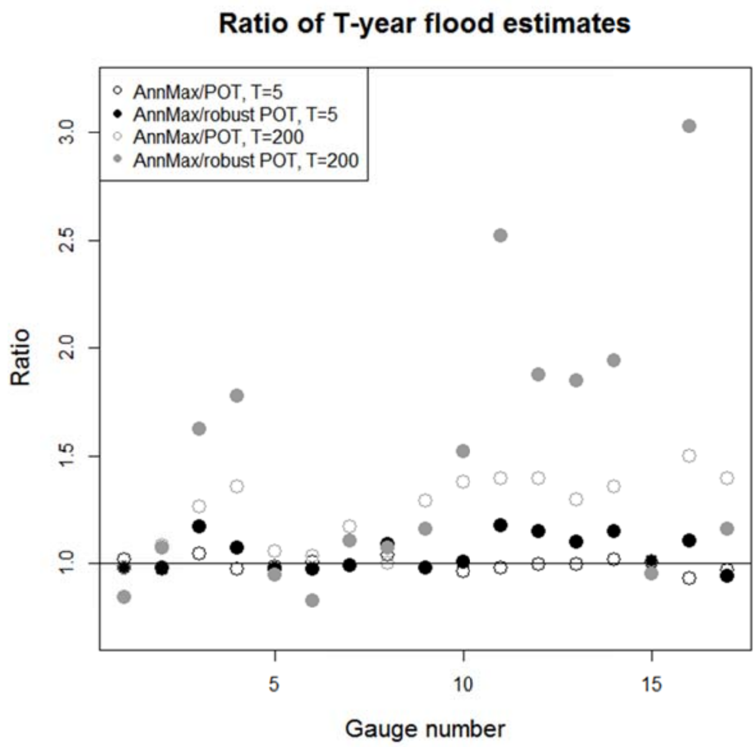

Figure 7: Ratio of the T-year flood estimation of the annual Maximum Approach and the POT (white dots) and the robust POT (filled dots) respectively.

We can see that the classical annual maximum estimation for a return period of T=200 years is mostly higher than the POT estimation leading to ratios greater than 1. The mean ratio is about 1.3 and so we have an averagely 30% higher estimation in the annual maximum model. Using robust estimation in the POT model we see even greater differences. Again the most estimations are smaller with the POT model than with the classical annual maximum model, up to 1/3. The single extreme event therefore influences the classical annual maximum estimation very much, whereas the POT approach gives an estimation where the influence of the extreme event is downweighted. This regards the possibility that the extreme event is an event occurring rarer than in 100 years, which cannot be seen in a data series of only 100 years. We can also see that the lower return periods, e.g. T=5, do not differ so much, regardless of the chosen estimation approach. So the more robust approaches only change the higher quantiles by downweighting the influence of extreme events. This supports the results concerning the reliability of the POT approach.

4 Conclusions

The distribution of flood quantiles by partial duration series is a combination of two distributions, one for the annual number of exceedances and the other for the magnitude of exceedances. Although the Poisson distribution does not seem to fit in a large number of cases, replacement by other, better fitting distributions does not make a difference in the estimation of annualities and can therefore be omitted.

In comparison with the classical annual maximum approach the POT approach forms a robust alternative, which has advantages especially in the occurrence of extreme events. It delivers a stable estimation of high quantiles over large time periods and avoids sudden jumps in estimation. Therefore it gives security in the estimation. The robust POT approach should only be used for samples with a length of at least 50 years, since in ***Figure 5*** we see that at least this sample length is needed for a stable estimation. Otherwise the efficiency of the estimator is much too low. It is therefore unsuitable for the most hydrological series. In the presence of extreme events the estimation based on the POT (robust and non-robust) is mostly lower than that based on the classical approach.

The dependence structure of the times series could play a crucial role in the estimation and a suitable model in the dependence case has to be studied further.

Acknowledgement

The financial support of the Deutsche Forschungsgemeinschaft (SFB 823, ”Statistical modelling of nonlinear dynamic processes”) is gratefully acknowledged.

References

Begueria, S. (2005) Uncertainties in partial duration series modelling of extremes related to the choice of the threshold value.

Cunnane, C. (1973) A particular comparison of annual maxima and partial duration series methods of flood frequency prediction. *Journal of Hydrology*, **18**(3–4), 257–271.

Cunnane, C. (1979) A note on the Poisson assumption in partial duration series models. *Water Resour. Res*, **15**(2), 489–494.

Dietrich, D. and Hüsler, J. (1996) Minimum distance estimators in extreme value distributions. *Communications in Statistics - Theory and Methods*, **25**(4), 695–703.

Garavaglia, F., Lang, M., Paquet, E., Gailhard, J., Garçon, R., and Renard, B. (2010) Reliability and robustness of rainfall compound distribution model based on weather pattern sub-sampling. *Hydrol. Earth Syst. Sci. Discuss.*, **7**(5), 6757–6792.

Gumbel, E. J. and Schelling, H. von (1950) The Distribution of the Number of Exceedances. *Ann. Math. Statist.*, 247–262.

Hosking, J. R. M., Wallis, J. R., and Wood, E. F. (1985) Estimation of the Generalized Extreme-Value Distribution by the Method of Probability-Weighted Moments. *Technometrics*, **27**(3), 251–261.

Kochanek, K., Renard, B., Arnaud, P., Aubert, Y., Lang, M., Cipriani, T., and Sauquet, E. (2013) A data-based comparison of flood frequency analysis methods used in France. *Nat. Hazards Earth Syst. Sci. Discuss.*, **1**(5), 4445–4479.

Kumar, P., Guttarp, P., and Foufoula-Georgiou, E. (1994) A probability-weighted moment test to assess simple scaling. *Stochastic Hydrology and Hydraulics*, **8**(3), 173-183.

Langbein, W. B. (1949) Annual floods and the partial-duration flood series. *Transactions, American Geophysical Union*, **30**(6), 879.

Madsen, H., Rasmussen, P. F., and Rosbjerg, D. (1997) Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events: 1. At-site modeling. *Water Resources Research*, **33**(4), 747–757.

Önöz, B. and Bayazit, M. (2001) Effect of the occurrence process of the peaks over threshold on the flood estimates. *Journal of Hydrology*, **244**(1–2), 86–96.

Rasmussen, P. F. and Rosbjerg, D. (1991) Prediction Uncertainty in Seasonal Partial Duration Series. *Water Resources Research*, **27**(11), 2875–2883.

Renard, B., Kochanek, K., Lang, M., Garavaglia, F., Paquet, E., Neppel, L., Najib, K., Carreau, J., Arnaud, P., Aubert, Y., Borchi, F., Soubeyroux, J.-M., Jourdain, S., Veysseire, J.-M., Sauquet, E., Cipriani, T., and Auffray, A. (2013a) Data-based comparison of frequency analysis methods: A general framework. *Water Resources Research*, **49**(2), 825–843.

Renard, B., Kochanek, K., Lang, M., Garavaglia, F., Paquet, E., Neppel, L., Najib, K., Carreau, J., Arnaud, P., Aubert, Y., Borchi, F., Soubeyroux, J.-M., Jourdain, S., Veysseire, J.-M., Sauquet, E., Cipriani, T., and Auffray, A. (2013b) Data-based comparison of frequency analysis methods: A general framework. *Water Resources Research*, **49**(2), 825–843.

Rosbjerg, D. (1985) Estimation in partial duration series with independent and dependent peak values. *Journal of Hydrology*, **76**(1–2), 183–195.

Rosbjerg, D., Madsen, H., and Rasmussen, P. F. (1992) Prediction in partial duration series with generalized pareto-distributed exceedances. *Water Resources Research*, **28**(11), 3001–3010.

Shane, R. M. and Lynn, W. R. (1964) Mathematical model for flood risk evaluation.

Stedinger, J. R. (1993) Frequency analysis of extreme events. *in Handbook of Hydrology*.

Taesombut, V. and Yevjevich, V. (1978) *Use of partial series for estimating the distribution of maximum annual flood peak*, Ft. Collins.

Tavares, L. V. and Da Silva, J. Evaristo (1983) Partial duration series method revisited. *Journal of Hydrology*, **64**(1–4), 1–14.

Wang, Q. J. (1991) The POT model described by the generalized Pareto distribution with Poisson arrival rate. *Journal of Hydrology*, **129**(1–4), 263–280.