An application of the Bivariate Generalized Pareto Distribution for the probabilities of low flow extremes estimation

W. Jakubowski

Hydrol. Earth Syst. Sci. Discuss., 3, 859–893, 2006
www.hydrol-earth-syst-sci-discuss.net/3/859/2006/

Hydrology and Earth System Sciences Discussions

Papers published in *Hydrology and Earth System Sciences Discussions* are under open-access review for the journal *Hydrology and Earth System Sciences*

An application of the Bivariate Generalized Pareto Distribution for the probabilities of low flow extremes estimation

W. Jakubowski

Dept. of Mathematics, Agricultural University of Wroclaw, Poland

Received: 3 March 2006 – Accepted: 23 March 2006 – Published: 6 June 2006

Correspondence to: W. Jakubowski (wj@ozi.ar.wroc.pl)

Abstract

The two-dimensional Bivariate Generalized Pareto Distribution (BGPD) of Tajvidi (1996) is applied in order to estimate the extreme values of the low flow deficit amounts and durations probabilities. Eight parameters BGPD depends on two one-dimensional distributions – Univariate Generalized Pareto Distributions (UGPDs). Each of these three parameter UGPDs describes the probability of one of low flow indices. To fit BGPD into observed data a three steps method of estimation is proposed: (1) For a given shift parameter of each UGPD two others are estimated by the maximum likelihood method. (2) For given shifts and the UGPD parameters estimated in the first step the remaining ones, connected to the bivariate distribution function formula, are also estimated by the maximum likelihood method. (3) The best shift pair is chosen by maximization of the correlation coefficient of the estimated BGPD. The results are applied to statistical description of the low flow index extremes behaviour at four different catchments profiles. To extract the low flow time series data the standard constant threshold level method is applied. Finally the estimated probabilities are compared to the Zelenhasić and Salvai (1987) model.

1 Introduction

Before starting the statistical elaboration of a low flow characteristic extreme, a method of the low flow indices extracting have to be chosen. It should be determined which flows are low and how the seasonality effects on the extracted low flows data time series. Then proper low flow indices should be defined. Full description of the low flow definitions can be found in monograph edited by Tallaksen and van Lanen (2004). Here only the indices extracted by the constant threshold level method are considered and probabilities of their extremes are estimated. To obtain such probabilities the two-dimensional Bivariate Generalized Pareto Distribution (BGPD) (Tajvidi, 1996) is used.

2 Low flow definition

By using the threshold level method two sequences of low flow indices can be determined from the time series of daily runoffs:

- deficit volumes [in m^3]: $D_n = \int\limits_{t_{n_o}}^{t_{n_1}} (Q_L - Q(t))\, dt$;
- durations [in days]: $T_n = t_{n_1} - t_{n_0} + 1$,

where t_{n_1} is the n-th low flow of the last day, t_{n_0} is the n-th low flow of the first day, $Q(t)$ stands for the runoff of t-th day, and Q_L is the threshold level.

In practice, such low flows time series consist of a big number of minor observations often mutually dependent on one another. That is why additional criteria should be applied. Following Zelenhasić and Salvai (1987), the following restrictions can be imposed; a single drought with duration shorter then the minimum drought duration or with a deficit lower then $\alpha D_{\max}$ are removed from the extracted low flow time series. Here $D_{\max}$ denotes the observed maximum deficit and the fractional coefficient α are often set to 0.005. The next restriction is set in the form of the inter-drought criterion: two droughts separated by an interval lasting shorter then the critical duration are pooled to one another. An example of a typical low flow series observed on the Odra River (years 1982–1984) in the Polish Lowland[1] is presented in Fig. 1.

3 Univariate Generalized Pareto Distribution of low flow indices maximum

Let $\{A_i\}_{i=1,\ldots,n}$ be a sequence of mutually independent, identically distributed random variables with distribution $F(x) = \Pr(A_i \le x)$. Sequence $\{A_i\}_{i=1,\ldots,n}$ stands either for

[1] Daily runoff data are obtained from Institute of Meteorology and Water Management, Wroclaw Branch, Poland.

HESSD

3, 859–893, 2006

Bivariate distribution of the low flow extremes estimation

W. Jakubowski

EGU

deficit volume or its duration. In order to determine of maximum values $\max_{i=1,\ldots,n} A_i = M_n$ the Univariate Generalized Pareto Distribution (UGPD) can be applied.

As in the work of Pickands (1975) UGPD is defined for $\kappa<0$ as follows:

$$H(x,\kappa,\sigma) = 1 - \left(1 - \kappa\frac{x}{\sigma}\right)^{\frac{1}{\kappa}}, \qquad \bar{H} = 1 - H, \tag{1}$$

It can be applied to determining the probability of the extremes estimation if and only if

$$\lim_{b\to x_F} \inf_{0<\sigma<\infty} \sup_{0\le x<\infty} |F_b(x) - H(x,\kappa,\sigma)| = 0, \tag{2}$$

where: $F_b(x) = \Pr(A<x+b\,|A>b)$ denotes the conditional distribution of the excess of the random variable $A=A_i$, $i=1,\ldots,n$, over the threshold b on the condition that $x_F = \sup\{x{:}F(x)<1\}$. The above limit (2) is equal to 0 under the assumption that for any $a_n>0$ and b_n there exists the limit

$$\lim_{n\to\infty} \Pr\left(\frac{M_n - b_n}{a_n} \le x\right) = \lim_{n\to\infty} F^n(a_n x + b_n) = G(x), \tag{3}$$

and $G(x)$ is a non-degenerate distribution function. Let us mention that $G(x)$ is called the Generalized Extreme Value Distribution.

For further investigation we denote by

$$H_b(x,\kappa,\alpha) = 1 - (1 - \alpha\kappa(x+b))^{\frac{1}{\kappa}}, \quad x \ge 0,\ \alpha > 0,\ \kappa < 0, \tag{4}$$

the shifted three parameter UGPD with transformed scale parameter α. It is easy to verify that the conditional distribution

$$\begin{aligned}\Pr(A \le x + \beta\,|A > \beta) &= \tfrac{H_b(x+\beta,\kappa,\alpha)-H_b(\beta,\kappa,\alpha)}{1-H_b(\beta,\kappa,\alpha)} \\ &= 1 - \left(1 - \tfrac{\alpha}{1-\alpha\kappa(b+\beta)}\kappa x\right)^{\frac{1}{\kappa}} = H_0(x,\kappa,\alpha_{b+\beta})\end{aligned} \tag{5}$$

is also a UGPD with another scale parameter and shift parameter b=0. The expected values for the probability density function $h_0(x,\kappa,\alpha)$ of $H_0(x,\kappa,\alpha)$ can be obtained by

$$\begin{aligned}
EA &= \int_0^\infty x h_0(x,\kappa,\alpha_b)dx = \frac{1}{\alpha_b(\kappa+1)}, && \text{for } \kappa \in (-1;0),\\
EA^2 &= \int_0^\infty x^2 h_0(x,\kappa,\alpha_b)dx = \frac{2}{\alpha_b^2(\kappa+1)(2\kappa+1)}, && \text{for } \kappa \in (-0,5;0),\\
E(A|A>\beta) &= \frac{1-\alpha_b\kappa\beta}{\alpha_b(\kappa+1)} = \frac{1}{\alpha_b(\kappa+1)} - \frac{\kappa\beta}{(\kappa+1)}, && \text{for } \kappa \in (-1;0).
\end{aligned} \tag{6}$$

Note that

- the expected values are finite only for limited values of κ;
- $E\left(A|a>\beta\right)$ is a linear function of β. It means that in stable conditions the expected value EA obtained for different shifts b should grow similarly to the conditional expected value $E\left(A|a>\beta\right)$;
- the shift parameter b set as a threshold level decreases the number of low flow events (its connection to the other threshold level Q_L used for low flow indices definition is only indirect) and because of estimation aspects cannot grow too high.

Practically, these conclusions are the basis of the method of UGPD unknown parameters estimation. For a given shift parameter b the two remaining parameters can be estimated by using the maximum likelihood method. Then estimators satisfying the above conditions have to be chosen. The linear increment of the conditional expected values suggests that the estimator $\hat{b}$ can be fixed by minimizing the mean square error. This approach has been applied to daily streamflow by Hisdal et al. (2002), Engeland et al. (2004) or Jakubowski (2005). The results show difficulties in proper estimation of the shift parameter b. For some low flow examples fitting the observed annual or seasonal maximum drought indices into estimated UGPD is not satisfactory. They are often distinctly worse than those obtained by the Zelenhasić and Salvai (1987) method.

4 Dependence between low flow indices

The defined above low flow indices of deficit amounts and durations are strictly dependent on one another. An example of this dependency is presented in Fig. 2. For extracting low flow events all the daily runoffs from the years 1966–2003 of the profile Cigacice on the Odra River is considered. The threshold level Q_L is put at 70% the other restriction parameters are set as follows:

- fractional coefficient $\alpha=0,005$;
- minimum drought duration – 5 days;
- drought separating interval duration – 3 days.

In Fig. 2 the asterisks denote the observed annual deficit or duration maximums, the crosses denote the other significant droughts. One can notice, especially for small low flows, that the dependence between these two indices is not linear. Along with the increase of the low flow durations the tendency of growing deficit increasing becomes even stronger. It seems that these nonlinear tendencies are the cause of the difficulties in estimating of the one-dimensional indices distribution.

5 Bivariate Generalized Pareto Distribution

5.1 Definition

Let $\left\{A_i^{(1)}, A_i^{(2)}\right\}_{i=1,\ldots,n}$ be a sequence of mutually independent identically distributed random variables with distribution function $F(x,t)$. Define as in the one-dimensional case $(M_n^{(1)}, M_n^{(2)})=(\max\limits_{i=1,\ldots,n} A_i^{(1)}, \max\limits_{i=1,\ldots,n} A_i^{(2)})$. If for any $a_n^{(1)}>0,\ a_n^{(2)}>0,\ b_n^{(1)},\ b_n^{(2)}$ exists

HESSD
3, 859–893, 2006

Bivariate distribution of the low flow extremes estimation

W. Jakubowski

EGU

the limit

$$\lim_{n\to\infty} \Pr\left(\frac{M_n^{(1)} - b_n^{(1)}}{a_n^{(1)}} \le x, \frac{M_n^{(2)} - b_n^{(2)}}{a_n^{(2)}} \le t\right) = G(x,t) \tag{7}$$

and if $G(x,t)$ is a non-degenerate distribution function, then $G(x,t)$ is a Bivariate Generalized Extreme Value Distribution (Resnick, 1987; Coles, 2001). According to Tajvidi (1996), $H{=}H(x,t)$ belongs to the family of Bivariate Generalized Pareto Distributions (BGPD) with positive support if

$$\bar{H}(x,t) = \frac{-\ln G(x+x_0, t+t_0)}{-\ln G(x_0,t_0)}, \quad x,t > 0, \tag{8}$$

$$H(x,t) = \begin{cases} 1-\bar{H}(x,t), & \text{where} \quad x,t>0, \\ 0, & \text{otherwise,} \end{cases}$$

for some extreme values distribution G with (x_0, t_0) in the support of G. It can be also shown by Tajvidi (1996) that

$$\bar{H}(x,t){=}1{-}H(x,t){=}\Pr\left((D_M,T_M)\nleq(x,t)\right) = \left(\frac{\bar{F}_d^p(x){+}k\bar{F}_d^{p/2}(x)\bar{F}_t^{p/2}(t){+}\bar{F}_t^p(t)}{\bar{F}_d^p(0){+}k\bar{F}_d^{p/2}(0)\bar{F}_t^{p/2}(0){+}\bar{F}_t^p(0)}\right)^{\frac{1}{p}}, \tag{9}$$

where

$$\bar{F}_d(x) = (1-\alpha_d\kappa_d(b_d+x))^{1/\kappa_d}, \quad \bar{F}_t(t) = (1-\alpha_t\kappa_t(b_t+t))^{1/\kappa_t}, \tag{10}$$

$$0 \le k \le 2(p-1), \quad p \ge 2, \quad \kappa_d, \kappa_t \in (-1;0), \quad \alpha_d, \alpha_t > 0,$$

belongs to the BGPD family. Note that the distribution functions $F_d(x)$, $F_t(t)$ are UGPDs and they describe each of the indices separately.

5.2 Some properties of Bivariate Generalized Pareto Distribution

1. Marginal distributions of $H(x,t)$. By letting t tend to infinity one can obtain

$$\bar{H}(x,\infty) = \Pr(D_M > x) = \frac{\bar{F}_d(x)}{\left(\bar{F}_d^p(0) + k\bar{F}_d^{p/2}(0)\bar{F}_t^{p/2}(0) + \bar{F}_t^p(0)\right)^{\frac{1}{p}}}, \tag{11}$$

whence the distribution $H(x,\infty)$ has a positive value (probability jump) for $x=0$. Marginal conditional probabilities

$$\Pr\left(D_M > x+x_0 | D_M > x_0\right) = \frac{\bar{H}(x+x_0,\infty)}{\bar{H}(x,\infty)} = \left(1 - \frac{\alpha_d}{1-\alpha_d\kappa_d(b_d+x_0)}\kappa_d x\right)^{\frac{1}{\kappa_d}} \tag{12}$$

are UGPDs with scale parameter $\alpha = \frac{\alpha_d}{1-\alpha_d\kappa_d(b_d+x_0)}$.

2. Conditional probability

$$\Pr\left((D_M,T_M) \nleq (x+x_0, t+t_0) | (D_M,T_M) \nleq (x_0,t_0)\right) = \frac{\bar{H}(x+x_0,t+t_0)}{\bar{H}(x_0,t_0)} \tag{13}$$

is also BGPD with shift parameters b_d+x_0 and b_t+t_0.

3. Support of $H(x,t)$. Positive shift parameters b_d and b_t (in Fig. 3 checked as simple perpendicular lines) divide a BGPD domain of researched maximums $(x_M, t_M) \in R_+^2$ into four areas – A, B, C, Z. Points

   - with $x_M \geq b_d$ and $t_M \geq b_t$, (A – area) obviously belong to the domain;
   - with $x_M < b_d$ and $t_M > b_t$ (B – area) or $x_M > b_d$ and $t_M < b_t$ (C – area) are projected on the straight lines $x_M = b_d$ or $t_M = b_t$ respectively;

– with $x_M \le b_d$ and $t_M \le b_t$ do not belong to the domain (Z-area).

Therefore, with the increasing of the shift parameters values, the number of events important for the low flow extreme indices estimation decreases. (Refer to Sect. 3.1.)

4. The density function of BGPD $H(x,t)\ x,t>0$ consists of three components: $h(x,t)$; $h(0,t)$; $h(x,0)$. Note that the first density function is two-dimensional, however the other two are one-dimensional only with the support in the shift lines. The integrals: $\int\limits_0^\infty \int\limits_0^\infty h(x,t)dxdt = \frac{\bar{F}_d(0)+\bar{F}_t(0)}{V} - 1$, $\int\limits_0^\infty h(x,0)dx = 1 - \frac{\bar{F}_t(0)}{V}$, $\int\limits_0^\infty h(0,t)dt = 1 - \frac{\bar{F}_d(0)}{V}$, where

$$V = \left(\bar{F}_d^p(0) + k\bar{F}_d^{\frac{p}{2}}(0)\bar{F}_t^{\frac{p}{2}}(0) + \bar{F}_t^p(0)\right)^{\frac{1}{p}}, \tag{14}$$

give the probability mass of each of the areas.

5. Correlation coefficient of BGPD (for precise formulas see Appendix A). It value is calculated for the area A only, and it will be apply below for the estimation of the shift parameters. Using the standard methods one obtains:

$$\begin{aligned}
ED_M &= \int\limits_0^\infty \int\limits_0^\infty x h(x,t)dxdt, && \text{for } \kappa_d \in (-1;0),\\
ED_M^2 &= \int\limits_0^\infty \int\limits_0^\infty x^2 h(x,t)dxdt, && \text{for } \kappa_d \in \left(-\tfrac{1}{2};0\right),\\
ED_MT_M &= \int\limits_0^\infty \int\limits_0^\infty xt h(x,t)dxdt, && \text{for } \kappa_d + \kappa_t \in (-1,0),
\end{aligned} \tag{15}$$

where $h(x,y)$ is the two-dimensional probability density function of $H(x,y)$. The above expected values depend on UGPDs: $\bar{F}_d, \bar{F}_t$ and integrals

$\int_0^{\alpha} \left(1+kt+t^2\right)^{\beta} t^{\gamma} dt$. The moments ET_M, ET_M^2 can be calculated similarly to ED_M, ED_M^2. It means that the correlation coefficient exists for the parameters $\kappa_d, \kappa_t \in \left(-\frac{1}{2}; 0\right)$ only.

6 Estimation of maximum low flow deficit and duration distributions

6.1 Estimation of the BGPD

Estimation is performed for two-dimensional observations of low flow deficits and durations. To estimate the distributions of extreme indices the above defined BGPD (Eq. 9) is applied. As it can be seen above, the BGPD depends on eight parameters. Six of them are connected with two 3-parameter UPGDs. The final two p, k are related to the form of the two-dimensional formula (Eq. 9). For the estimation the following method is applied:

1. For a given pair of shift parameters b_d, b_t the four of them $(\hat{\alpha}_d, \hat{\kappa}_d, \hat{\alpha}_t, \hat{\kappa}_t)$ are estimated by the maximum likelihood method, for each of the one-dimensional indices separately. Each sequence of the index observations is decreased by a shift parameter then the standard maximum likelihood method is applied. The goodness of fit for each of the UPGDs is achieved by using the chi-square test. To return to the initial values the estimated UGPD $\hat{\alpha}$ parameter is converted to $\hat{\alpha}'$ using the conditional probability formula (Eq. 5). For further estimation the pairs of the shift parameters which do not reject the goodness of fit tests are considered only. The two last $\hat{p}$, $\hat{k}$ are estimated by the BGPD using the maximum likelihood method as well. The chosen shift parameters sequences are equal to the successive ordered low flow deficits and durations.

2. The best shift pair (b_d, b_t) is chosen by the maximization of the correlation coefficient. This assumption is made because of the non-homogeneity of observed

low flows. Upon observation, it is suffice to analyse the nonlinear dependence between deficit amounts and durations (for instance see Fig. 2 above). Along with the increase of the deficit duration the volumes grow much quicker – this is clearly visible for short durations. By taking the maximum correlation coefficient the homogeneity of processed low flow observations is stabilized. To determine the correlation coefficient the integrals $\int_0^{\alpha}\left(1+kt+t^2\right)^{\beta} t^{\gamma} dt$ should be computed (see Appendix A). The Gauss-Jacobi quadrature method is applied. All computations are carried out for shift pairs with connected estimators $\hat{\kappa}_d, \hat{\kappa}_t$ in the interval $(-0.5; 0)$ only. Other pairs, where at least one $\hat{\kappa}$ stays outside the interval $(-0.5; 0)$, are omitted.

3. The goodness of fit of the estimated one-dimensional marginal distributions of extreme annual or seasonal index extremes is obtained by λ-Kolmogorov goodness of fit test.

6.2 Application of BGPD for determining the probabilities of low flow extremes indices

For the presentation of the above estimation method four catchments are chosen. The first catchment – about 40 thousands km^2 on the Odra River (Cigacice gauges) is situated in the Polish Lowland. The second, is a small (50 km^2) Sudety Mountains catchment – Międzylesie profile on the Nysa Klodzka River. The third, is a New Zealand highland catchment Kuripapango on the Ngaruroro River (370 km^2) and the last one Colwick, UK, on the Trent River (about 7500 km^2). For profiles Cigacice[2], Colwick and Kuripapango all daily observed runoffs are taken into consideration. For Międzylesie, however only the summer (May–October) daily runoffs are considered. The annual average precipitation is varying from 580 mm (Cigacice) to over 2000 mm at Kuripapango.

[2]Polish daily runoff data are obtained from Institute of Meteorology and Water Management Wroclaw Branch, Poland, other from assembled by the ASTHyDA project Global Data Set.

For the two remaining profiles the average precipitation amount to 760 mm (Colwick) and 870 mm (Międzylesie).

Using the threshold level method with the Zelenhasic and Salvai restrictions the observed low flow deficit amounts and durations are extracted. The threshold level is set at $Q_{70\%}$, minimum drought duration is put at 5 days, separation criteria at 3 days and the coefficient α is set at 0.005. The estimated best pairs of shift parameters (b_d, b_t) and the correlation coefficients are shown in Table 1.

The estimation results are shown in Figs. 4–7. They present estimated two-dimensional probability plots of extreme low flow deficit amounts and durations. As in Fig. 2 asterisks denote the observed annual or summer maximums of deficits or durations, crosses refer to other significant observed droughts. Straight lines depict the estimated best shift parameters. The quantile curves – constant value probability lines determine the areas laying left or below them whose estimated probabilities of non-exceedance (Eq. 9) are equal to 50, 80, 90 and 95%.

Note that shift lines are generally dividing the observed low flow events in two classes. The first one (in Fig. 3 denoted as Z-area) consists of many small insignificant low flows. The deficit amount is slower in its increase depending on their duration. The second one (in Fig. 3 A-area) contains smaller number of greater low flows with quicker deficit amount increasing. Such a difference confirms the earlier assumed non-homogeneity of the observed low flows. Practically, it means that these smaller low flows are caused by other hydrological processes than the greater ones. Because of few observed events, the two remaining areas (B, C) are exerting a small influence on the bivariate probability behaviour.

Considered above non-homogeneity is observable in the graphs of the quantile curves similarly. Especially for lower probabilities (in figures for Pr =50%), they tend to be moved to the left. This theoretically causes the lowering of the estimated duration quantiles. As it can be seen below, this effect does not transmit to the marginal distributions, so for further calculation it is not taken into consideration.

Taking the marginal distributions (Eq. 11) of the estimated BGPD one can compute the probability of each of the examined maximum values of the indices. In Figs. 8–11 the probabilities of the non-exceedance of the low flow extreme deficit amounts and durations are presented. Because of the positive probability values of the marginal distribution $H(x,\infty)|_{x=0}$ and $H(\infty,t)|_{t=0}$ the one-dimensional distributions of low flow indices have a discontinuous probability jump set at the estimated shift parameter. These jumps, as it was shown in Eq. (11), depend on behaviour of the both researched indices. And of course they are distinctly lowering the probabilities values obtained by the investigation of the UGPD only. We note that the probability jumps of the deficit amount are laying much lower than the respective probabilities for the low flow durations. It also follows from the non-homogeneity of the observed magnitudes of the low flow indices.

Goodness of the fitting into the marginal BGPDs are achieved by making use of the λ – Kolmogorov test. The results are shown in Table 2. Hypotheses of the goodness of fit are not rejected at any of the investigated cases.

6.3 Comparison to Zelenhasić and Salvai (1987) model (ZS model)

For the determination of the probabilities of maximum low flow indices by Zelenhasić and Salvai model the formulae

$$G(x) = \Pr(E=0) + \sum_{k=1}^{\infty} F^k(x)\Pr(E=k) \tag{16}$$

is applied, where the probabilities are estimated for annual or seasonal data and

- $G(x)$ is a searched distribution function of maximum low flow indices;
- E is an estimated number of low events in the season;
- $F(x)$ is an estimated distribution function of the low flow indices.

The Figs. 12–15 show the differences between the results. A tendency to overestimate quantiles in ZS model (NIZOWKA program – Jakubowski and Radczuk, 2004) is observed, especially for high probabilities of non-exceedance. Only in the case of the Cigacice profile the probability of the observed maximum indices (low flow deficit amount) is nearly equal in the both methods of estimation. All other indices show lower probabilities of the non-exceedance. It means that ZS model permits too high low flow threat than it is in reality.

7 Conclusions

1. The proposed Bivariate Generalized Pareto Distribution fits very well in the observed shifted extremes of the low flow indices (deficit and duration). The only problem of the estimation are the lower than −0.5 values of $\hat{\kappa}_d$ estimator. There are some profiles where $\hat{\kappa}_d$ is always lower then −1. This tendency is not observable for the $\hat{\kappa}_t$ estimator.

2. Because of the high values of shift parameters (in all researched cases) the observed low flows show distinct statistical non-homogeneity. It means that the different hydrological processes results in the arising of the huge and minor low flows.

3. The estimated correlation between two low flow indices (deficit amount and duration) is very high – in each case it oversteps 0.9. This is why the distribution estimation of the extremes of the low flow indices should not be considered alone one from another.

4. Some of profiles show substantial seasonal non-homogeneity.

Appendix A

Expected values of BGPD

1. First moment:

$$ED_M = \int_0^\infty \int_0^\infty x h(x,t)\,dx\,dt = \frac{1-\alpha_d\kappa_d}{\alpha_d\kappa_d}\left(\frac{\bar{F}_d(0)+\bar{F}_t(0)}{V}-1\right)$$

$$+\frac{k(2-p)}{2Vp\alpha_d\kappa_d(1+\kappa_d)}\left\{\bar{F}_d^{1+\kappa_d}(0)I\left(\frac{\bar{F}_t^{\frac{p}{2}}(0)}{\bar{F}_d^{\frac{p}{2}}(0)},\frac{1}{p}-1,0\right)+\bar{F}_t^{1+\kappa_d}(0)I\left(\frac{\bar{F}_d^{\frac{p}{2}}(0)}{\bar{F}_t^{\frac{p}{2}}(0)},\frac{1}{p}-1,\frac{2\kappa_d}{p}\right)\right\}$$

$$+\frac{(4-k^2)(1-p)}{2Vp\alpha_d\kappa_d(1+\kappa_d)}\left\{\bar{F}_d^{1+\kappa_d}(0)I\left(\frac{\bar{F}_t^{\frac{p}{2}}(0)}{\bar{F}_d^{\frac{p}{2}}(0)},\frac{1}{p}-2,1\right)+\bar{F}_t^{1+\kappa_d}(0)I\left(\frac{\bar{F}_d^{\frac{p}{2}}(0)}{\bar{F}_t^{\frac{p}{2}}(0)},\frac{1}{p}-2,\frac{2\kappa_d}{p}+1\right)\right\}.$$

   The first moment is finite when $\kappa_d \in (-1;0)$.

2. Second moment

$$ED_M^2 = \int_0^\infty \int_0^\infty x^2 h(x,t)\,dx\,dt = \left(\frac{1-\alpha_d\kappa_d}{\alpha_d\kappa_d}\right)^2\left(\frac{\bar{F}_d(0)+\bar{F}_t(0)}{V}-1\right)$$

$$+\frac{(1-\alpha_d\kappa_d)k(2-p)}{Vp\alpha_d^2\kappa_d^2(1+\kappa_d)}\left\{\bar{F}_d^{1+\kappa_d}(0)I\left(\frac{\bar{F}_t^{\frac{p}{2}}(0)}{\bar{F}_d^{\frac{p}{2}}(0)},\frac{1}{p}-1,0\right)+\bar{F}_t^{1+\kappa_d}(0)I\left(\frac{\bar{F}_d^{\frac{p}{2}}(0)}{\bar{F}_t^{\frac{p}{2}}(0)},\frac{1}{p}-1,\frac{2\kappa_d}{p}\right)\right\}$$

HESSD

3, 859–893, 2006

$$
+\frac{(1-\alpha_d\kappa_d)(4-k^2)(1-p)}{Vp\alpha_d^2\kappa_d^2(1+\kappa_d)}\left\{\bar{F}_d^{1+\kappa_d}(0)I\left(\frac{\bar{F}_t^{\frac{p}{2}}(0)}{\bar{F}_d^{\frac{p}{2}}(0)},\frac{1}{p}-2,1\right)\right.
$$

$$
\left.+\bar{F}_t^{1+\kappa_d}(0)I\left(\frac{\bar{F}_d^{\frac{p}{2}}(0)}{\bar{F}_t^{\frac{p}{2}}(0)},\frac{1}{p}-2,\frac{2\kappa_d}{p}+1\right)\right\}
$$

$$
-\frac{k(2-p)}{2Vp\alpha_d^2\kappa_d^2(1+2\kappa_d)}\left\{\bar{F}_d^{1+2\kappa_d}(0)I\left(\frac{\bar{F}_t^{\frac{p}{2}}(0)}{\bar{F}_d^{\frac{p}{2}}(0)},\frac{1}{p}-1,0\right)\right.
$$

$$
\left.+\bar{F}_t^{1+2\kappa_d}(0)I\left(\frac{\bar{F}_d^{\frac{p}{2}}(0)}{\bar{F}_t^{\frac{p}{2}}(0)},\frac{1}{p}-1,\frac{4\kappa_d}{p}\right)\right\}
$$

$$
-\frac{(4-k^2)(1-p)}{2Vp\alpha_d^2\kappa_d^2(1+2\kappa_d)}\left\{\bar{F}_d^{1+2\kappa_d}(0)I\left(\frac{\bar{F}_t^{\frac{p}{2}}(0)}{\bar{F}_d^{\frac{p}{2}}(0)},\frac{1}{p}-2,1\right)\right.
$$

$$
\left.+\bar{F}_t^{1+2\kappa_d}(0)I\left(\frac{\bar{F}_d^{\frac{p}{2}}(0)}{\bar{F}_t^{\frac{p}{2}}(0)},\frac{1}{p}-2,\frac{4\kappa_d}{p}+1\right)\right\}.
$$

The second moment is finite when $\kappa_d \in \left(-\frac{1}{2};0\right)$.

Moments ET_M and ET_M^2 can be calculated in a similar way.

HESSD
3, 859–893, 2006
Bivariate distribution of the low flow extremes estimation
W. Jakubowski

3. Mixed moment

$$ED_M T_M = \int_0^\infty \int_0^\infty xth(x,t)dxdt = \frac{1-\alpha_t\kappa_t}{\alpha_t\kappa_t} ED_M + \frac{1-\alpha_d\kappa_d}{\alpha_d\kappa_d} ET_M$$
$$- \frac{1-\alpha_t\kappa_t}{\alpha_t\kappa_t}\frac{1-\alpha_d\kappa_d}{\alpha_d\kappa_d}\left(\frac{\bar{F}_d(0)+\bar{F}_t(0)}{V}-1\right)$$
$$- \frac{k(2-p)}{2Vp\alpha_d\kappa_d\alpha_t\kappa_t(1+\kappa_d+\kappa_t)}\left\{\bar{F}_d^{1+\kappa_d+\kappa_t}(0) I\left(\frac{\bar{F}_t^{\frac{p}{2}}(0)}{\bar{F}_d^{\frac{p}{2}}(0)}, \frac{1}{p}-1, \frac{2\kappa_t}{p}\right)\right.$$
$$\left. + \bar{F}_t^{1+\kappa_d+\kappa_t}(0) I\left(\frac{\bar{F}_d^{\frac{p}{2}}(0)}{\bar{F}_t^{\frac{p}{2}}(0)}, \frac{1}{p}-1, \frac{2\kappa_d}{p}\right)\right\}$$
$$- \frac{(4-k^2)(1-p)}{2Vp\alpha_d\kappa_d\alpha_t\kappa_t(1+\kappa_d+\kappa_t)}\left\{\bar{F}_d^{1+\kappa_d+\kappa_t}(0) I\left(\frac{\bar{F}_t^{\frac{p}{2}}(0)}{\bar{F}_d^{\frac{p}{2}}(0)}, \frac{1}{p}-2, \frac{2\kappa_t}{p}+1\right)\right.$$
$$\left. + \bar{F}_t^{1+\kappa_d+\kappa_t}(0) I\left(\frac{\bar{F}_d^{\frac{p}{2}}(0)}{\bar{F}_t^{\frac{p}{2}}(0)}, \frac{1}{p}-2, \frac{2\kappa_d}{p}+1\right)\right\}.$$

The mixed moment is finite when $\kappa_d+\kappa_t \in (-1;0)$.

The parameter V is defined by Eq. (14) in Sect. 5.2 and the integral: $I(\alpha,\beta,\gamma)=\int_0^\alpha \left(1+kt+t^2\right)^\beta t^\gamma dt$. Since $\gamma>-1$, the integral I is always finite.

HESSD
3, 859–893, 2006

EGU

References

Coles, S.: An Introduction to Statistical Modeling of Extreme Values, Springer-Verlag London Limited, 2001.

Engeland, K., Hisdal, H., and Frigessi, A.: Practical Extreme Value Modelling of Hydrological Floods and Droughts: A Case Study, Extremes, 7, 5–30, 2004.

Hisdal, H., Tallaksen, L. M., and Frigessi, A.: Handling non-extreme events in extreme value modelling of streamflow droughts, IAHS Publ. No 274, 281–288, 2002.

Jakubowski, W. and Radczuk, L.: Estimation of Hydrological Drought Characteristics NIZOWKA2003, software, on accompanying CD to Hydrological Drought, Processes and Estimation Methods for Streamflow and Groundwater, edited by: Tallaksen, L. M. and van Lanen, H. A. J, 2004, Elsevier, Amsterdam, 2004.

Jakubowski, W.: Zastosowanie Uogólnionego Rozkladu Pareto do wyznaczania rozkladów maksymalnych charakterystyk niżówek, ZNAR we Wroclawiu (in Polish), 520, 29–41, 2005.

Pickands, J.: Statistical inference using extreme order statistics, Ann. Statist., 3, 119–131, 1975.

Resnick, S. I.: Extreme values, Regular Variation and Point Processes, Berlin Springer-Verlag, 1987.

Tajvidi, N.: Characterisation and Some Statistical Aspects of Univariate and Multivariate Generalized Pareto Distribution, PhD Thesis, available at: http://www.maths.lth.se/matstat/staff/nader/fullpub.html, 1996.

Tallaksen, L. M. and van Lanen, H. A. J. (Eds.): Hydrological Drought, Processes and Estimation Methods for Streamflow and Groundwater, Elsevier, Amsterdam, 2004.

Zelenhasić, E. and Salvai, A.: A Method of Streamflow Drought Analysis, Water Resour. Res., 23(1), 156–168, 1987.

Table 1. The best shift parameters estimated by the correlation coefficient maximization.

Profile	Years of the observed runoffs	Observed maximum of		Shift parameter for		Correlation coefficient
		deficit in millions m^3	duration in days	b_d-deficit in millions m^3	b_t-duration in days	
Cigacice	1966–2003	1371.686	270	74.650	39	0.9312
Colwick	1959–2000	266.722	180	10.353	17	0.9063
Kuripapango	1965–2000	39.917	113	3.859	23	0.9386
Międzylesie	1966–2003	2.471	178	0.134	24	0.9315

Table 2. The λ-Kolmogorov test; goodness of fit into marginal BGPD.

Profile	Number of the low flow events over shift parameters		λ-Kolmogorov test; values for low flow	
	deficit amounts	durations	deficit amounts	durations
Cigacice	30	29	0.477	0.400
Colwick	71	70	0.661	0.504
Kuripapango	54	52	0.365	0.261
Międzylesie	35	34	0.374	0.419

HESSD
3, 859–893, 2006

Bivariate distribution of the low flow extremes estimation

W. Jakubowski

EGU

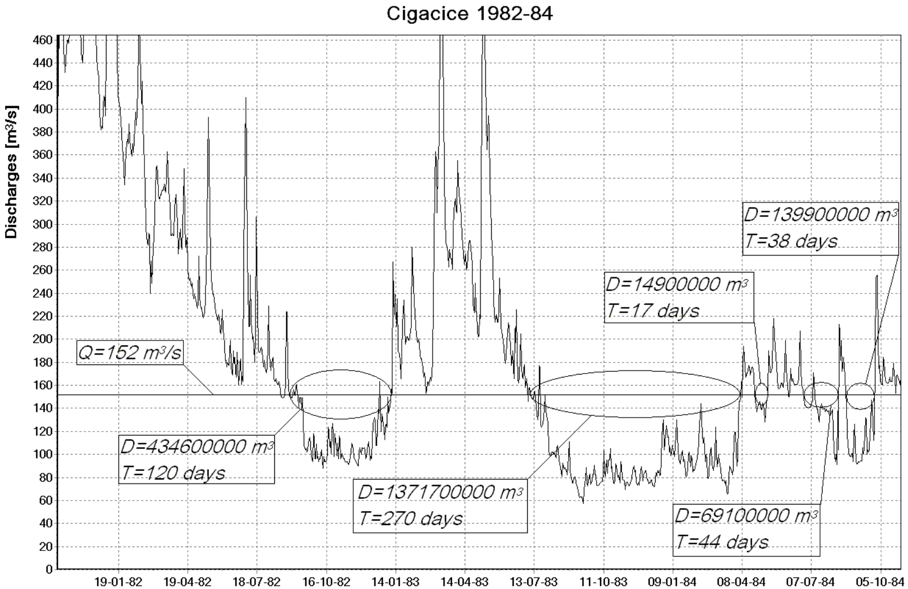

Fig. 1. The Odra River, Cigacice profile; an example of the low flow indices.

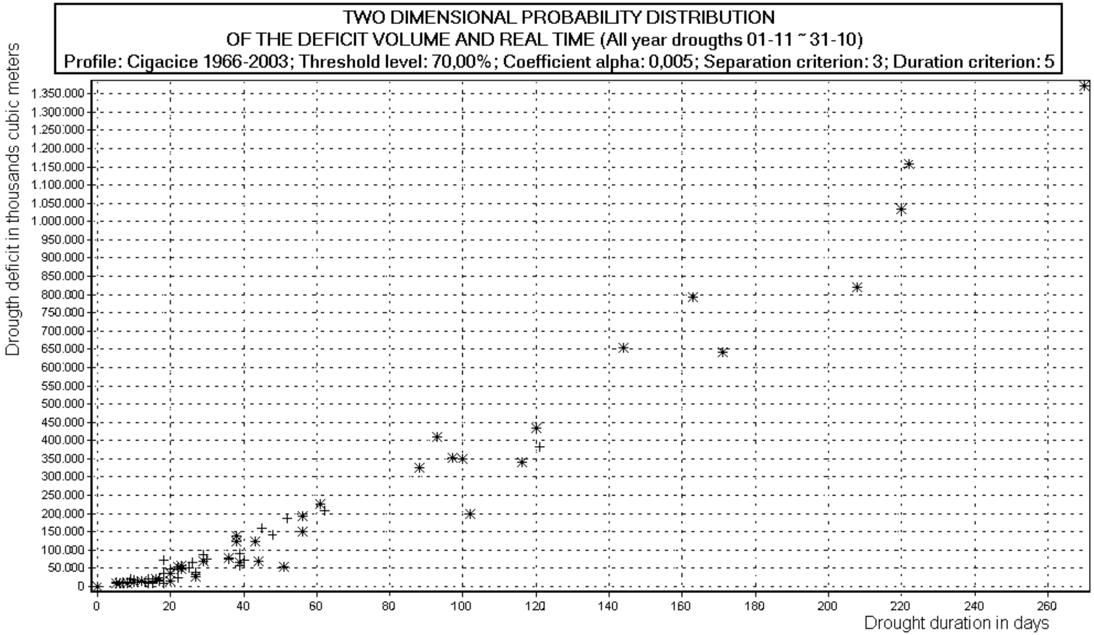

Fig. 2. The Odra River, Cigacice profile; an observed low flow deficits and durations.

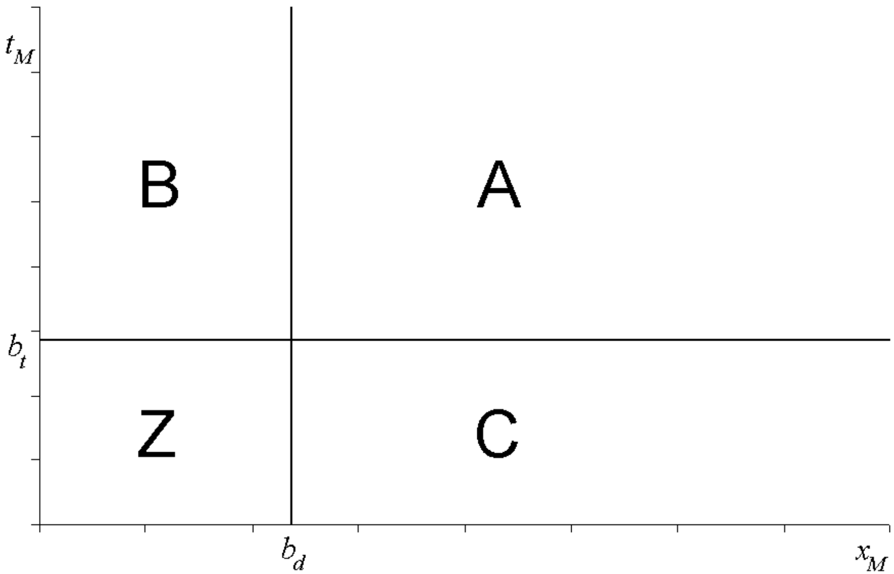

Fig. 3. Support of the Bivariate Generalized Pareto Distribution (Tajvidi, 1996).

HESSD
3, 859–893, 2006

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

EGU

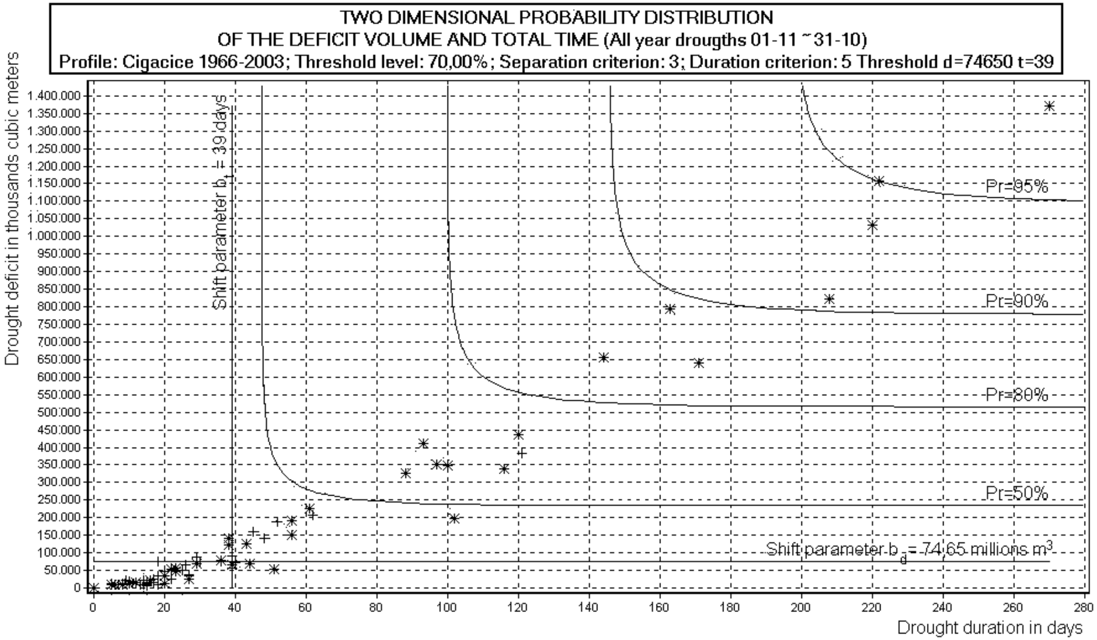

Fig. 4. The Odra River, Cigacice profile; an annual low flows – fitting into the Bivariate Generalized Pareto Distribution.

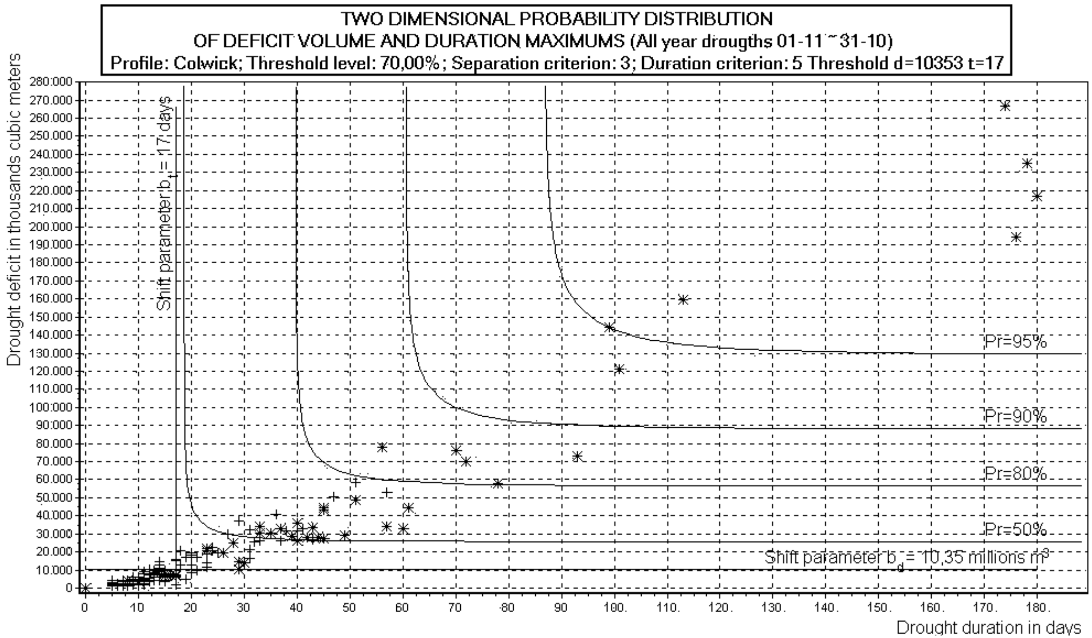

Fig. 5. The Trent River, Colwick profile; an annual low flows – fitting into the Bivariate Generalized Pareto Distribution.

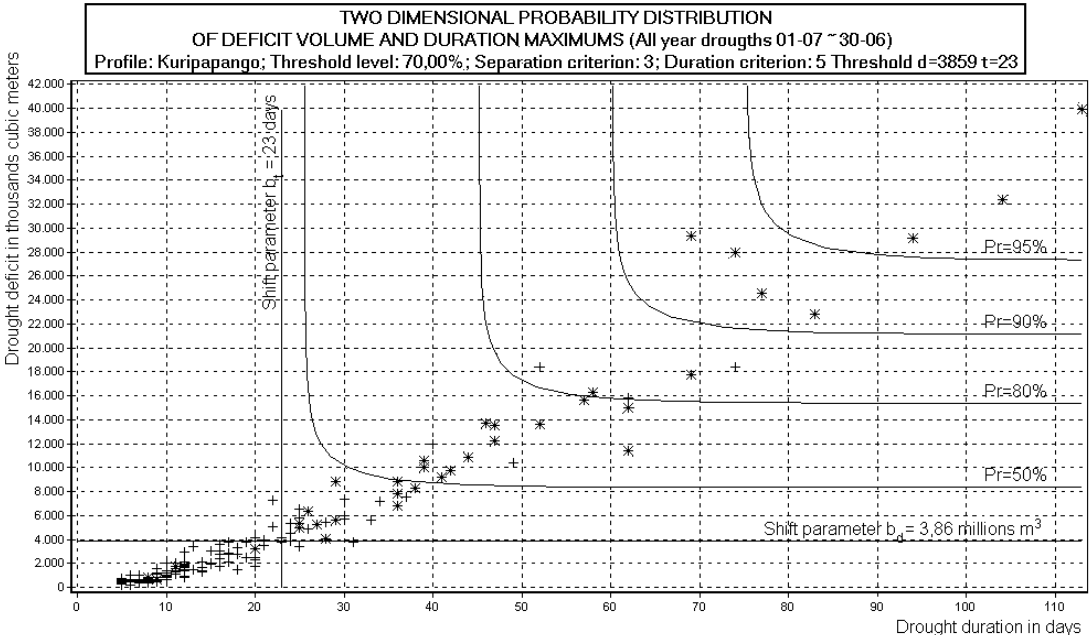

Fig. 6. The Ngaruroro River, Kuripapango profile; an annual low flows – fitting into the Bivariate Generalized Pareto Distribution.

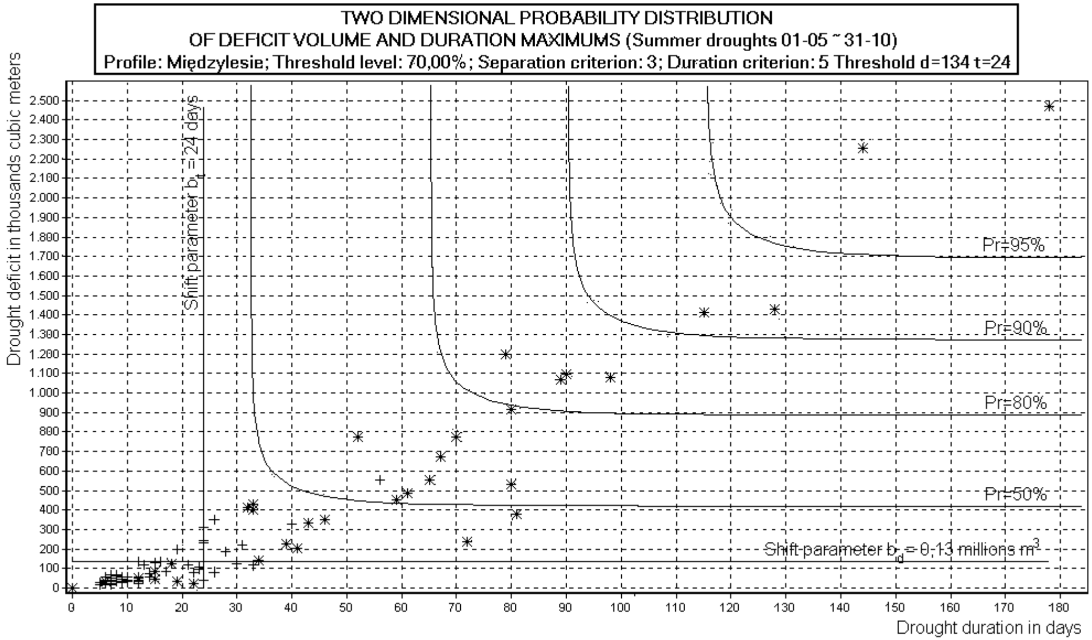

Fig. 7. The Nysa Klodzka River, Międzylesie profile; a summer low flows – fitting into the Bivariate Generalized Pareto Distribution.

HESSD
3, 859–893, 2006

Bivariate distribution of the low flow extremes estimation

W. Jakubowski

EGU

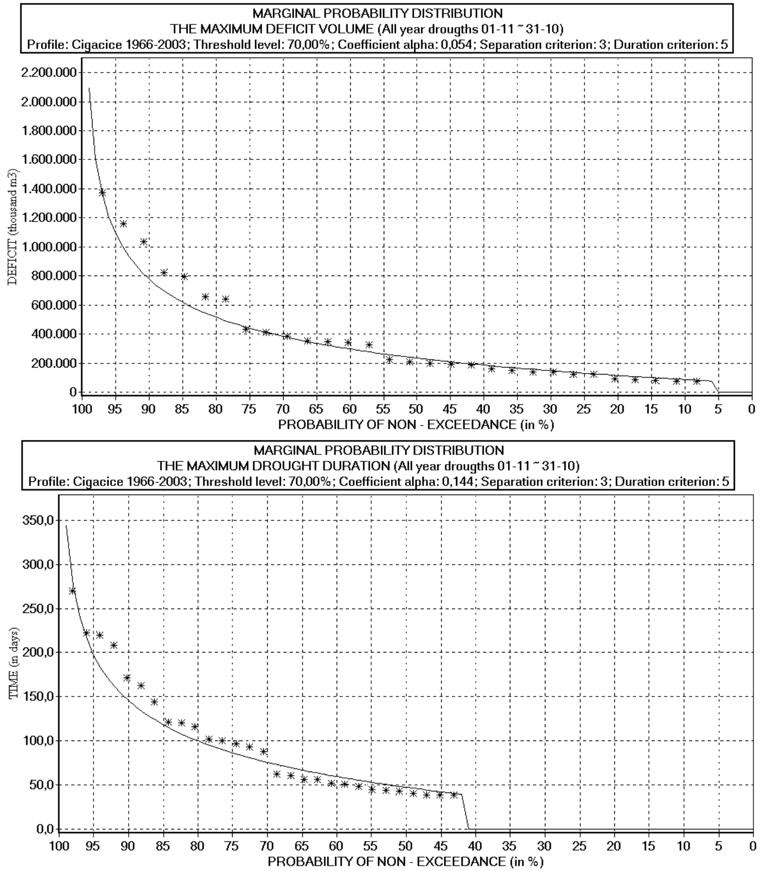

Fig. 8. The Odra River, Cigacice profile; an annual low flows – fitting into marginal distributions of the low flow maximum deficit volume and duration.

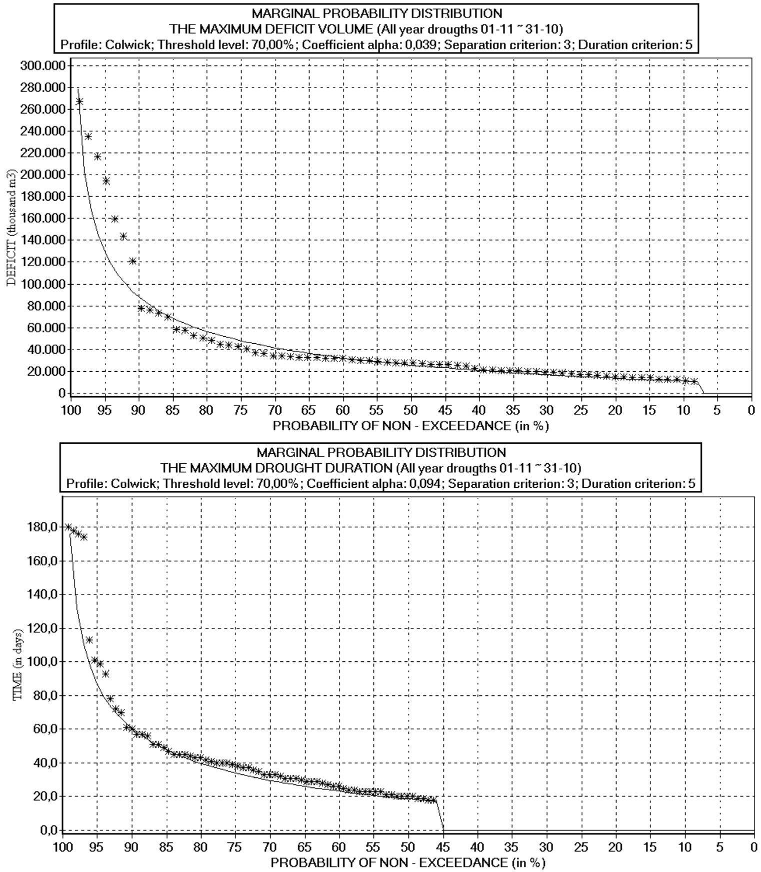

Fig. 9. The Trent River, Colwick profile; an annual low flows – fitting into marginal distributions of the low flow maximum deficit volume and duration.

HESSD
3, 859–893, 2006

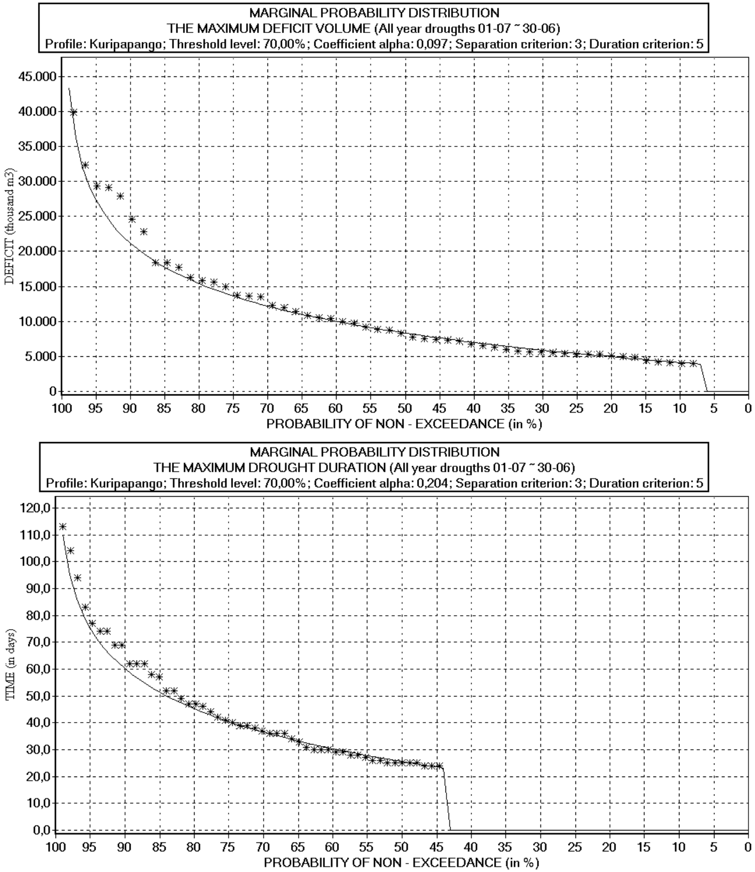

Fig. 10. The Ngaruroro River, Kuripapango profile; an annual low flows – fitting into marginal distributions of the low flow maximum deficit volume and duration.

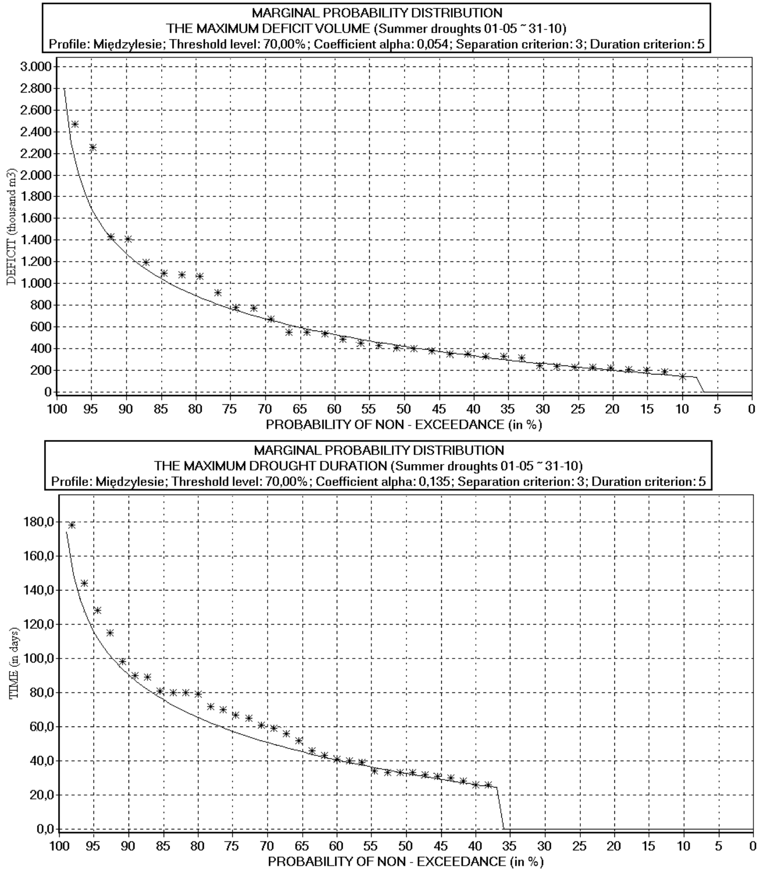

Fig. 11. The Nysa Klodzka River, Międzylesie profile; a summer low flows – fitting into marginal distributions of the low flow maximum deficit volume and duration.

HESSD
3, 859–893, 2006

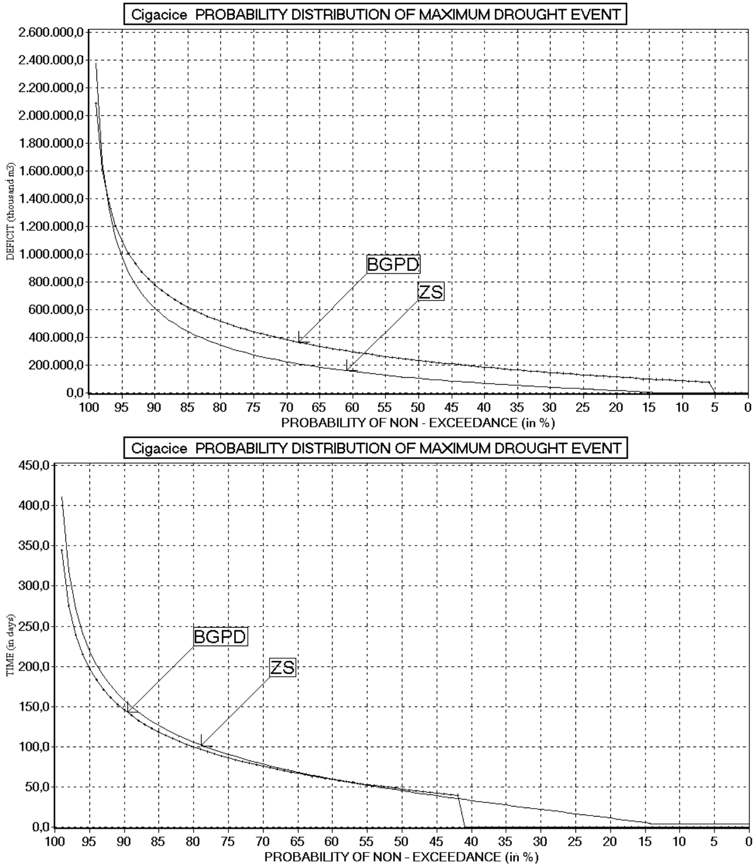

Fig. 12. The Odra River, Cigacice profile; an annual low flows – comparison between the low flow deficit and duration quantiles obtained by fitting into the BGPD and ZS model.

HESSD
3, 859–893, 2006

Bivariate distribution of the low flow extremes estimation

W. Jakubowski

EGU

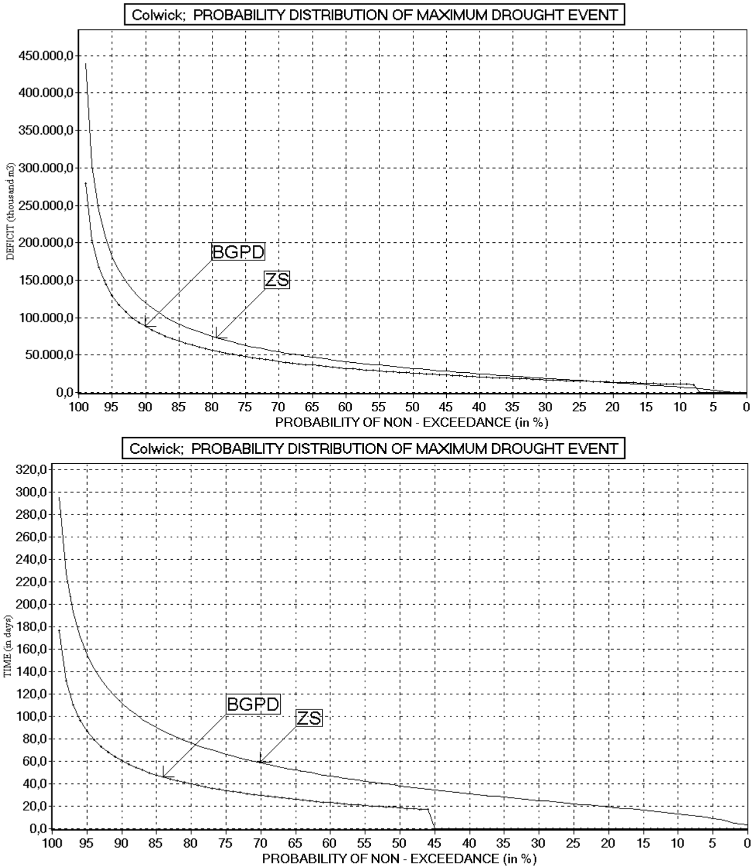

Fig. 13. The Trent River, Colwick profile; an annual low flows – comparison between the low flow deficit and duration quantiles obtained by fitting into the BGPD and ZS models.

HESSD

3, 859–893, 2006

Bivariate distribution of the low flow extremes estimation

W. Jakubowski

EGU

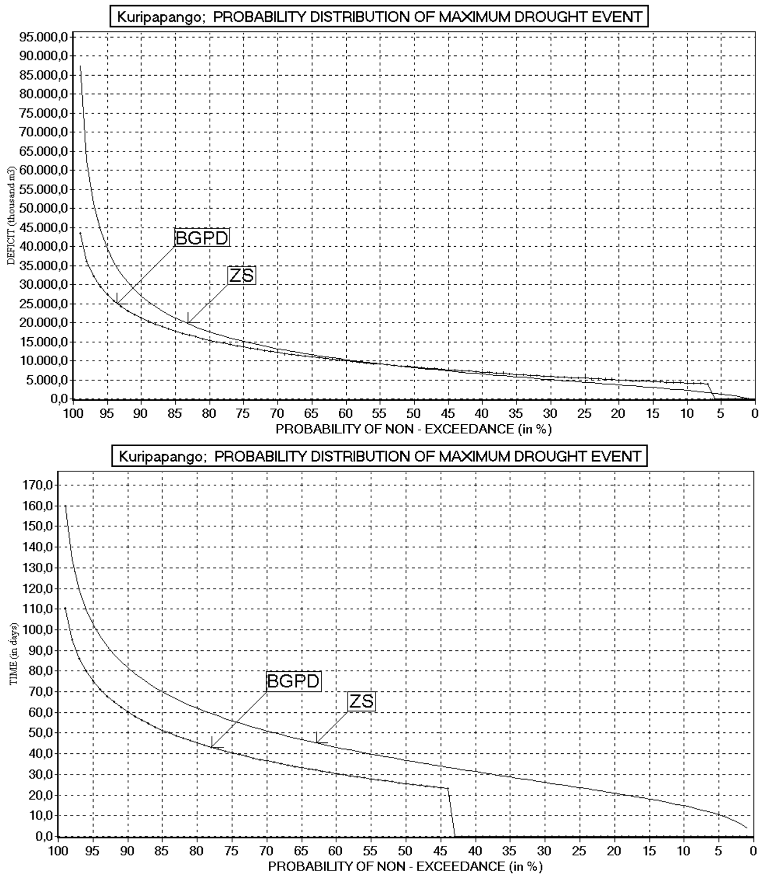

Fig. 14. The Ngaruroro River, Kuripapango profile; an annual low flows – comparison between the low flow deficit and duration quantiles obtained by fitting into the BGPD and ZS models.

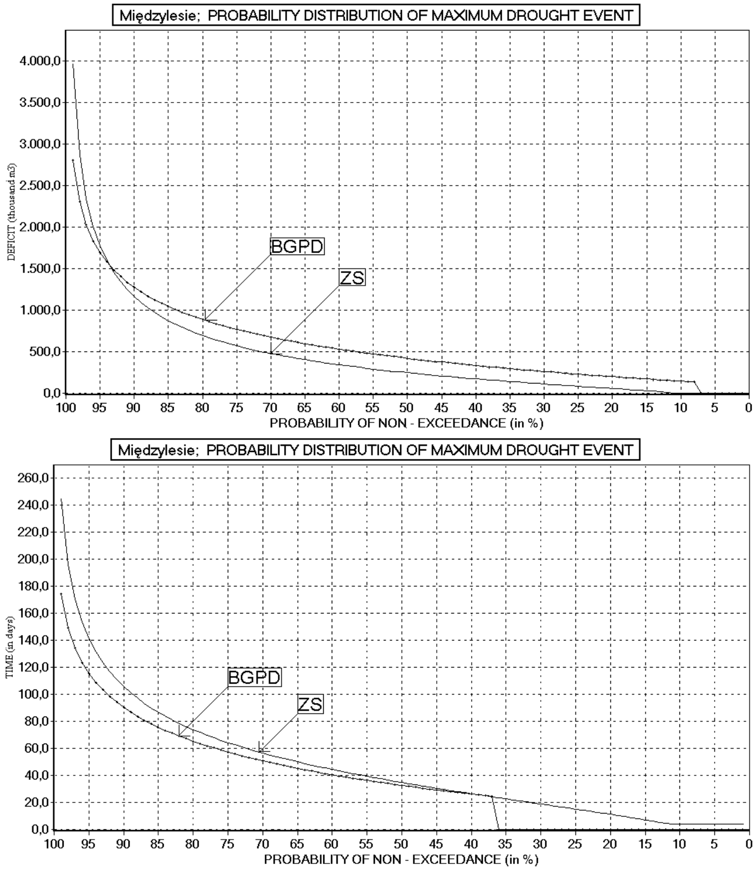

Fig. 15. The Nysa Klodzka River, Międzylesie profile; a summer low flows – comparison between the low flow deficit and duration quantiles obtained by fitting into the BGPD and ZS models.