

Probable annual floods in a small lowland river estimated with the use of various sets of data

KAZIMIERZ BANASIK, ANDRZEJ BYCZKOWSKI

Department of Water Engineering and Environmental Restoration
Warsaw University of Life Sciences – SGGW,

Abstract: *Probable annual floods in a small lowland river estimated with the use of various sets of data.* The results of estimation of the probable annual flood flows with the use of various sets of data from a small agricultural lowland river are presented. The traditional statistical series are formed from the annual maximum (AM) flows of hydrological years. After examination of the homogeneity of the series, two sets of data were formed (AM1 of 34 elements and AM2 of 40 elements), which differed in the significance level of one of the stationary tests. The other series for flood frequency analysis has been formed by selecting peaks over threshold discharges (POT). A computer program, developed by IMGW (Institute of Hydrology and Water Management), and spreadsheet were applied for frequency analysis with the use of AM series and POT series, respectively. Results of computations with the use of AM1 and AM2 data indicate for nonsignificant differences in probable flood flows, and significant with the probable flood flow estimated with the POT data.

Key words: floods, small catchment, design floods.

INTRODUCTION

The first step in designing culverts, bridges or hydraulic structures and in planning floodplain management is to determine the design floods, i.e. flood flows having given recurrences intervals or probabilities of exceedance. This is equally important for larger and smaller

ivers and needs to be determined for gauged and ungauged locations (Banasik et al., 2003; Byczkowski, 1999; WMO 1994). The problem of estimating flood flows has also an increasing importance in various environmental studies (Cunderlink et al., 2006; Cyberski et al., 2006; Pinter et al., 2006). The presented paper deals with methods used for estimation of flood discharges with the probability of assumed exceedances in small gauged river. Traditionally, long term annual maximal flows (AM) are used for such purposes (Banasik, Byczkowski, 2006; Byczkowski, 1999). Accuracy of estimation of such approach depends on sample size (i.e. number of years of flow record). This is specially important for $p \leq 1\%$ (i.e. for probability of exceedance not larger than 1%). In practical application, sample size is not often large enough to estimate flood flows $Q_p > Q_{1\%}$. It is to be pointed out that the sample size could be increased by taking into account not only the maximal annual discharges, but also other flow peaks – smaller than annual maximum (Strupczewski, 1967; Beran, Nozdryn-Plotnicki, 1977). All flow peaks which are equal to or higher than a threshold discharge are taken into account ($Q_{\text{peak}} \geq Q_0$), so the procedure is called the peaks over threshold (POT) method. It is not allowed to include all

flow peaks into the sample as some of the events may not keep the principle of independence. Usually as threshold discharge Q_0 , it is assumed the smallest value of the maximal annual discharges $\{\min Q_{\max}\}$. However, some researchers, i.e. Aleksejew, suggest assuming the threshold discharge as $Q_0 = 2 \cdot SSQ$ (Strupczewski 1967a, b).

– SGGW, described in details in previous publications (Banasik et al., 2003; Banasik, Byczkowski, 2006). The annual maxima (AM) values of the 40-year period are shown in Figure 1. Figure 2 illustrates the peaks over a threshold (POT) procedure, which leads to significant increase in the sample size in comparison to AM method.

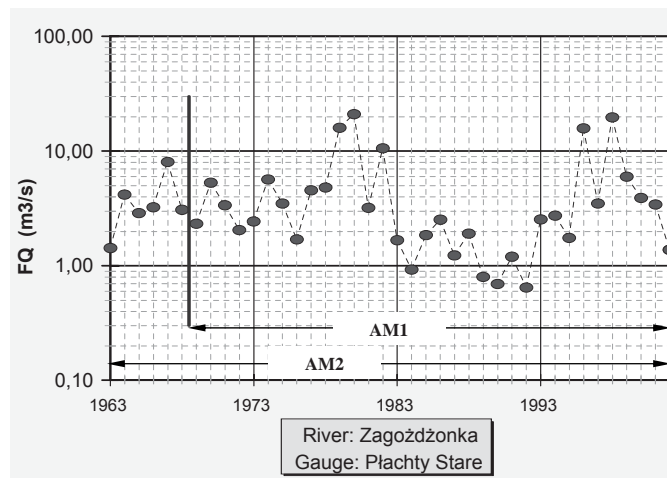


FIGURE 1. Annual peak floods of Zagożdżonka River at Plachty Stare gauge for the hydrological years of 1963–2002 (displaying two analyzed set of data AM1 and AM2)

The aim of this paper is to compare the annual flood flows of assumed probability of exceedance estimated with the use of AM (traditional) method with the values estimated with the use of POT method. Investigation has been based on the flow data of a period of 40 years (1963–2002) recorded in a small agricultural river basin of Zagożdżonka ($A = 82 \text{ km}^2$), located in the center of Poland (ca 100 km south of Warsaw). The Zagożdżonka river basin is research area of Department of Water Engineering and Environmental Restoration (earlier: Department of Hydraulic Structures) of Warsaw University of Life Sciences

DESCRIPTION OF THE SETS OF DATA USED IN THE ANALYSIS

The first set of data has been formed from annual maximum (AM) floods, without taking into account their genesis. The process of performing the frequency analysis based on the AM floods, with the use of IMGW (Institute of Meteorology and Water Management) computer program described by Ozga-Zielińska et al. (1999), is being conducted after prior examination of the homogeneity of the series with the use of statistical methods. The methods included:

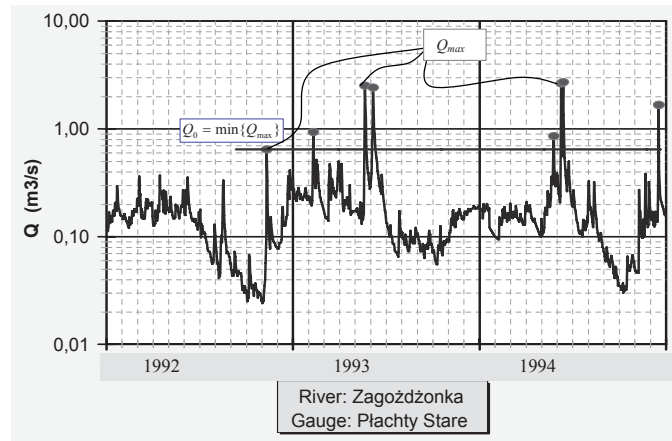


FIGURE 2. Example of selection of the flood flows for peaks over threshold (POT) series analysis

- a) investigation of the outliers with the use of Grubbs-Beck test,
- b) investigation of the independence of elements of the sample with the use of the run test,
- c) investigation of the stationarity of the series with the use of the three following tests:
 - the Kruskal-Wallis one-way analysis of variance by ranks,
 - the Spearman test for trend of mean value,
 - the Spearman test for trend of variance.

The examination of the homogeneity of the 40-year AM series (1963–2002) showed that the only critical test was that one of the Kruskal-Wallis, indicating significant change of mean value (nonstationarity) at the 0.05 level. However, after eliminating the oldest values from the series, it has been found that 34-year series (1969–2002), called AM1 set of data, fulfilled all of the conditions of homogeneity. As the 40-year AM series fulfilled the Kruskal-Wallis condition at 0.01 significance level, so it has been also considered as other set of data, called AM2, for

frequency analysis. The respective data for both sets (AM1 and AM2) are shown in Figure 1. The problem of accepting the series to, or rejecting it from the frequency analysis, depending on the fulfilling the respective homogeneity criteria, has been recently discussed by Węglarczyk (2007) and Ozga-Zielińska et al. (2007).

Having applied the Akaike criterion (Mutua, 1994), a statistical distribution has been chosen from the following: the Pearson type III (i.e. 3-parameter gamma), the log-normal, the Weibull, and log-Pearson type III distributions. With the use of computer program QMAX, developed by Ozga-Zielińska et al. (1999), the log-normal distribution has been selected for the annual values sample as producing the lowest value of the Akaike criterion for both the AM1 and AM2 set of data. For estimation of the distribution parameters the method of maximum likelihood has been used. The probable annual maxima flood flows have been presented in Figures 3 and 4 for the AM1 and AM2 set of records respectively, and in Table 1.

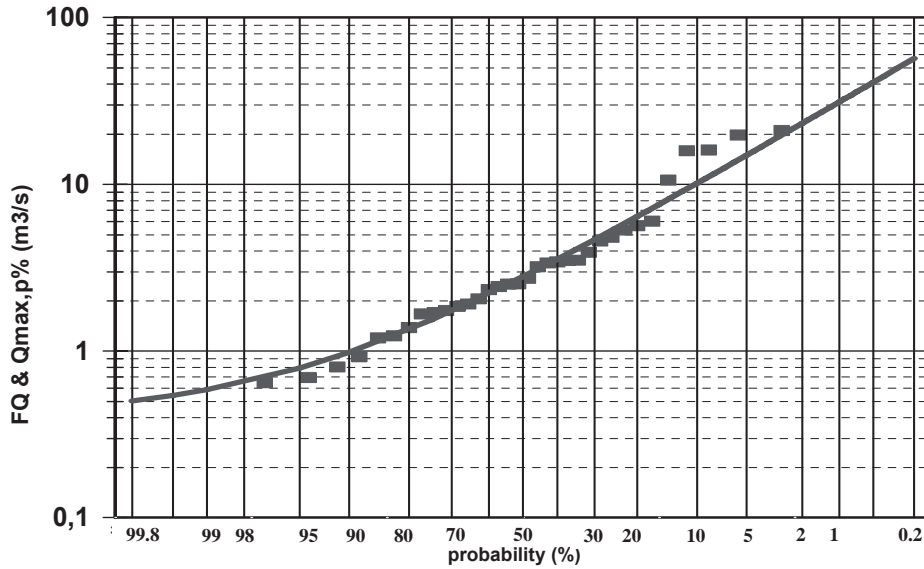


FIGURE 3. The probable annual maxima of flood flows for the Zagożdżonka river at Płachty Stare with the use of 34-year period (AM1 set of data, i.e. 1969–2002)

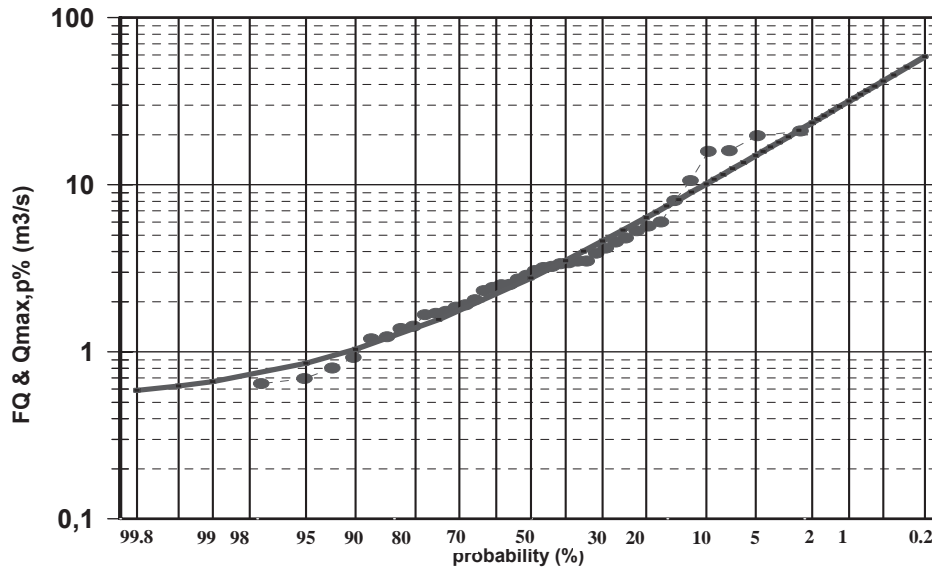


FIGURE 4. The probable annual maxima of flood flows for the Zagożdżonka river at Płachty Stare with the use of 40-year period (AM2 set of data, i.e. 1962–2002)

TABLE 1. Probable flood flows estimated with the use of log-normal distribution and with various sets of records for the Zagożdżonka river at Plachty Stare

Return period T (year)	Probability p (%) (1-CDF)	Flood flows (m ³ /s) estimated on the base of		
		annual maxima for period of		annual exceedances (POT)
		34 years (AM1)	40 years (AM2)	i.e. 261 events from the period of 40 years
1	2	3	4	5
1000	0.1	71.4	74.3	48.8
100	1.0	31.1	31.7	22.2
20	5.0	14.9	15.0	11.7
10	10	10.2	10.1	8.52
2	50	2.79	2.77	3.34
1.11	90	0.98	1.04	1.71
1.01	99	0.59	0.57	1.19

The other set of data has been formed by accepting all peak flows higher than a threshold discharge ($Q_{\text{peak}} \geq Q_0$), from the 40-year period (POT) record. In the sample formed in this way, the probability of exceedance of each peak flow (p') is estimated. Those values are not equivalent to annual probability (p). If assumed that peak flow above accepted threshold discharge appears at least once in each year (i.e. $Q_{\text{max}} \geq Q_0$ when $Q_0 = \min Q_{\text{max}}$), the annual exceedance probability could be computed according to Langbein formula (Strupczewski, 1967a):

$$p = 1 - (1 - p')^\lambda \tag{1}$$

where:

p – annual probability of exceedance,
 p' – probability of exceedance $Q_{\text{peak}} \geq Q_0$,
 λ – mean annual frequency of occurrence of peak flows, computed from equation:

$$\lambda = \frac{n}{N} \tag{2}$$

where:

n – sample size (number of peak flows),
 N – number of years.

Based on the POT set of data and with the use of the formulas (1) and (2) the annual maximal flood flows can be computed. The lowest value of the maximal annual flood was: $\min\{Q_{\text{max}}\} = 0.646 \text{ m}^3/\text{s}$. The set of peak flows has 261 elements. It has been checked for stationarity by the Kruskal-Wallis test, and passed the test at the 0.05 level. The mean annual frequency of occurrence of peak flows within the 40 year period of record was $\lambda = 261/40 = 6.5$.

On the basis of the set of POT values, the probability in the sample p' and annual probability p (according to equation 1) has been computed. The lower boundary (ϵ) of the assumed log-normal distribution has been found on the way of maximizing the likelihood function as $\epsilon = 0.638 \text{ m}^3/\text{s}$, from the formula:

$$L = f(x_1; \varepsilon; \mu, \sigma) \cdot f(x_2; \varepsilon; \mu, \sigma) \dots \cdot f(x_n; \varepsilon; \mu, \sigma) \quad (3)$$

where:

L – likelihood function,
 f(·) – probability density function of the log-normal distribution,
 x_1, x_2, \dots, x_n – sample of n POT exceedances (i.e. 261 peak flows),
 ε, μ, σ – parameters of the distribution.

The annual exceedances of flood flows have been presented in Figure 5 and in Table 1.

homogeneity indicated the values of 2.79 m³/s, 10.2 m³/s, and 31.1 m³/s as floods of probability of exceedances as 50, 10 and 1%, respectively. Nonsignificant differences have been noticed when the 40 year (AM2) set of data, passing the Kruskal-Wallis test at 0.01 significance level, was used.

The further analysis of the differences of the probable flood flows computed from the AM2 sets of data i.e. annual

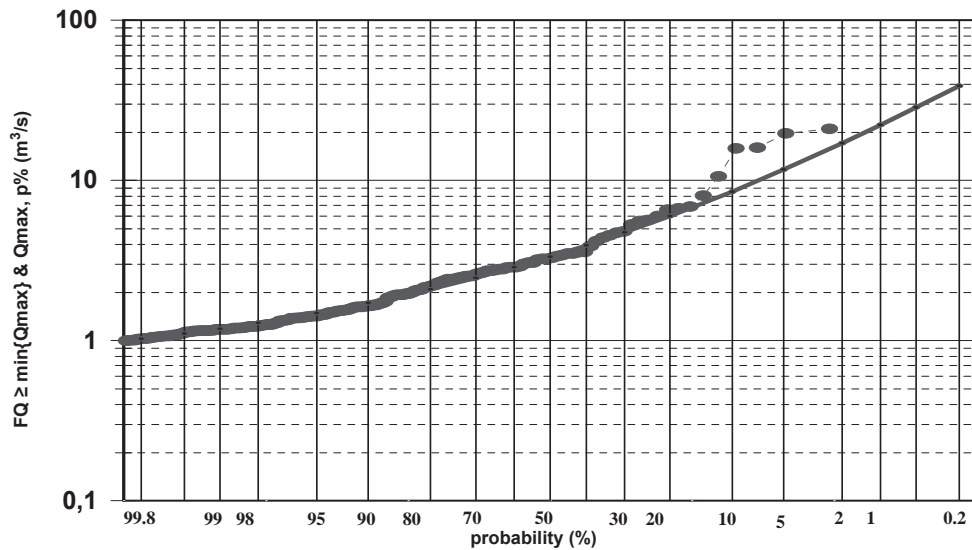


FIGURE 5. The probable annual exceedances of flood flows for the Zagożdżonka river at Płachty Stare (POT set of data)

COMPUTATION RESULTS AND CONCLUDING REMARKS

Flood flow data recorded in a small agricultural lowland river of Zagożdżonka at the gauge Płachty Stare in the period 1963–2002 has been analyzed to estimate design floods.

The annual maximum set of data of the period 1969–2002 (AM1), which passed all the statistical tests for

maxima floods and peak flows over threshold (POT), both from the period of 40-years, are the basis to formulate the following conclusions:

- for the probability of exceedances of $p < 10\%$ the flood flows estimated with the use of annual maxima events are larger than computed with the POT events,
- for the probability of exceedances of $p > 50\%$ the results are opposite.

The POT method requires using much larger sample of events than the annual maxima methods, which give a base for more precise estimation of skewness of the distribution, which depends on the size of the set data. Taking this into account one can assume that values estimated with the POT method could be more reliable.

ACKNOWLEDGMENT

The study described in this paper has been carried out within research project 1010/P01/2006/30, founded by Ministry of Science and Higher Education. The financial support provided by this organization is gratefully acknowledged.

REFERENCES

- BANASIK K., BYCZKOWSKI A., GŁADECKI J. 2003: Prediction of T-year flood discharge for a small river basin using direct and indirect methods. *Annals of Warsaw Agricultural University, Land Reclamation*. No 34, p. 3–8.
- BANASIK K., BYCZKOWSKI A. 2006: Estimation of T-year flood discharge for a small lowland river using statistical method. *Annals of Warsaw Agricultural University, Land Reclamation*. No 37, p. 27–32.
- BERAN M.A., NOZDRYN-PLOTNICKI M.J. 1977: Estimation of low return period floods. *Hydrological Sciences Journal*, Vol. XXII, No 6, p. 275–282.
- BYCZKOWSKI A. 1999: *Hydrologia*, t. 2. Wyd. SGGW, Warszawa.
- CUNDERLINK J.M., OUARDA T.B.M.J. 2006: Regional flood-duration-frequency modeling in the changing environment. *Journal of Hydrology*. Vol. 318, Issues 1–4, p. 276–291.
- CYBERSKI J., GRZEŚ M., GUTRY-KORYCKA M., KUNDZEWICZ Z.W. 2006. History of floods on the River Vistula. *Hydrological Sciences Journal*, Vol. 51, Issue 5, p. 799–817.
- MUTUA F.M. 1994: The use of the Akaike information criterion in the identification of an optimum flood frequency model. *Hydrological Science Journal*, Vol. 39 No 3 p. 235–244.
- OZGA-ZIELIŃSKA M., BRZEZIŃSKI J., OZGA-ZIELIŃSKI B. 1999: Zasady obliczania największych przepływów rocznych o określonym prawdopodobieństwie przewyższenia – przy projektowaniu obiektów budownictwa hydrotechnicznego (Guidelines for flood frequency computation with small probability of exceedances – for design of hydrotechnical structures). *Materiały Badawcze, Seria: Hydrologia i Oceanologia*. IMGW, Warszawa.
- OZGA-ZIELIŃSKA M., BRZEZIŃSKI J., OZGA-ZIELIŃSKI B. 2007: Określenie prawdopodobieństwa przepływów ekstremalnych rocznych genetycznie niejednorodnych – metoda alternatywy zdarzeń (Determining the probability of genetically heterogeneous yearly floods – method of the alternative events). *Gospodarka Wodna* No 5, p. 191–196.
- PINTER N., ICKES B.S., WLOSINSKI J.H., VAN DER PLOEG R.R. 2006: Trends in flood stages: Contrasting results from the Mississippi and Rhine River systems. *Journal of Hydrology*. Vol. 331, Issues 3–4, p. 554–566.
- STRUPCZEWSKI W. 1967a: Determination of the probability of repeating phenomena. *Acta Geophysica Polonica*, Vol. XV, No 2, p. 147–158.
- STRUPCZEWSKI W. 1967b: Determination of the probability annual distribution of some events of all their occurrences. *Acta Geophysica Polonica*, Vol. XV, No 3, p. 247–263.
- WEGLARCZYK S. 2007: Metoda alternatywy zdarzeń identyfikacji rozkładu prawdopodobieństwa przepływów maksymal-

nych w roku - analiza krytyczna (Metod of alternative events for identification of probability distribution of maximum flow values in the year – a critical analysis). Gospodarka Wodna No 5, p. 185–190.

WMO (World Meteorological Organisation) 1994: Guide to Hydrological Practices, WMO-No 168.

Streszczenie: Powodzie prawdopodobne w małej rzece nizinnej określone przy wykorzystaniu różnych zbiorów danych. Na podstawie 40-letniego ciągu przepływów dobowych zarejestrowanych w profilu Płachty Stare na Zagożdżoncu ($A = 82,4 \text{ km}^2$) utworzono zbiory danych do analizy częstotliwości przepływów powodziowych. Po statystycznej analizie jednorodności rocznych przepływów maksymalnych utworzono dwa zbiory AM1 i AM2 (annual maxima), odpowiednio 34- i 40-elementowe, różniące się poziomem zgodności jednego z testów stacjonarności ciągów. Trzeci zbiór danych utworzono ze wszystkich wezbrań okresu 40-letniego, nie mniej-

szych niż NWQ. Obliczenia wykonano stosując do zbiorów AM1 i AM2 program komputerowy IMGW, identyfikujący dla obydwu zbiorów rozkład logarytmiczno-normalny (spośród czterech) jako najbardziej zgodny z danymi pomiarowymi. Analizę trzeciego zbioru danych (zawierających 261 przepływów) przeprowadzono przy wykorzystaniu arkusza kalkulacyjnego. Na podstawie wyników obliczeń stwierdzono znikome różnice w wartościach przepływów maksymalnych prawdopodobnych przy wykorzystaniu zbiorów AM1 i AM2 oraz znaczne różnice w porównaniu z wartościami uzyskanymi przy wykorzystaniu trzeciego zbioru danych.

MS. received November 2007

Authors' address:

Kazimierz Banasik, Andrzej Byczkowski
Warsaw University of Life Sciences – SGGW
Department of Water Engineering and
Environmental Restoration
02-776 Warszawa, ul. Nowoursynowska 159
Poland
e-mail: kazimierz_banasik@sggw.pl