Relation between design floods based on daily maxima and daily means: use of the Peak Over Threshold approach in the Upper Nysa Kłodzka Basin (SW Poland)

Agnieszka Rutkowska a, Patrick Willems b and Tomasz Niedzielski c

aDepartment of Applied Mathematics, Faculty of Environmental Engineering and Land Surveying, University of Agriculture in Kraków, Kraków, Poland; bHydraulics Section, Department of Civil Engineering, KU Leuven, Leuven, Belgium; cDepartment of Geoinformatics and Cartography, Institute of Geography and Regional Development, University of Wrocław, Wrocław, Poland

ABSTRACT
The estimation of flood quantiles is crucial in the assessment of the magnitude and frequency of floods. We carried out a comparative analysis of design discharges estimated from both daily maximum flows and daily mean flows for four mountainous catchments located in the Upper Nysa Kłodzka river basin (SW Poland). After separation of baseflow, split of the riverflow time series in independent events, and selection of the Peak Over Threshold sample, the parameters of the Generalized Pareto Distribution were estimated using the Hill statistic, after bias correction, and considering asymptotic properties. The comparison was performed for various return periods, where the long return periods were of main concern. The jack-knife approach was used to assess the uncertainty of the predicted flood quantiles, and comparison was made with an alternative approach based on annual maxima. We found a meaningful level of differences between daily maximum and mean design discharges and between the rate of change of flood magnitude for which the level (i) stabilized with increasing return period, (ii) decreased downstream, and (iii) was large for catchments susceptible to flooding and with high elevation change. Results are useful in practice when daily maximum discharge is not routinely recorded.

1. Introduction

The main objective of Flood Frequency Analysis (FFA) is to estimate flood quantile magnitudes for given return periods $T$. Such estimates are applied in flood hazard and risk assessment, and in designing water protection devices, including levees, dams, barrages, canals, floodgates and polders or other structures such as bridges and basements for viaducts.

In support of these applications, FFA delivers Flood Frequency Curve (FFC) that depicts relationships between flood magnitude and $T$. In an FFC, the response behaviour of a basin to rainfall is represented because its shape reflects the dynamics of floods. Different methods, however, exist to construct such FFC and analyse its shape (Kellagher 2013).

In the Peak Over Threshold (POT) method a discharge $X$ over a certain threshold $x_i$ ($X > x_i$) is assumed to follow the Generalized Pareto Distribution (GPD). The approach is based on the Pickands–Balkema–de Haan theorem (Balkema & de Haan 1974, Pickands III 2000) stating that the
GPD is the limiting distribution function of $X$ if $x_i \to \infty$. The cumulative distribution function (CDF) of the GPD is

$$G(x) = \begin{cases} 
1 - \left(1 + \frac{x - x_\tau}{\sigma}\right)^{-\frac{1}{\gamma}} & \text{if } \gamma \neq 0, \\
1 - \exp\left(-\frac{x - x_\tau}{\sigma}\right) & \text{if } \gamma = 0,
\end{cases}$$

where $\gamma \in \mathbb{R}$, $\sigma > 0$ are parameters of shape and scale, respectively. The tail $1 - G$ is heavy if $\gamma > 0$, normal if $\gamma = 0$ (exponential CDF) and light (truncated) if $\gamma < 0$. The shape $\gamma$ is known as the extreme value index. The scale parameter depends linearly on the threshold $x_\tau$, with proportionality factor equal to the shape parameter for sufficiently high $x_\tau$, $\sigma = \gamma x_\tau$. High quantiles of GPD are especially responsive to $\gamma$ and very sensitive to deviations of $\gamma$ from 0 (Davidson & Smith 1990, Brilhante 2004). Therefore, the issue of how to distinguish $\gamma = 0$ from $\gamma > 0$ is of great importance in FFA and was addressed in the context of rainfalls (Gomes & Van Montfort 1987, Groisman et al. 2005, Kozubowski et al. 2008) and discharges (Van Montfort & Witter 1985, Smith 1989).

The main advantage of the POT approach in comparison to the use of block maxima (e.g. annual maxima AM) is an increase of the sample size by considering all independent peak flows in the time series instead of only the AM. Next, the GPD can cover a wide range of tail weights. The disadvantage is the inclusion of a subjective judgment into the POT selection and GPD calibration process, and the strong sensitivity of high quantile estimates to the threshold choice. The main difficulties in selecting the POT series are to ensure independence between successive floods and to select the threshold that produces the lowest bias of quantile estimates.

The problem of the choice of independent events and the truncation level are related to each other as a threshold should be high enough to ensure independence. However, it should not be too high to reach much profit from a large sample size. A two-step approach consisting of the identification of the sample of independent events (physical declustering) and setting a threshold high enough to yield convergence to GPD (statistical optimization) usually leads to a reliable estimation of the parameters of the GPD (Bernardara et al. 2014). To provide independent events, Cunnane (1979) proposed to select consecutive peaks the time between which is at least three times longer than the average time to peak and with the smallest peak lower than two-thirds of the highest peak, the Water Resources Council (USWRC 1982) advised that the interexceedance time is not shorter than $5 + \ln (A)$ (where $A$ is the catchment area in square miles) and the drop of the interevent discharge below 75% of the lowest peak (Madsen et al. 1997, Lang et al. 1999, Bačová-Mítková & Onderka 2010, Bezák et al. 2014). Yet another method was introduced by Willems (2009) where the extraction of approximately independent events was based on split of the series into quick and slow flow hydrograph periods. Since then, this method has been increasingly adopted both for river discharges and rainfall intensities.

In relation to the problem of the threshold selection, Langbein (1949) proposed the smallest AM discharge as the initial threshold which is then increased to meet independence criteria, and Cunnane (1979) included a test of the Poissonian number of exceedances based on dispersion index. The issue of the mean number of exceedances has been widely discussed (Dalrymple 1960, Madsen & Rosbjerg 1997b, Lang et al. 1999, Robson & Reed 1999). Those procedures may suffer from some subjectivity and from lack of unique guidelines (Serinaldi & Kilsby 2014). In the method used by Willems et al. (2007), the condition of optimal goodness-of-fit between the estimated and observed quantiles was introduced, as an objective judgment, on the threshold selection.

Calibration of the GPD was conducted using various methods. Madsen and Rosbjerg (1997a) compared the precision of the 100-year quantile assessment for the method of moments, probability weighted moments and maximum likelihood (ML). Martins and Stedinger (2001) proposed the Generalized Maximum Likelihood (GML) method which combines ML and Bayesian approaches. Another approach to estimating of the GPD parameters was adopted by Willems et al. (2007) who
used the Hill estimator of $\gamma$ for POT models of river discharges. This method has been successfully applied for various world regions both to river discharges and rainfall intensities (Boniphace & Willems 2011, Onyutha & Willems 2015, Taye & Willems 2011). The distribution of excess values over a threshold converges to the GPD if the threshold increases to infinity. Thus, the calibration of the GPD may lead to a bias of the estimator of $\gamma$ which also leads to a bias of quantile estimates. Therefore, a bias correction method should be included to reveal a more reliable shape parameter estimate and lessen the bias of high flood quantile estimates. In this paper, a bias correction method based on slowly varying function was used (Beirlant et al. 1999). This method was implemented by Willems et al. (2007) to river discharges.

The POT models were applied to various regions of Poland by Zielińska (1965), Strupczewski (1967), Banasik and Byczkowski (2007), Byczkowski et al. (2008), Rutkowska and Banasik (2014). In this paper, the method of Willems et al. (2007), Willems (2009) was used based on the generalization of the Chapman filter for extraction of the slow and quick runoff fractions and the split of the time series in independent quick runoff event based on independency criteria. For calibration of the GPD, the Hill estimates and the bias correction were used.

The main objective of this paper is to find the relationship between $X_b$, the means-based design discharge, and $X_a$, the maxima-based design discharge based on discharge data from four river gauging stations located within the upper part of the Nysa Kłodzka river basin. The differences between shape parameter estimates, design discharges per unit area and rate of change of FFC per unit area are studied. Additionally, catchment characteristics that have the strongest impact on the difference between $X_a$ and $X_b$ are also identified.

1.1. Rationale for the study

Not uncommonly, even if daily discharge maxima are needed for certain applications (e.g. dam management, hydrologic modelling), they are/were not measured routinely at hydrologic gauges. For instance, historical data, acquired before the era of automatic hydrologic stations, may be available as mean daily discharge calculated from very few readings a day, none of which captures a daily maximum. Although recent automatic hydrometric gauges allow us to measure water levels or discharges with unprecedented sampling intervals of minutes, there exist hydrologic observational networks, or their parts, in which measurements are carried out only a few times a day. This may occur either at sites where there is no automatic gauge or at gauges in which sensors are temporarily out of order. In such cases, an experienced observer is responsible for carrying out the measurements in question, and this is routinely done a few times per day (e.g. 6:00 UTC, 6:00 + 18:00 UTC, 6:00 + 12:00 + 18:00 UTC). This kind of sampling may lead to omissions of observations of daily peak flows. When we know the relationship between $X_a$ and $X_b$ from previous studies and when we accept that calculating a mean value from these small data-sets produces a meaningful estimate of the true mean, we are able to apply the method elaborated in this paper to reconstruct the peak flow. Therefore, knowledge gained through the reconstruction may be important for various types of water engineering applications, such as for the design of hydraulic structures (peak discharges remain the key information for engineers) or for hydrologic forecasting (when mean is forecasted and the prognosis of the peak flow may be established using the approach outlined in this article). In order to confirm that the described application of our method is feasible, we refer to the official hydrometeorological network of Poland which is built and maintained by the Institute of Meteorology and Water Management National Research Institute (Instytut Meteorologii i Gospodarki Wodnej, IMGW–PIB). Even though this network is advanced, with over one thousand fully automatic weather or hydrologic gauges (Dubicki 2007), there are individual sites in which an observer conducts water level or flow measurements. This is currently the case in the town of Zgorzelec in southwestern Poland where water levels are observed every 12 hours.

Thus, having such mean daily discharge data, it is possible to process them in order to estimate the daily maxima time series. The problem is that these two types of data over certain thresholds
represent two different random variables. The first variable, the maxima-based discharge \((X_a)\) is stochastically larger than the second variable, the means-based discharge \((X_b)\). Thereby, for a certain return period \(T\), the design discharge estimated from \(X_b\) is lower than the design discharge estimated from \(X_a\), and the use of \(X_b\) instead of \(X_a\) can result in an incorrect assessment of flood risk. The difference between \(X_a\) and \(X_b\) is substantial for rivers with high flood dynamics. In this case, if only mean daily discharges \(X_b\) are available, these may be transformed to maxima-based discharges \(X_a\). A key issue then is the recognition of the proper transformation between \(X_b\) and \(X_a\). The transformation is catchment-specific.

In this paper, discharges of rivers in the Nysa Kłodzka basin are studied. The basin is a part of the Odra basin, the second longest river in Poland. The necessity of proper management of flood risk is undeniable, as the region is very susceptible to flooding. The region has suffered from severe floods for centuries (Dubicki et al. 2005, Kasprzak 2010, Czaja et al. 2014). The Odra basin was investigated by many authors. Among others, Sen and Niedzielski (2010), Niedzielski (2007, 2010, 2011) as well as Niedzielski and Miziński (2016) examined statistical characteristics of riverflow variability in the basin and attempted to forecast the hydrographs. It was found that discharges at gauges located in the upper and middle Odra basin revealed heavy-tailed dynamics and may be controlled by many processes among which El Niño/Southern Oscillation has been identified. However, the predictability of the hydrographs with data-based models is acceptable for short lead times and is rather limited to a rising limb of hydrograph, with daily peak discharge being usually mismodelled. Hence, there is a need for seeking methods that may help resolve a daily maximum discharge in the process of forecasting high flow episodes when input data are daily means rather than daily maxima.

1.2. Studies on the relationship between instantaneous peak flows (IPF) and mean daily flows (MDF)

Studies on this relationship have been carried out by Langbein (1944), Sangal (1983), and Fill and Steiner (2003) who estimated IPF using a linear combination of MDF from days prior and next to the flow peak. Next, the ratio of MDF to IPF as a function of catchment area was studied by Fuller (1914), Tucci (1991), and Ellis and Gray (1966). Regionalization of this ratio was also proposed enabling the introduction of a series of flows that are shorter than those required by the local approach and that could be used to estimate design floods in places without instantaneous peak flows measurements. Canuti and Moisello (1982) included basin geomorphic parameters into this relation for groups of rivers in Tuscany (Italy). Taguas et al. (2008) found, using the Principal Component Analysis technique, several regional linear equations for rivers in the South Mediterranean Basin of Spain and used such characteristics as rainfall, catchment area and slope. Dastorani et al. (2013) and Shabani and Shabani (2012) estimated IPF from MDF for rivers in Iran using Artificial Neural Networks. Ding et al. (2016) compared various methods to estimate the distribution function of IPF using HBV hydrologic model while IPF quantiles were estimated from MDF quantiles using multiple regression scheme. A common formula between IPF and MDF for groups of catchments was found by Ellis and Gray (1966) (Canada), Taguas et al. (2008) (Spain), Cotecchia (1965) and Tonini (1969) (Italy). In the studies mentioned above, MDF is the mean daily flow at the day of peak. Also note that in this paper no such assumption was introduced because of our focus on design flows, i.e. flows for given return period. The days of maximum flows \(X_a\) and the days of mean flows \(X_b\) selected for POT samples may differ. This problem is discussed in Sections 3.5 and 4.5.

2. Study area and data

The research focuses on the rivers draining the Kłodzko Valley (SW Poland), a mid-mountain abasement situated in the Sudetes Mountains. The specific geographical setting of the Kłodzko Valley is associated with its location between several mountain chains, with considerable relative elevations observed in respect to the bottom of the valley. The main river of the Kłodzko Valley is Nysa
Kłodzka which, after passing through the trough in Bardo at the approximate height of 260 m a.s.l., flows into the Odra River (the second largest river in Poland). There are several tributaries of the Nysa Kłodzka river, most of which are mountainous rivers that are flood-prone (Kasprzak 2010) and reveal short flood wave propagation periods. Extreme peak flows were recorded along most of them, but the Nysa Kłodzka river and its right tributary Biała Łądecka are particularly endangered (Latocha & Parzóch 2010). Their upper parts drain the entire Śnieżnik Massif, the highest mountain chain in the study area. This Massif experiences rainfall episodes which are usually more intensive than in other mountain chains in the Kłodzko Valley. For these reasons, we selected these two rivers for this study, with three gauging stations along the Nysa Kłodzka river (at Bystrzyca Kłodzka, Kłodzko, Bardo) and one along the Biała Łądecka river (at Żelazno). The studied catchments were shown in Figure 1.

Although the two rivers in their upper parts drain the entire Śnieżnik Massif, they have several differences. The most profound difference is a dissimilar land cover of the basins, with various percentages of forested areas. Three sub-basins upstream of Bystrzyca Kłodzka, Kłodzko and Bardo reveal similar values of the Terrain Ruggedness Index (TRI), while the remaining sub-basin upstream of Żelazno is more rugged than the former three basins. The TRI is a measure of terrain elevation heterogeneity. It is the root mean square elevation change between any point on a grid and its surrounding area (Riley et al. 1999). The catchment characteristics are juxtaposed in Table 1.

Figure 1. Studied catchments.
The difference in river defences is also noticeable. The number of settlements along the Biała Lądecka river banks is observed to be much higher than along the Nysa Kłodzka, thus the river course of the former has been much more protected than of the latter. Along the Biała Lądecka river, many buildings, bank reinforcements and groynes are tied together side by side to form a continuous system of protection measures; mostly along the middle and lower river reaches (Latocha and Witek 2013, Witek 2012). This causes an increase in the steepness of the river banks, and of the river flow response time to rainfall. Along the Nysa Kłodzka river, in turn, the length of the regulated banks is much lower because the reinforcements are mostly located along the main cities.

Water levels at 22 automatic river gauging stations in the Kłodzko Valley and its vicinity are measured in real time with approximately 15-minute temporal resolution by the Local System for Flood Monitoring (Lokalny System Oślony Przeciwpowodziowej – LSOP) of Kłodzko County. The system was designed to provide a riverflow monitoring platform available both for authorities and citizens. The LSOP system has recently been coupled with the innovative HydroProg infrastructure (Niedzielski et al. 2014) to calculate and publish predictions of water levels in real time (Niedzielski and Miziński 2016). The gauges selected for this study – namely Bystrzyca Kłodzka, Kłodzko, Bardo and Żelazno – are covered by the LSOP infrastructure, and hence the unique 15-minute water level data are available at these sites. The data were pre-processed within the HydroProg system to: (1) produce equally sampled 15-minute water-level time series, (2) express time in UTC (Universal Time Coordinated) and (3) remove artefacts from a raw water-level time series using Rosner’s test (Rosner 1983). We limited our analysis to the hydrologic years 2006–2014, namely to the time period 01/11/2005–31/10/2014 according to the Polish definition of hydrologic year.

Each 15-minute, water-level time series was subsequently transformed into discharge. This procedure was feasible since the above-mentioned four gauges are common both for the LSOP gauging network and the state hydrological network which is managed and maintained by the Institute of Meteorology and Water Management – National Research Institute (Instytut Meteorologii i Gospodarki Wodnej – Państwowy Instytut Badawczy, IMGW–PIB, Poland). Tabulated rating curves for these sites have been acquired from IMGW–PIB. After obtaining the 15-minute discharge time series, we produced two daily series, of daily maximum and daily mean discharges.

Additionally, empirical AM discharges from three gauging stations: Bystrzyca Kłodzka, Kłodzko and Żelazno were included in the study. The AM data were obtained from IMGW–PIB. AM were not available for Bardo.

Of note is that measurements from LSOP and from IMGW–PIB are independent of each other. We used the series from LSOP for the estimation of design discharge by means of the POT method and the series from IMGW–PIB for validation. That data differed by origin, which is reflected in different annual maxima peak values in the common data period 2006–2014. The length of AM series was 63 years (Bystrzyca Kłodzka, Kłodzko) and 53 years (Żelazno).

### 3. Methods

#### 3.1. Calibration of the GPD and estimation of design discharge

The calibration of the GPD to POT values, starting from a daily flow series, involves different steps (Willems 2009): (i) separation of the flow time series in quick and slow subflows, (ii) separation of

### Table 1 Catchment characteristics.

| River/gauging station          | Area (km²) | FR (%) | TRI (m) | RD (km/km²) |
|-------------------------------|------------|--------|---------|-------------|
| Nysa Kłodzka/Bystrzyca Kłodzka | 259.9      | 37.9   | 8.5     | 1.3         |
| Nysa Kłodzka/Kłodzko          | 1080.0     | 46.8   | 9.3     | 1.2         |
| Nysa Kłodzka/Bardo            | 1744.1     | 42.1   | 9.2     | 1.2         |
| Biała Lądecka/Zelazno         | 303.1      | 62.3   | 12.5    | 1.1         |

Note: FR, forest ratio; TRI, mean terrain ruggedness index; RD, river density.
the quick subflow series in independent events and selection of the POT values and (iii) calibration
of the GPD.

In (i), a generalization of the recursive digital filter of Chapman (1991) was used to derive the
darflow component in the total flow series, and use it to separate the quick flow from the slow flow
(darflow) runoff subseries.

In (ii), the following criteria were used: the peaks had to exceed a given threshold (higher than
darflow of the quick flow component); the inter-event time between two successive events had to
be greater than the recession constant; and the difference between the lower discharge and the
samples event times, thresholds and fractions were considered. Finally, as a result of (ii), several discharge
optimization (Bernardara et al. 2014).

Based on the criteria in (ii), only independent hydrograph events were considered for (iii). This
were considered in stages (i) and (ii), while only a purely probabilistic approach was used in (iii).
Based on the criteria in (ii), only independent hydrograph events were considered for (iii). This
multi-step procedure is in line with the need both for the physical declustering and for statistical
optimization (Bernardara et al. 2014).

Subsequently, for every subsample \( x_1, ..., x_i (t \leq m) \), the shape parameter \( \gamma \) was estimated
in (iii) as the Hill statistic (Beirlant et al. 1996, Willems 1998, Willems et al. 2007, Onyutha and
Willems 2015),

\[
\hat{\gamma} = \frac{1}{t-1} \sum_{j=1}^{t-1} \ln x_j - \ln x_t
\]

which is the mean excess of the log-transformed data over the threshold \( x_n \), where \( t = \sum_{x_j \geq x_i} 1 \). In
assessing the degree of agreement between the empirical and theoretical GPD cumulative distribu-
tion functions, the Pareto quantile-quantile plot (qq plot) was helpful because the shape parameter
estimate \( \hat{\gamma} \) is equal to the slope of this plot. The Pareto qq plot depicts the logarithms of POT values
versus minus logarithms of exceedance probabilities, \((-\ln (1 - G(x)), \ln x)\). Another helpful plot
was obtained after regressing \( \hat{\gamma} \) against \( x_t \). Two disjoint regions of \( x_t \) were observed in this plot while
\( t \) varied from 1 to \( m \), namely \( S_1 \) where \( \hat{\gamma} \) was approximately stable and \( t \) was not much less than \( m \),
and \( S_2 \) where \( \hat{\gamma} \) strongly fluctuated and \( t \) was far from \( m \). The region \( S_2 \) was of no interest in further
considerations because of the high variance of \( \hat{\gamma} \) for very high thresholds. The final choice of the
POT sample was based on the optimal threshold \( x_n \) selected from \( S_1 \), by minimizing the asymptotic
mean squared error whose estimate is Beirlant et al. (1996)

\[
MSE = \frac{1}{t-1} \sum_{j=1}^{t-1} w_j \left( \frac{\ln x_j}{x_i} - \frac{\hat{\gamma} \ln t}{j} \right)^2.
\]

assuming the Hill weights \( w_j = -\frac{1}{\ln j}, j = 1, ..., t - 1 \). The optimal threshold is a trade-off between
low and high threshold values where the former cause a high bias of \( \hat{\gamma} \) and the latter cause a high
variance of \( \hat{\gamma} \). In case of small POT samples, the very careful choice of \( x_t \) is especially advisable. The
main advantage of the Hill weights is pulling regression towards matching data of low probability of
exceedance as such data are especially important for flood frequency analysis. The Hill estimator \( \hat{\gamma} \)
is consistent and its bias is of order \( \left( \frac{n}{t-1} \right)^{-\sigma \gamma} \) (Beirlant et al. 1996), hence the higher \( \sigma \) the quicker
the convergence of \( \hat{\gamma} \) to \( \gamma \). The asymptotic variance of \( \hat{\gamma} \) is \( \text{Var} (\hat{\gamma}) = \frac{\gamma^2}{t-1} \).

The number of POT values during a period of 1 year is assumed to follow a Poisson distribution.
To verify this hypothesis, a test based on dispersion index \( \Psi \) was used. This index equals 1 for the
poissonian number of exceedances over threshold. Theoretically, \( \Psi \cdot (n - 1) \) asymptotically follows
a \( \chi^2 \) distribution with \( n - 1 \) degrees of freedom where \( n \) is the sample size (Cunnane 1979, Lang
et al. 1999). In this study, many divisions of the observation period in 1-year time slots were considered while exceedances were summed up in every time slot. The test was then completed for every division.

The return period $T$ of a given flow level $x$ was derived as the mean recurrence interval between exceedances over $x$, which is the inverse of the probability that a flow is not less than that level, with respect to $\mu$, the mean number of exceedances per year (Rosbjerg 1985),

$$T(x) = \frac{1}{\mu} \cdot \frac{1}{1 - G(x)}.$$  \hspace{1cm} (4)

The theoretical design discharge $x$ (theoretical flood magnitude) of a certain return period $T$ is a quantile of order $1 - \frac{1}{T\mu}$ of the GPD distribution, assuming that $T\mu > 1$. Provided that $\gamma \neq 0$ and $\sigma = \gamma x_\mu$, this quantile equals to

$$x(T) = x_\mu (T\mu)^{\gamma} \hspace{1cm} (5)$$

Observe that if $\gamma > 0$ then the condition $T\mu > 1$ must hold to ensure discharges to be greater than $x_\mu$.

If the threshold $x_\mu$ is low, bias correction (BC) should be introduced to the GPD parameter estimation. Such a low threshold can be obtained for small sample size when the calibrated GPD not sufficiently follows high quantiles and when the slope $\hat{\gamma}$ in qq plot is considerably different than its asymptotic value. The bias correction was based on the slowly varying function and was modelled according to the formula (Willems et al. 2007)

$$z = -\gamma \infty \ln p - \xi(p) = -\gamma \infty \ln p \left(1 + \frac{\xi(p)}{\gamma \infty \ln p}\right)$$  \hspace{1cm} (6)

where $z = \ln x$, $\lim_{p \to 0} \frac{\xi(p)}{\gamma \infty \ln(p)} = 0$ and $\xi(p)$ is a slowly varying function, $\lim_{p \to 0} \frac{\xi(\lambda p)}{\xi(p)} = 0$ for any $\lambda > 0$. The simplified assumption on $\xi$ was introduced for $p$ values close to 0, namely $\xi(p) = cp^{-\rho}$ (Willems et al. 2007) where $c, \rho$ were constant values. Moreover, $\rho$ is always less than 0 and is more negative for small bias and near to 0 for high bias. Therefore, the estimate of the asymptotic slope $\gamma \infty$ was obtained after shifting of logs of observations according to the equation

$$z + \xi(p) = -\gamma \infty \ln p.$$  \hspace{1cm} (7)

This means that the shift $\xi(p)$ is high for low observations, small for high observations and zero for the highest observation. If $c = 0$, no correction is introduced. In calibration of $\xi$, the minimum of MSE was required. Willems et al. (2007) pointed out that BC usually yields lower threshold value and lower shape estimate than the model without BC.

As a final result of (iii), the parameter estimates of the calibrated GPD were obtained, both for daily maxima flows and daily mean flows.

The goodness of fit between empirical and theoretical distribution functions was tested using the Anderson–Darling (AD) test (Anderson & Darling 1954, Laio 2004),

$$AD = -n - \frac{1}{n} \sum_{j=1}^{n} \left( (2j - 1) \ln G(x_j) + (2n + 1 - 2j) \ln \left(1 - G(x_j)\right) \right)$$  \hspace{1cm} (8)

where $n$ is the sample size.
Remembering that \( p = \frac{1}{T\mu} \), the design discharge was estimated as

\[
x(T) = x_t(T\mu)^{\gamma^\infty} \cdot e^{-c(T\mu)^p}.
\]

Observe that Equation (9) is a generalization of Equation (5) for a model with BC. The FFC was obtained from Equation (9) for every POT sample. Special attention was paid to \( \gamma^\infty \), the strongest factor influencing the FFC for long \( T \).

The shape of FFC was assessed using the derivative of \( x \) with respect to \( T \),

\[
\frac{dx}{dT} = x_t\mu^{\gamma^\infty} T^{\gamma^\infty - 1} \cdot (\gamma^\infty - \rho c(T\mu)^p) \cdot e^{-c(T\mu)^p}
\]

which is the rate of change of the flood magnitude as a function of \( T \), or, in other words, it is the slope (steepness measure) of the FFC.

In Equations (9) and (10) the values \( x_t, \gamma^\infty \) were replaced by their estimates, \( x_t, a, \gamma_a \) calculated from daily maxima and \( x_t, b, \gamma_b \) calculated from daily means. The values of \( \mu_a, \mu_b \) represented the mean number of exceedances per year in POT maxima and POT means sample.

Equation (10) was used for assessing the flood dynamics. High values of \( \frac{dx}{dT} \) reflect the fact that even small changes of return period cause large changes in design discharge. If applied to flood risk assessment, a high value of \( \frac{dx}{dT} \) is a sign of a catchment’s high susceptibility to flooding. Note that if \( 0 < \gamma^\infty < 1 \) then increase of \( T \) to \( +\infty \) causes decrease of \( \frac{dx}{dT} \) to 0, which means that the FFC stabilizes with increasing \( T \).

### 3.2. Comparisons

Our study focused on comparisons between:

(i) \( \gamma_a \) and \( \gamma_b \) at every station,
(ii) \( X_a \) (\( X_b \)) per unit area for every station,
(iii) ln \( X_a \) and ln \( X_b \) for every station,
(iv) rate of change \( \frac{dx}{dT} \) (\( \frac{dx}{dT} \)) per unit area and rate of change for every station,
(v) catchment characteristics that are supposed to mostly influence the magnitude of the difference between \( X_a \) and \( X_b \).

In (ii), (iii) and (iv) the comparisons were made for various return periods.

In (iii), the variable \( U = \ln (X_a) - \ln (X_b) \) was estimated from Equation (9) as:

\[
U(T) = \ln (X_a) - \ln (X_b) = \delta + (\gamma_a - \gamma_b)\ln T - c_a(T\mu_a)^{p_a} + c_b(T\mu_b)^{p_b},
\]

where \( \delta = \ln \frac{x_t\mu_a^{\gamma_a}}{x_t\mu_b^{\gamma_b}} \) is a constant value for every catchment. If \( T \) tends to \( +\infty \) then the last two terms tend to zero. If, additionally, \( \gamma_a > \gamma_b \) then \( U \) increases linearly with logarithm of \( T \). The slower the increase the nearer to zero the difference \( \gamma_a - \gamma_b \) is. Therefore, \( U(T) \) is for large \( T \) mainly influenced by \( \delta \) and \( \gamma_a - \gamma_b \). It is worth pointing out that \( U \) can be considered as the term which is to add to the logarithm of the design flood magnitude estimated from daily mean discharge to receive the logarithm of the design flood magnitude of the same return period estimated from daily maximum discharge. In other words, \( V = \exp(U) \) is the factor by which \( X_b \) has to be multiplied to obtain \( X_a \). The factor \( V \) represents the transformation between the variables \( X_b \) and \( X_a \). This factor increases with \( T \) while the rate of increase depends on the parameters of the GPD. For FFA applications, it is crucial that this factor be properly estimated. From practical point of view, the
consideration of Equation (11) may be limited to the first two terms for large $T$, especially if $\rho_a, \rho_b < -1$, i.e. when the influence of the bias correcting terms is low.

### 3.3. Uncertainty and accuracy of quantile estimates

We also assessed the uncertainty both for $X_a$ and $X_b$. This problem was of high importance because small samples often lead to highly uncertain parameter estimates which, in turn, lead to low confidence of design discharge estimates. This was expressed through the 90% confidence interval around the FFC. For every catchment, the leave-one-out jack-knife bootstrapping method (Efron & Tibshirani 1993, Castellarin et al. 2004, Sahinler & Topuz 2007) was used to estimate the confidence intervals (CIs). Each element, in turn, was dropped from the sample of independent events to obtain the bootstrap jack-knife sample of independent events, and the different GPD calibration steps repeated:

1. Derivation of the parameters of the GPD for threshold values successively increasing from the lowest to the highest value in the bootstrap jack-knife sample,
2. Selection of the final bootstrap threshold by minimizing the MSE (Equation (3)). The threshold was selected from $S_1$, the region of stability of the shape parameter estimate. In that manner, the final POT bootstrap jack-knife sample of values greater or equal to the bootstrap threshold was selected,
3. Estimation of flood quantile magnitudes (Equation (9)) using parameters of the GPD which were derived from the POT bootstrap jack-knife sample.

To include BC, the range of variability of $c$ and $r$ was studied in the original POT sample for threshold values varying in $S_1$. It was noticed that $c$ and $r$ varied by ±30% then. Therefore, such a variability of $c$ and $r$ was imposed when flood quantile magnitudes were estimated from Equation (9).

As a result, a sample of flood magnitudes was obtained for every $T$. The lower and upper confidence limits were calculated as the 5% and 95% quantiles of that sample.

The relative bias was derived to assess the accuracy of quantile estimates. First, probabilities of exceedance and return periods of empirical, ordered quantiles $x_j, j = 1, ..., t$ (j = 1 to the highest) of the original POT sample were calculated as $p_j = \frac{j}{t}, T_j = \frac{1}{m} \cdot \frac{1}{j}$ respectively. Next, $x_j$ were compared to $x(T_j)$, the theoretical quantiles of the same return period (Equation (9)), using

$$
\text{Bias}(\%) = \frac{1}{t} \sum_{j=1}^{t} \frac{x_j - x(T_j)}{x_j} \cdot 100.
$$

The probability plot correlation coefficient (PPCC) was also calculated; this is the Pearson correlation coefficient between theoretical and empirical quantiles.

### 3.4. Validation based on AM data

The model for $X_a$ was validated using the AM data for Bystrzyca Kłodzka, Kłodzko and Żelazno. It should be pointed out that the source of AM discharges differed from the source of daily discharges on the basis of which the GPD was calibrated, as stated in Section 2. AM discharges $y_i, i = 1, ..., n$ were compared to theoretical design discharges $X_a$ of the same return period $T_i$ through the relative bias and root mean squared error, namely

$$
\text{Bias}(\%) = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i - X_a(T_i)}{y_i} \cdot 100, \quad \text{RMSE}(\%) = \left( \frac{1}{n} \sum_{i=1}^{n} \frac{y_i - X_a(T_i)}{y_i} \right)^{\frac{1}{2}} \cdot 100.
$$

Subsequently, Generalised Extreme Value (GEV) distribution was fitted to AM data wherein the parameters of the GEV were estimated using the ML method, and compared with the calibrated
GPD. The shape parameters of the GEV and GPD were also compared because, from theoretical point of view, they should be equal to each other.

Calculations were carried out in WETSPRO, the program designed at the Hydraulics Section, KU Leuven, Belgium (Willems 2009), and in R program (R Core Team 2015).

3.5. Comparison to other studies

Results moreover were compared with results from existing calculation methods of IPF from MDF. The equation $IPF = MDF \cdot (1 + 2.66 \cdot A^{-0.3})$ was proposed by Fuller (1914), where $A$ (km$^2$) is the catchment area. However, according to Fuller (1914) and Dastorani et al. (2013), the $MDF$ should consist of daily mean flows of the days that contain the peaks. It should be noted that days of peaks of $X_a$ and $X_b$ are not always identical. In some cases, the mean daily flow $X_b$ from the POT sample is observed one to three days before or after the day of peak flow $X_a$. The dates are sometimes inconsistent because no $X_b$ is observed in the POT sample around the time of peak flow $X_a$. This inconsistency is discussed in Section 4.5. Summarizing, $MDF \neq X_b$ while $IPF = X_a$. Therefore, to compare results to other studies, the sample $MDF$ had to be extracted directly from the series of daily mean flows while the $IPF$ was assumed as $X_a$.

Other studies were based on $IPF, Q_2$ (equal to $MDF$) and $Q_1(Q_3)$, the instantaneous peak flow, the daily mean flow of the day of peak and the mean daily flow in the previous (next) day to peak, respectively. Two formulas were used, formula (S): $IPF = \frac{Q_1(MDF-Q_3)}{2}$ (Sangal 1983) and formula (FS): $IPF = 0.8 \cdot MDF + 0.25(Q_1+Q_3)/K$, where $K = 0.9123X + 0.3620, X = (Q_1+Q_3)/(2 \cdot MDF)$ (Fill & Steiner 2003).

The approach of Taguas et al. (2008) aimed at establishing regions with a similar rate $k = IPF/MDF$, where $k$ was computed using the Least Squares method. These rates from basins in southeastern Spain and from the four studied catchments were compared. Of note is that the rate $k$ is not equivalent to the factor $V = \exp(U)$ (Equation (11)) because $k$ is based on $X_a$ and $MDF$ while $V$ is based on $X_a$ and $X_b$.

4. Results and discussion

4.1. Calibration of the GPD

To separate the baseflow and to extract independent events, various recession constants, lengths of independency period and other filter parameters were considered. The GPD parameters were calculated using the methods discussed in Section 3.1. The AD test showed a good fit of the POT samples to theoretical GPDs. The AD test statistic varied from 0.12 for daily means based sample from Bardo to 0.52 for daily maxima based sample from Bardo, being less than the critical value equal to 0.6 (Arshad et al. 2003). Subsequently, the BC was introduced. For all series, the condition of minimum MSE yielded lower shape parameter estimates and exactly the same threshold values as before correction. In Figure 2, the qq plot and the bias corrected qq plot for the POT sample extracted from the daily maxima flows in Kłodzko are presented. The slope of the former is equal to $\gamma$ (Equation (2)) while the slope of the latter is equal to $\gamma^\infty$ obtained after shifting of natural logarithms of observations by $\xi(p)$ (Equation (7)). It is confirmed in this plot that $\gamma^\infty < \gamma$. The variance of the shifted sample is lower than that of the original sample. The shift $\xi(p)$ vanishes with increasing $-\ln p$. Similar graphs were obtained for other series.

The final choice of the maxima-based (means-based) POT samples resulted in the following record lengths for Bystrzyca Kłodzka, Kłodzko, Bardo and Żelazno, respectively: 22 (19), 17 (17), 19 (19) and 18 (13), in the following threshold values $x_{\min}(x_{\max})$ [m$^3$·s$^{-1}$]: 33.5192 (22.6987), 82.1256 (52.7362), 97.6497 (69.2408) and 25.7129 (24.1022), and in the following estimates of $\gamma_a(\gamma_b)$: 0.497 (0.470), 0.552 (0.533), 0.536 (0.533) and 0.703 (0.653).
To complete the test based on the dispersion index, the 9-year observation period was first divided into 1-year time slots for every POT sample. To yield the second division, all these slots were shifted simultaneously by the time between the first two events in the POT sample so that the first time slot started at the time of the second event while the first event was omitted. Then, the next division was obtained by shifting the slots by the time between the second and the third event, and so on. To summarize, eight sets of several 1-year slots were considered for every POT sample, creating 64 sets in total. In as many as 53 cases, the test based on the dispersion index did not reject the hypothesis. Therefore, the distribution of the number of exceedances over threshold was found to follow the Poisson distribution in every POT sample.

4.2. Comparisons

4.2.1. Between $\gamma_a$ and $\gamma_b$
Zelazno is featured by the largest difference between these two values, Bystrzyca Kłodzka and Kłodzko by the second and third largest, and for Bardo the parameters are nearly the same. The high difference between $\gamma_a$ and $\gamma_b$ suggests a quick increase of the difference between $x_a$ and $x_b$ with $T$, according to Equation (11). The shape parameter estimates differed considerably from 0 thus the tail of the GPD distribution is heavy for both maxima-based and means-based design discharges.

4.2.2. Between $X_a$ ($X_b$) per unit area for various stations
In Figure 3 the theoretical discharges $X_a$ and $X_b$ per unit area were depicted in FFCs while $T$ varied from 20 to 500. The highest discharge value per unit area for $X_a$ was recognized at the Zelazno outlet. From up- to downstream along the Nysa Kłodzka river, the highest value of $X_a$ per unit area was obtained for the first gauge (Bystrzyca Kłodzka), followed by Kłodzko and Bardo. Hence, more downstream locations show lower discharge magnitude for $X_a$ per unit area. Similar properties were observed for $X_b$. However, for downstream located stations, the differences between $X_b$ per unit area were lower than the differences between $X_a$ per unit area.

The $T$ - year flood magnitudes for $T = 20, 50, 100$ (years) were depicted in Figure 4. The largest values of $X_a$ per unit area are observed for Zelazno, followed by the three gauges more downstream along the Nysa Kłodzka river.

When observing Figure 4 we conclude that the 50-year design discharge $X_a$ per unit area on the Biała Łądecka at Zelazno is comparable to the 100-year design discharge $X_a$ per unit area on the Nysa Kłodzka at Bystrzyca Kłodzka. Further comparisons reveal that the 50-year discharge per unit area at Zelazno is similar to the 200-year flood at Kłodzko.
4.2.3. Between $X_a$ and $X_b$ for every station

For every station, the direct comparison between $X_a$ and $X_b$ per unit area was conducted for various $T$ values. The largest absolute differences for $T = 20, 50$ and $100$ years were noticed at Żelazno, and the lowest at Bardo. The largest relative differences for $T = 20, 50$ and $100$ years were noticed at Bystrzyca Klodzka, reaching 76%, 80% and 84% and the lowest at Bardo, namely 42%, 43%, 43%. What clearly distinguishes Żelazno from other stations is the strong increase of the relative difference from 59% for $T = 20$ years to 72% for $T = 100$ years.

Graphical depiction of the $X_a$ and $X_b$ for $T = 20, 50$ and $100$ years is shown in Figure 5. The quickest increase of the frequency curves is apparent for Żelazno. When logarithmized values are considered, the increase of the curves is limited for high $T$ and the curves are nearly parallel.

A graph of the difference $U = \ln X_a - \ln X_b$ can be seen in Figure 6(a). The largest $U$ is found for Bystrzyca Klodzka, followed by Żelazno, Klodzko and Bardo. The relative difference $(\ln X_a - \ln X_b)/\ln X_b$ is pictured in Figure 6(b). Of note is that it is higher at Bystrzyca Klodzka and
Zelazno, reaching a level of 10.22% and 8.33%, but lower at Kłodzko (7.69%) and Bardo (4.73%). The curves stabilize with increasing $T$.

A limited increase of $U$ with $T$ from 20 to 200 years is observed, namely from 0.5652 to 0.6282 (Bystrzyca Kłodzka), from 0.5189 to 0.5573 (Kłodzko), from 0.3556 to 0.3607 (Bardo) and from 0.4625 to 0.5773 (Zelazno). The large $U$ values at Bystrzyca Kłodzka are due to the large first term $\delta$ in Equation (11), and at Zelazno by the high difference $\gamma_a - \gamma_b$ (second term in Equation (11)). In Table 2 the values of factor $V = \exp(U)$ are displayed for various return periods. Remembering that $V$ is the factor by which $X_b$ has to be multiplied to obtain $X_a$, these values can be used in practice in estimation of $X_a$ from $V$ and $X_b$. The use of the factor $V$ is advised in the future when only daily mean discharges are available to the POT sample and daily maxima values are not routinely recorded at the four gauging stations.

Figure 5. Flood Frequency Curves of $X_a$ and $X_b$ and the 90% confidence intervals: (a) Bystrzyca Kłodzka, (b) Kłodzko, (c) Bardo, (d) Zelazno.

Figure 6. Difference between ln($X_a$) and ln($X_b$): (a) difference, (b) relative difference.
4.2.4. Between rates of change

In Figure 7, the rates of change \( \frac{dX_a}{dT} \), \( \frac{dX_b}{dT} \) (Equation (10)) divided by the catchment areas were depicted. The lowest rate of change of \( X_a \) is revealed for Bardo and the highest for Želazno. This implies that even small change in \( T \) causes a huge change in design flood magnitude for Želazno or, in other words, discharges may go up to very large values in this catchment. The rate of change of \( X_a \) is lower on the Nysa Kłodzka river downstream, hence similar change in \( T \) causes much larger change in design flood per unit area at Bystrzyca Kłodzka than at Bardo. With increasing \( T \), the rate approaches 0 which shows the tendency of the FFC to stabilize with increasing \( T \).

For every catchment, \( \frac{dX_a}{dT} \) is greater than \( \frac{dX_b}{dT} \). Therefore, the \( X_a \) based FFC is featured by a stronger tendency to changes than the \( X_b \) based FFC. Along the Nysa Kłodzka river, the largest discordancy between \( \frac{dX_a}{dT} \) and \( \frac{dX_b}{dT} \) is noticed at Bystrzyca Kłodzka and the lowest at Bardo. For Bystrzyca Kłodzka, Kłodzko, and Bardo, the relative differences between \( \frac{dX_a}{dT} \) and \( \frac{dX_b}{dT} \) are similar for \( T \) from 50 to 200 years, reaching values of 98%, 81% and 44% for \( T = 200 \) years. At Želazno, an increase of this difference is observed, from 79% to 92%. To summarize, Želazno (Bardo) is featured by the largest (lowest) difference between the \( X_a \) based rate of change than of the \( X_b \) based rate of change.

The difference between the rates of change is related to the variability of peak streamflows, which is stronger in small than in large drainage areas (Fill and Steiner 2003, Malamud and Turcotte 2006). However, results from the comparison between \( X_a \) and \( X_b \) and between the rates of change indicate that area is not the only factor to influence the relation between \( X_a \) and \( X_b \) because the catchments at Bystrzyca Kłodzka and at Želazno have a comparable area but a different relation between \( X_a \) and \( X_b \).

4.2.5. Catchment characteristics versus the magnitude of the difference between \( X_a \) and \( X_b \)

The largest magnitude of difference between \( X_a \) and \( X_b \), and the large rate of change of the flood magnitudes along the Biała Łądecka river reflect the fact that this catchment is the most flash-flood prone one among the four catchments. It is worth reminding that the Biała Łądecka catchment at

---

**Table 2** Factor \( V \) for various return periods (in years).

| Catchment  | \( T = 20 \) | \( T = 50 \) | \( T = 100 \) | \( T = 200 \) |
|------------|-------------|-------------|-------------|-------------|
| Bystrzyca K. | 1.7598      | 1.8045      | 1.8390      | 1.8743      |
| Kłodzko    | 1.6897      | 1.7073      | 1.7260      | 1.7470      |
| Bardo      | 1.4287      | 1.4325      | 1.4355      | 1.4387      |
| Želazno    | 1.5881      | 1.6620      | 1.7201      | 1.7813      |

---

**Figure 7.** Rate of change of the \( T \) — year flood magnitude per unit area.
Zelazno is featured by a much larger TRI value than the other three catchments. The issue that high ruggedness can increase susceptibility to flooding was considered, for example, by Patton and Baker (1976) for basins in North Central Utah, USA, and by Bhatt and Ahmed (2014) for catchments located in the Krishna basin, India.

The difference in river regulations is supposed to be another factor that affects the difference between the $X_a$ and $X_b$ based design discharges and the rate of change. The continuous reinforcements along a substantial part of the Biała Łądecka river causes an increase of water velocity and an increase of the instantaneous maximum discharge in comparison to rivers with low number of reinforcements. The mean discharge, in turn, is not as sensitive to changes in velocity as the maximum value, which explains why the difference mentioned above is higher for the Biała Łądecka river than for the Nysa Kłodzka river.

Results of the study of the difference between $X_a$ and $X_b$ suggest that similar comparison could be completed for larger group of catchments. It should be intended to regionalize this group according to factor $V$.

4.3. Uncertainty and accuracy of quantile estimates

The bootstrapping method based CIs of ln $X_a$ and ln $X_b$ are depicted in Figure 5. The uncertainty increases with $T$ and is higher for $X_a$ than for $X_b$. For $20 \leq T \leq 500$ and both for $X_a$ and $X_b$, Klodzko is featured by the largest width of CI, Zelazno by the second largest, followed by Bardo and Bystrzyca Kłodzka. However, the largest width of CI divided by the catchment area is observed at Zelazno (1.00, 1.85, 3.35), followed by Bystrzyca Kłodzka (0.50, 0.80, 1.27), Klodzko (0.43, 0.72, 1.19) and Bardo (0.25, 0.43, 0.70) for $T = 50, 100, 200$ years. Substantial differences between catchments suggest that the estimates are the most reliable for Bardo and the least reliable for Zelazno. The width of CI partially reflects the variability of the shape parameter estimate (Section 3.1) because the largest variance of $\hat{\gamma}$, equal to $\sigma^2_{\hat{\gamma}}$, is detected at Zelazno, followed by Klodzko, Bardo and Bystrzyca Kłodzka, namely 0.028, 0.019, 0.016, 0.012, respectively.

Results of the accuracy analysis are displayed in Table 3. The low values of Bias and high values of PPCC reflect a low discrepancy between observed and estimated quantiles. This shows a high quality of the calibrated GPD model.

4.4. Validation

For Bystrzyca Kłodzka, Klodzko and Zelazno, the AM discharges are depicted in Figure 8 together with the design discharges $X_a$ and design discharges obtained from the GEV, the AM-based $X_{GEV}$. The validation was not completed for Bardo due to the unavailability of AM data. The $X_a$-based calibrated GPD fits very well the highest AM value for all three catchments and fits well the two highest AM values for Bystrzyca Kłodzka and Zelazno. The fit between the lower AM values and the $X_a$ is poorer. Bias and RMSE values are shown in Table 4. The best fit is evident for Bystrzyca Kłodzka, followed by Klodzko and Zelazno. However, for Bystrzyca Kłodzka and Zelazno, this fit is much better for $T > 10$ years.

The shape parameter of the GEV distribution is equal to $0.32 \pm 0.12$, $0.41 \pm 0.13$ and $0.77 \pm 0.17$ for Bystrzyca Kłodzka, Klodzko and Zelazno, respectively. For Bystrzyca Kłodzka and Klodzko, it is

| Table 3 Results of accuracy analysis. |
|--------------------------------------|
| Catchment    | Bias of $X_a$ (%) | Bias of $X_b$ (%) | PPCC of $X_a$ (-) | PPCC of $X_b$ (-) |
|--------------|--------------------|--------------------|--------------------|--------------------|
| Bystrzyca K. | 8.34               | 2.44               | 0.97               | 0.97               |
| Klodzko      | 9.42               | 10.46              | 0.98               | 0.99               |
| Bardo        | 7.96               | 5.75               | 0.95               | 0.99               |
| Zelazno      | 8.42               | 7.72               | 0.95               | 0.98               |
somewhat lower than the GPD shape parameter. However, for Želazno a greater congruence is observed. The GEV fit yields a strong underestimation of the highest AM values in all three catchments and a very strong underestimation of the second highest AM value in Bystrzyca Kłodzka. The difference between the two estimation methods is reflected in Figure 8 which demonstrates how the Hill estimation pulls the FFC leading to a better matching of the highest observations while the ML method leads to underestimation of those highest observations. A possible reason of this underestimation is that the ML method attaches the same importance to every AM observation.

4.5. Comparison to other studies

The comparison, reported in Section 3.5, was carried out between various peak flow estimation methods, namely using the Fuller (1914), Sangal (1983), Fill and Steiner (2003) methods, (F), (S), (FS), and the method based on Hill’s estimates (H), i.e. the formula $X_s = V \cdot X_p$ and Equation (11). It is worth remembering that the days of $MDF$ are the days of peaks, which makes $MDF$ not equal to $X_p$. Thus, in the first three methods, the estimation was not based on the POT means sample but on the $MDF$ at the days of peak, i.e. at the days of $X_s$ while $MDF$ were extracted directly from the raw data of the mean flows. Only in the fourth (H) method, the POT means sample was used. After transformation of the mean daily flows to the maximum daily flows using the four methods, these flows were compared with the POT maxima. Then, the diagnostic measures were used. As evident
in the results presented in Table 5, the method based on (H) estimates yielded a much lower relative bias and relative root mean square error \( \text{rmse} \) between estimated and observed values than any of the methods of (F), (S) and (FS). However, it is worth remembering that both daily maxima and daily means are used in this study to estimate \( V \) while only daily means are used in the (F), (S) and (FS) methods. For this reason, these three methods yield lower quality results.

Regarding the rate \( k = \frac{\text{IPF}}{\text{MDF}} \), this value is equal to 1.65, 1.54, 1.44, 2.18 for Bystrzyca Kłodzka, Kłodzko, Bardo and Żelazno, respectively, and is equal to 1.03 and 1.64 (Group 1) and varies from 2.91 to 4.77 (Group 2) for the 12 catchments located in southeastern Spain studied by Taguas et al. (2008). The two catchments in Group 1 have a small area and a moderate slope and elevation. This proves that some similarity can be found in \( k \) between Group 1 and the four catchments studied. However, the areas of the studied catchments are much larger than those of Group 1. Because Taguas et al. (2008) studied the factor \( k \) only in the context of the relation between sample values of IPF and MDF, it is constant for every catchment. In this study, however, the factor \( V \) between \( X_a \) and \( X_b \) was considered. It is intended for use in FFA, thus to be applicable for design discharges for various return periods. As shown in Table 2, it is noteworthy that \( V \) is different from \( k \) in Bystrzyca Kłodzka, Kłodzko, Bardo and Żelazno, and nearly equal to \( k \) in Bardo. A deeper investigation into the POT maxima and POT means samples revealed that the dates of these two series are inconsistent. The ratio of the number of days during which both a maximum and mean value belong to the POT samples to the length of the POT means sample is only 0.74 (Bystrzyca Kłodzka), 0.71 (Kłodzko), 0.63 (Bardo) and 0.77 (Żelazno). This low degree of agreement precludes the applicability of the (F), (S) and (FS) methods of derivation of peak flows from daily mean flows to the POT means sample in order to estimate design discharges in these catchments. Therefore, these three methods cannot be applied together with the POT approach.

5. Conclusions

The discrepancies between \( X_a \) and \( X_b \) enable us to draw the conclusion that attention should be paid to lacking availability of maximum daily discharges and when design flood discharges are assessed on the basis of daily means instead. This statement is crucial for rivers with high flow dynamics, as the Biała Łądecka river at Żelazno and the Nysa Kłodzka river at Bystrzyca Kłodzka.

Results of this study allowed us to make the following conclusions and recommendations:

1. For long \( T \), the FFCs of the logarithmized \( X_a \) and \( X_b \) design flows vary, approximately, by a constant value. Therefore, the flood magnitudes \( X_a \) can, for given return periods, be obtained by multiplication of \( X_b \) by the exponent of this value,
2. The difference between \( X_a \) and \( X_b \) based design discharges depends on area and on other catchment characteristics,
3. The highest absolute and relative differences between \( \ln X_a \) and \( \ln X_b \) and the highest difference between rate of change of FFC are obtained for catchment with large TRI, and for river with continuous system of protection measures due to many human settlements along river banks,
4. The magnitude of the differences decreases from up- to downstream,
(5) the magnitude of the differences reflects the susceptibility to flooding of the studied catchments,
(6) using the GPD calibration and bias correction method presented in this paper, the factor $V$
between the daily maxima based and daily means based discharges can be obtained,
(7) the methods of Fuller (1914), Sangal (1983), Fill and Steiner (2003) for calculation of peak
flows from daily mean flows cannot be applied to POT means samples for the estimation of
design discharges in the four catchments studied,
(8) in case of accessibility of data, the estimation of factor $V$ should be completed again for a
longer period in the four catchments studied,
(9) using the tools used in this paper, the estimation of $V$ can be carried out also for other catch-
ments in order to compare results from various catchments of different hydrological
regimes,
(10) results of the study suggest the possibility of regionalizing catchments according to the fac-
tor $V$.

Acknowledgements
The investigation described in this paper has been partially performed during the first author’s research visit to the
Hydraulics Section of the KU Leuven, in 2014 within the EU-funded Project POKL entitled ‘Wzmocnienie potencjału
dydaktycznego UR’, No. UDA-POKL.04.03.00-00-00-165/12-00. This research was also partially supported by the
Ministry of Science and Higher Education of the Republic of Poland. The research has also been supported by the
National Science Centre, Poland, through grant no. 2011/01/D/ST10/04171 under leadership of Dr hab. Tomasz Nied-
zielski, Professor at the University of Wroclaw, Poland. All these supports are gratefully acknowledged. The authors
kindly acknowledge the authorities of the County Office in Kłodzko for providing the hydrometeorological data col-
llected by the Local Flood Monitoring System (Lokalny System Ochrony Przeciwpowodziowej – LSOP). The authors
also acknowledge IMGW–PIB for AM data and for rating curves. We are indebted to Dr Małgorzata Wieczorek who
kindly provided us with values of TRI and RD characteristics. We thank Mrs Matylda Witek for discussions on river
regulation in the upper Nysa Kłodzka basin. We also thank Mr Bartłomiej Miziński who participated in the design
and implementation of the pre-processing data modules for the HydroProg system. We thank Caroline Wojtylak for
her English language support and proofreading help.

Disclosure statement
No potential conflict of interest was reported by the authors.

Funding
EU-funded Project POKL entitled ‘Wzmocnienie potencjału dydaktycznego UR’ [grant number UDA-
POKL.04.03.00-00-00-165/12-00]; Ministry of Science and Higher Education of the Republic of Poland; National Sci-
ence Centre, Poland [grant number 2011/01/D/ST10/04171].

ORCID
Agnieszka Rutkowska  http://orcid.org/0000-0002-5418-5659
Patrick Willems  http://orcid.org/0000-0002-7085-2570
Tomasz Niedzielski  http://orcid.org/0000-0001-8236-9861

References
Anderson T, Darling D. 1954. A test of goodness of fit. J Am Stat Assoc. 49: 765–769.
Arshad M, Rasool M, Ahmad M. 2003. Anderson Darling and modified Anderson Darling tests for generalized Pareto
distribution. J Appl Sci. 3: 85–88.
Balkema A, de Haan L. 1974. Residual life time at great age. Ann Probab. 2: 792–804.
Latocha A, Witek M. 2013. Dawne obiekty hydrotechniczne na ziemi kłodzkiej: stan zachowania i rola w przebiegu procesów fluwialnych, w tym wezbrań ekstremalnych [Old hydrotechnical constructions in Kłodzko Land; present condition and role in controlling fluvial processes, including extreme high flows]. In: Kościć E. editor. Gdy nadtacjała wielka woda: klęski powodzi na ziemiach polskich na przestrzeni wieków [When high water was approaching: flood catastrophes in Poland over centuries]. Wydawnictwo GAJT; pp. 105–125. Polish.

Madsen H, Rasmussen P, Rosbjerg D. 1997. Comparison of annual maximum series and partial duration series methods for modelling extreme hydrologic events. 1: at-site modelling. Water Resour Res. 33:747–757.

Madsen H, Rosbjerg D. 1997a. Generalized least squares and empirical Bayes estimation in regional partial duration series index-flood modeling. Water Resour Res. 33:771–181.

Madsen H, Rosbjerg D. 1997b. The partial duration series method in regional index-flood modeling. Water Resour Res. 33:737–746.

Malamud B, Turcotte D. 2006. The applicability of power-law frequency statistics to floods. J Hydrol. 322:168–180.

Martins E, Stedinger J. 2001. Generalized maximum-likelihood Pareto-Poisson estimators for partial duration series. Water Resour Res. 37:2551–2557.

Niedzielski T. 2007. A data-based regional scale autoregressive rainfall-runoff model: a study from the Odra River. Stoch Environ Res Risk Assess. 21:649–664.

Niedzielski T. 2010. Empirical hydrologic predictions for southwestern Poland and their relation to ENSO teleconnections. Artif Sat. 45:11–26.

Niedzielski T. 2011. Is there any teleconnection between surface hydrology in Poland and El Nino/Southern Oscillation? Pure Appl Geophys. 168:871–886.

Niedzielski T, Miziński B. 2016. Real-time hydrograph modelling in the upper Nysa Kodzka river basin (SW Poland): a two-model hydrologic ensemble prediction approach. Stoch Environ Res Risk Assess. 168:871–886.

Niedzielski T, Miziński B, Kryza M, Netzef P, Wieczorek M, Kasprzak M, Kosek W, Migoni P, Szymanowski M, Jezierska J, Witek M. 2014. HydroPro: a system for hydrologic forecasting in real time based on the multimodelling approach. Meteorology Hydrology Water Manag – Res Oper Appl. 2:65–72.

Onyutha C, Willems P. 2015. Uncertainty in calibrating generalized Pareto distribution to rainfall extremes in lake Victoria basin. Hydrol Res. 46:356–376.

Patton P, Baker V. 1976. Morphometry and floods in small drainage basins subject to diverse hydrogeomorphic controls. Water Resour Res. 12:941–952.

Pickands III J. 2000. Statistical inference using extreme order statistics. Ann Statist. 3:119–131.

R Core Team. 2015. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.

Riley S, DeGloria S, Elliot R. 1999. A terrain ruggedness index that quantifies topographic heterogeneity. Int J Sci. 5:23–27.

Robson A, Reed D. 1999. Statistical procedures for flood frequency estimation. Volume 3 of the flood estimation handbook. Wallingford: Centre for Ecology & Hydrology.

Rosbjerg D. 1985. Estimation in partial duration series with independent and dependent peak values. J Hydrol. 76:183–195.

Rosner B. 1983. Percentage points for a generalized ESD many-outlier procedure. Technometrics, 25:165–172.

Rutkowska A, Banasik K. 2014. The shape parameter of the GEV and GP distributions of Annual Maxima and Peak Over Thresholds discharges – statistical analysis. Monografie Komitetu Gospodarki Wodnej PAN z XX; p. 95–106. Polish.

Sahinler S, Topuz D. 2007. Bootstrap and jackknife resampling algorithms for estimation of regression parameters. J Appl Quant Methods. 2:188–199.

Sangal B. 1983. Practical method of estimating peak flow. J Hydraulic Eng. 109:549–563.

Sen A, Niedzielski T. 2010. Statistical characteristics of riverflow variability in the Odra River basin, Southwestern Poland. Polish J Environ Stud. 19:387–397.

Serinaldi F, Kilsby C. 2014. Rainfall extremes: toward reconciliation after the battle of distributions. Water Resour Res. 50:336–352.

Shabani M, Shabani N. 2012. Application of artificial neural networks in instantaneous peak flow estimation for Kharestan Watershed, Iran. J Resour Ecol. 3:379–383.

Smith J. 1989. Regional flood frequency analysis using extreme order statistics of the annual peak record. Water Resour Res. 25:311–317.

Strupczewski WG. 1967. Determination of the probability of repeating phenomena. Acta Geophys Polonica. 15:147–158.

Taguas EV, Ayuso JL, Pena A, Yuan Y, Sanchez MC, Giraldez JV, Pérez R. 2008. Testing the relationship between instantaneous peak flow and mean daily flow in a Mediterranean area Southeast Spain. Catena. 75:129–137.

Taye M, Willems P. 2011. Influence of climate variability on representative QDF predictions of the upper Blue Nile basin. J Hydrology. 411:355–365.
Tucci C. 1991. Regionalização de Vazões no Rio Grande do Sul [Flow Regionalization in Rio Grande do Sul]. Porto Alegre: IPH/UFRGS. Portuguese.

USWRC. 1982. Guidelines for determining flood flow frequency. Bull. 17B. Washington: Hydrol. Comm., Water Resour. Counc.

Van Montfort M, Witter J. 1985. Testing exponentiality against generalised Pareto distribution. J Hydrol. 78:305–315.

Willems P. 1998. Hydrological applications of extreme value analysis. In: Wheater H, Kirby C, editors. Hydrology in a changing environment. vol. 3. Chichester: John Wiley; p. 15–25.

Willems P. 2009. A time series tool to support the multi-criteria performance evaluation of rainfall-runoff models. Environ Model Software. 24:311–321.

Willems P, Guillou A, Beirlant J. 2007. Bias correction in hydrologic GPD based extreme value analysis by means of a slowly varying function. J Hydrology 338:221–236.

Witek M. 2012. Wpływ zabudowy hydrotechnicznej na współczesne kształtowanie rzeźby koryt rzek ziemi kłodzkiej [The impact of hydraulic engineering structures on present-day morphology of Kłodzko County river channels]. Landform Anal. 19:91–102. Polish.

Zielińska M. 1965. Sposoby zestawiania danych hydrometeorologicznych dla opracowań statystycznych. Gospodarka Wodna 7:245–251. Polish.