

# A comparison of three parameter estimation methods of the gamma distribution of annual maximum low flow duration and deficit in the Upper Vistula catchment (Poland)

Katarzyna Baran-Gurgul<sup>1,\*</sup>

<sup>1</sup>Cracow University of Technology, Department of Environmental Engineering, Institute of Water Engineering and Water Management, ul. Warszawska 24, 31-155 Cracow, Poland

**Abstract.** Based on 30-year 24-hour flow sequences at 69 water gauging stations in the Upper Vistula catchment, it was determined that the probability distributions of the low flow duration and its maximum annual deficit can be described by the gamma distribution with the estimated parameters by the methods: MOM, the method of moments, LMOM, the method of linear moments, and MLE, the method of maximum likelihood. The stationarity of the time series was tested by the Mann-Kendall correlation using the Hamed and Rao variance correction. The low flows were defined by the SPA method, with the limit flow  $Q_{70\%}$ . The quality of the match was tested by the Anderson-Darling goodness of fit test. This test allowed accepting the gamma distribution in all analysed cases, regardless of the method used to estimate the distribution parameters, since the  $p_v$  ( $p$ -values) values were greater than 5% (over 18% for  $T_{max}$  and 7.5% for  $V_{max}$ ). The highest  $p_v$  values for individual water gauging stations, as well as the highest 90%  $T_{max}$  and  $V_{max}$  quantiles were noted using LMOM to estimate the gamma distribution parameters. The highest 90%  $T_{max}$  and  $V_{max}$  quantiles were observed in the uppermost part of the studied area.

## 1 Introduction

**The river low flow** is usually referred to as the time period in which the water flow in a given section of the river is lower than the conventionally assumed flow threshold [1–5]. This method of determining the low flow (threshold level methods) was introduced in the United States by Yevjevich [6] and in Poland by Zielińska [7].

Low flow is usually defined by one of two methods: PUT (Pit Under Threshold) [8] or SPA (Sequent Peak Algorithm) [2]. The beginning of the low flow occurs in both methods at the time  $t_p$  of the decrease of the flow below the limit  $Q_g$ . Differences between the use of these methods are evident in the determination of the end time  $t_k$  of the low flow duration and its deficit. In the PUT method, it is assumed that the end of the low flow occurs when the flow in the river rises back above the  $Q_g$  level. In the SPA method, the differences  $Q_g - Q_i$  are added (starting with  $t_p$ ) in subsequent days until the sum obtained ceases to be positive, i.e. until the resulting water deficit is balanced by flows higher than  $Q_g$ . The day on which the sum reaches the maximum value is the  $t_k$  day end of the SPA low flow and determines both the  $T$ -time of the SPA low flow duration and its deficit  $V$  [2, 9, 10].

**The maximum annual low flow.** An important issue in hydrology is the risk assessment of low flow hazards, and the maximum values of the duration and deficit of a low flow are of particular importance. The maximum

values are usually determined by one of two methods: annual maximum series or partial duration series. The annual maximum series (AMS) contains only the highest values that occur in each year of the examined multi-year, while the partial duration series (PDS) covers all values in the considered period, above the set limit value.

Estimation of one-dimensional distributions of duration and deficits of the maximum low flow (using the PDS method) were studied by, *inter alia*, Zelenhasić and Salvai [11], while in Poland the issue was studied primarily by Jakubowski [12, 13]. In order to determine the probability of not exceeding the volume of the low flow and its duration, Jakubowski and Radczuk, developed the computer program NIZOWKA2003 [14], in which the number of low flows (in a year or season) is estimated by the following discrete distributions: Poisson and Pascal, and the duration of low flows and their deficit by the following continuous distributions: Pearson type III, Weibull, lognormal, Johnson, double exponential, and generalized Pareto.

This paper uses the annual maximum series of the duration and deficit of the low flow.

Finding the probability distribution of time variables  $T_{max}$  duration and volume  $V_{max}$  of the maximum low flows allows estimating the value of exceeding the pre-set  $T_{max}$  and/or  $V_{max}$  with a given return period, i.e. a numerical risk/hazard assessment manifested as a low flow with a defined characteristic. The use of the gamma distribution to describe the maximum duration and the

\* Corresponding author: [Katarzyna.Baran-Gurgul@iigw.pk.edu.pl](mailto:Katarzyna.Baran-Gurgul@iigw.pk.edu.pl)

maximum deficit of a low flow can be found in many works [4, 10, 13–17].

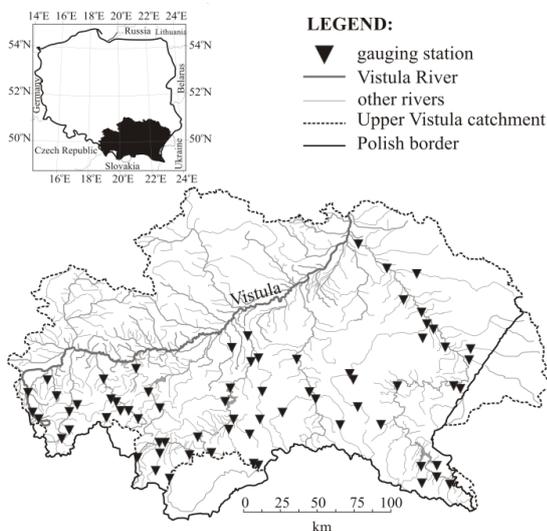
The most commonly used methods of estimating the distribution parameters of the characteristics of the low flows are: the method of maximum likelihood MLE [12–14], the method of moments MOM [18] and the method of linear moments LMOM [18]. Each of these methods has its advantages and disadvantages. The method of moments cannot be used for non-simple tests, and the estimators of parameters are sensitive to outliers [19]. The most frequently used method in Poland is the method of maximum likelihood that can be used for non-simple tests. The method of linear moments is less sensitive to significant outlier values in the sample [19]. Sharma and Panu [18] suggest that for small sample size of drought events (<30), it is advisable to use LMOM, and for a large sample — MOM.

The goal of this work is to investigate whether the gamma distribution can be used to describe the probability distribution of the duration and volume of the maximum low flow in the Upper Vistula catchment. In the work, 90% gamma distribution quantiles were determined with parameters estimated in three ways: the method of maximum likelihood, the method of moments and the method of linear moments. The influence of the method of estimation of gamma distribution parameters on the quality of distribution adjustment was also investigated.

## 2 Materials and methods

### 2.1 Study area and data

The study is based on daily discharge series for the period from 1.11.1983 to 31.10.2013. This yields a total of 30 hydrologic years and 10,958 discharge values obtained at each of 69 water gauging stations located on the right-bank tributaries of the Upper Vistula River in southern Poland (Fig. 1).



**Fig. 1.** Study area and location of selected water gauging stations in the Upper Vistula catchment.

The data were made available by the Institute of Meteorology and Water Management-National Research Institute (IMWM-NRI).

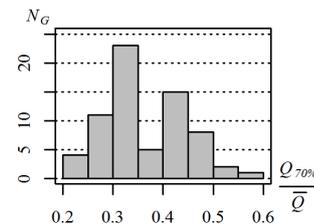
The gauging cross-sections close catchments of various area ranging from 24.7 km<sup>2</sup> to 16,824 km<sup>2</sup>, and the elevation of gauging stations varies from 139.1 m a.s.l. to 965.6 m a.s.l. About half of the catchment gauging station elevations (more than 35 out of 69) do not exceed 300 m a.s.l.; 53% of catchment areas are in between 100 and 1000 km<sup>2</sup>.

The area of the Upper Vistula is diverse in terms of climate, mainly due to significant morphological differences, chiefly owing to the varied land relief and large area denivelations [20, 21]. In the right-bank part of the Upper Vistula catchment, three main climatic regions can be distinguished. Because the southern part of the area lies in the Carpathians and the Beskid Mountains, the region has a mountain climate, rich in precipitation. The boundary of this region coincides with the annual isotherm of 7 °C and the annual isohyet of 900 mm. Located more to the north is the climate of the Carpathian Foothills — moderately warm and humid, with an average annual temperature of 7–8 °C and an annual rainfall of 700–900 mm. The climate of the region of Kotliny Podgórskie [Sub-mountain Basins], which is located in the northernmost part of the studied area, is warmer and moderately dry [21].

### 2.2 Selection of maximum low flow

Based on the 30-year daily flow sequences, at each of the 69 water gauging stations, using the SPA method and the limit flow  $Q_{70\%}$  read from the flow duration curve, two characteristics of the low flow were determined: duration  $T$  [day] and deficit  $V$  [day]. The low flow deficit is regulated by a multi-year average flow  $\bar{Q}$  and is therefore expressed in 24 hours. The determination of the volume expressed in days shows how many days, at a given water gauging station, the low flow deficit could be filled with the average flow.

The normalized limit flow  $Q_{70\%}$  calculated, for the 69 water gauging stations, varies between 0.2 and 0.6, and in most cases (43 per 69) it is smaller than 0.4 (Fig. 2).



**Fig. 2.** Distribution of the number of catchments,  $N_G$ , versus normalized flow,  $Q_{70\%}$ , at the 69 gauging stations.

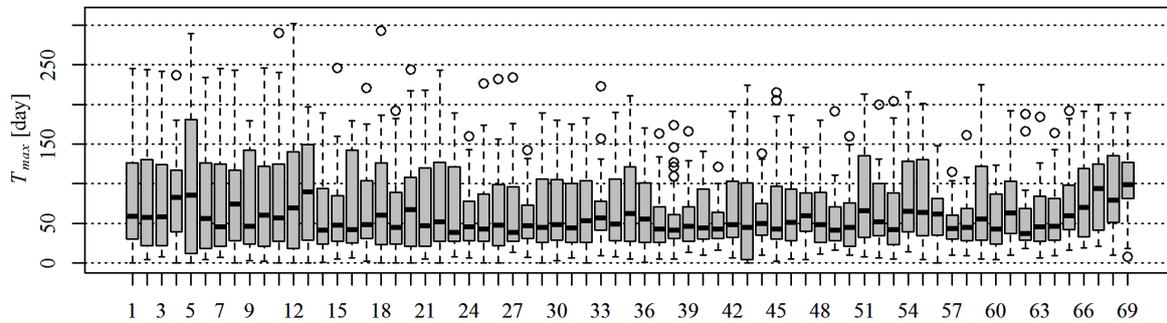
The maximum annual low flows are understood in two ways: as a low flow with the longest annual duration in a year,  $T_{max,i}$ ,  $i = 1, 2, \dots, 30$ , and as a low flow with the highest annual deficit  $V_{max,i}$ ,  $i = 1, 2, \dots, 30$ . If a low flow circuit started in one hydrological year and ended in the next one, it was not divided, but assigned in its entirety to the year in which its mean was observed. This

approach was proposed by Zelenhasić and Salvai [11]. 138 (69 water gauging stations, two definitions of the maximum low flows) time series of maximum low flows were obtained.

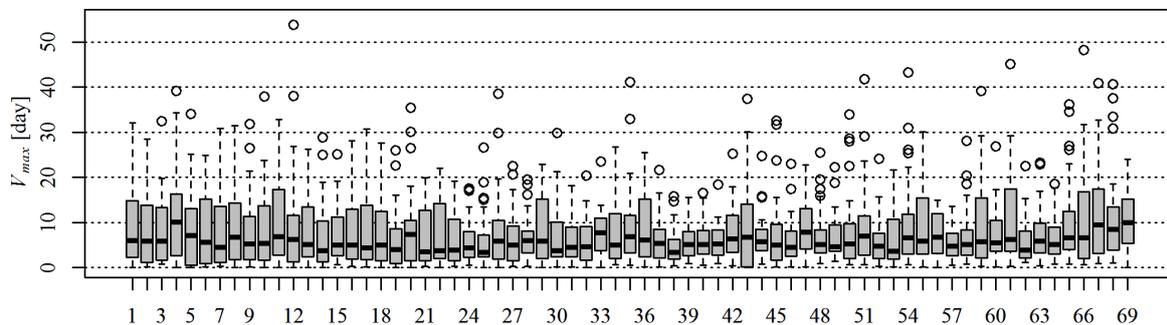
During the considered period (1983–2013), the maximum low flows lasted, on average, from 37 days to 98.5 days, with a maximum of 302 days (Fig. 3). Low flow deficit, depending on the water gauging station limit, ranged from 3.4 days to 10 days on average. There is no dependence of the median  $T_{max}$  and the median  $V_{max}$

on the height  $H$  of the gauge (the Pearson correlation coefficient between the median  $T_{max}$  and  $H$  is 0.16, and the median  $V_{max}$  and  $H$  – 0.31) (Figs. 3 and 4).

There is no dependence of the median  $T_{max}$  and the height  $H$  of the gauge (the Pearson correlation coefficient between the median  $T_{max}$  and  $H$  is 0.16) (Fig. 3). The Pearson correlation coefficient between the median  $V_{max}$  and  $H$  is 0.31 is significant (by the significance level, 0.05) (Fig. 4).



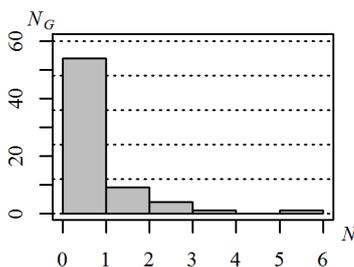
**Fig. 3.** Time  $T_{max}$  of the low flow duration at the 69 water gauging stations.



**Fig. 4.** Deficit  $V_{max}$  of the low flow at the 69 water gauging stations.

The maximum low flows usually start in September, the lowest probability of their occurrence is in March and April. In the multi-year surveys, 2010 was the "wettest" year (with the lowest number of low flow days) and 2012 the "driest" (with the highest number of low flow days).

At 54 (out of 69) water gauging stations, low flows were observed annually (Fig. 5), while at the remaining 22%, at least in one year all low flows were higher than the assumed limit flow. At one of the gauging stations (Mikuszowice, the Biała River) no low flows were observed for 6 years out of the 30 measured.



**Fig. 5.** Distribution of the number of water gauging stations,  $N_G$ , versus the number of years without low flows,  $N$ .

## 2.3 Methods

Because low flows need not be found in every year, zero values may appear in the ranks of duration and deficit of low flows, determined by the AMS method. In this situation, Stedinger et al. [22] recommend the use of a conditional probability model. An additional parameter describes the probability  $p_0$  that the duration or deficit of the low flow is zero. The distribution function  $F(x)$ , for any value  $x > 0$ , is as in the formula:

$$F(x) = p_0 + (1 - p_0)G(x) \quad (1)$$

where  $G(x)$  is a continuous cumulative distribution function for non-zero values of  $X$  (a random variable  $X$  representing  $T_{max}$  or  $V_{max}$ ).

All statistical calculations were performed using the GNU R software package [23]. For all the tests considered in the paper, we fixed the significance level at 0.05.

### 2.3.1 Mann–Kendall test

The stationarity of the time series duration  $T_{max}$  and the deficit  $V_{max}$  of the maximum annual low flows was tested

by the Mann–Kendall test [24–26]. The test statistic is defined as:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{sgn}(x_j - x_i) \quad (2)$$

where  $x$  value ( $T_{max}$  or  $V_{max}$ ) with sample size  $n$ :

$$\text{sgn}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases} \quad (3)$$

When  $H_0: S = 0$  is true (i.e. stationarity), the mean and variance of  $S$  are given by:

$$E(S) = 0 \quad (4)$$

$$\text{var}(S) = \frac{n(n-1)(2n+5)}{18} \quad (5)$$

The above equality (5) occurs when there are no repeated elements of a time series (so-called ties). If there are more than a few repeating elements, then the formula (5) needs to be adjusted, which yields the following form [27]:

$$\text{var}(S) = \frac{n(n-1)(2n+5)}{18} - \frac{\sum_{m=1}^k t_m \cdot m(m-1)(2m+5)}{18} \quad (6)$$

where  $t_m$ ,  $m = 1, 2, \dots, k$ , indicates that there exists a number of groups with  $k$  repeating elements.

Kendall [25] proved that when  $n > 10$ , then the test statistic:

$$U = \frac{S - \text{sgn}(S)}{\sqrt{\text{var}(S)}} \quad (7)$$

approximately follows the standard normal distribution  $N(0;1)$ . If the  $p_v$  value of the hypothesis test is less or equal to the chosen significance level  $\alpha_v$ , the null hypothesis must be rejected.

The basic assumption when using this test is a lack of autocorrelation in the data series. If this condition is not met and autocorrelation is positive, which is a frequent situation in hydrology, then the variance  $\text{var}(S)$  is underestimated. Bayley and Hammersley [26] introduced a variance adjustment:

$$\text{var}^*(S) = \frac{n}{n_S^*} \text{var}(S) \quad (8)$$

when using “an effective number of observations  $n_S^*$ ”, which Hamed and Rao [28] suggest to determine as:

$$\frac{n}{n_S^*} = 1 + \frac{2 \sum_{i=1}^{n-1} (n-i)(n-i-1)(n-i-2)\rho_S(i)}{n(n-1)(n-2)} \quad (9)$$

where  $n$  is the number of observations and  $\rho_S(i)$  is the autocorrelation function of ranks of observations. The variance adjustment is only calculated for data whose autocorrelations are significant at the 5% level (according to the Ljung–Box test [23]).

The variance  $\text{var}(S)$  in (6) is replaced by  $\text{var}^*(S)$  from (8) and the rest of the testing process is the same as in (7) (i.e. comparing the values of  $U$  with the quantiles of  $N(0;1)$  corresponding to the type I error of the test).

The description of the Mann–Kendall test with the Hamed and Rao correction, together with the use of this test to determine the stationarity of the series of low flows can be found in the work of Baran–Gurgul [29].

### 2.3.2 Gamma distribution

To obtain the probability distribution of the maximum low flow, the two-parameter gamma distribution was used. The probability density function of the gamma distribution is [23]:

$$f_{\Gamma}(x; \alpha, \lambda) = \frac{\alpha^{\lambda}}{\Gamma(\lambda)} x^{\lambda-1} e^{-\alpha x}, \quad x > 0, \alpha > 0, \lambda > 0 \quad (10)$$

where  $\lambda$  is the shape,  $1/\alpha$  – the scale parameter, and  $\Gamma(\cdot)$  – the gamma function.

### 2.3.3 Estimation of gamma distribution parameters

The paper uses three gamma distribution parameter estimation methods: MOM, LMOM and MLE.

Using individual methods, the estimators of gamma distribution parameters can be calculated according to formulas [20, 22, 23, 30]:

$$\hat{\lambda}_{\text{MOM}} = \left(\frac{\bar{x}}{s}\right)^2, \quad \hat{\alpha}_{\text{MOM}} = \frac{\hat{\lambda}_{\text{MOM}}}{\bar{x}} = \frac{\bar{x}}{s^2} \quad (11)$$

$$\text{if } 0 < t < 0.5 \text{ then } c = \pi t^2 \text{ and } \hat{\lambda}_{\text{LMOM}} = \frac{1-0.308 c}{c-0.05812 c^2+0.01765 c^3}$$

$$\text{if } 0.5 \leq t < 1 \text{ then } c = 1-t \text{ and } \hat{\lambda}_{\text{LMOM}} = \frac{0.7213 c-0.5947 c^2}{1-2.1817 c+1.2113 c^2} \quad (12)$$

$$\hat{\alpha}_{\text{LMOM}} = \frac{\hat{\lambda}_{\text{LMOM}}}{\bar{x}}$$

$$\hat{\lambda}_{\text{MLE}} = \frac{1 + \sqrt{1 + (4/3)(\ln \bar{x} - \overline{\ln x})}}{4(\ln \bar{x} - \overline{\ln x})}, \quad \hat{\alpha}_{\text{MLE}} = \frac{\hat{\lambda}_{\text{MLE}}}{\bar{x}} \quad (13)$$

where  $\bar{x}$  and  $s$  are the sample mean and standard deviation, respectively;  $t = t_2 / t_1$ , and  $t_r - r$ th linear moment of data sample ( $r = 1, 2$ ).

### 2.3.4 Anderson–Darling test

The quality of matching the gamma distribution to the data was tested with the Anderson–Darling test. For a random ordered sample:  $(x_1, x_2, \dots, x_n)$ , the  $AD$  statistic for this test is expressed by the formula [23]:

$$AD^2 = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) [\ln F(x_i) + \ln(1-F(x_{n-i+1}))] \quad (14)$$

where  $F$  is the cumulative distribution function of the specified distribution.

The critical values for the Anderson–Darling test are dependent on the specific distribution that is being tested. Tabulated values and formulas have been published in the various papers by Stephens (e.g. [31]) for a few specific distributions (normal, lognormal, exponential, Weibull, logistic, extreme value type I).

## 3 Results and discussion

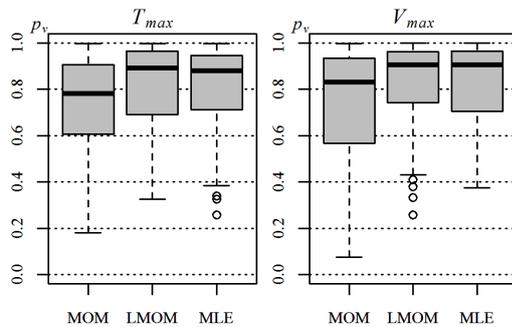
Investigation of the hypothesis about the stationarity of the time series of  $T_{max}$  duration and  $V_{max}$  deficit of the maximum annual low flows at the 69 water gauging stations was carried out with the non-parametric Mann–Kendall test. In all analysed cases, the Mann–Kendall test  $p_v$  values are greater than 5%, which allows acceptance

of the hypothesis about the stationarity of the time series of the maximum low flows  $T_{max}$  and  $V_{max}$ . This acceptance allows creating probability distributions of these variables and to calculate the values of  $T_{max}$  and  $V_{max}$  variables exceeded with a given probability.

Parameters of the gamma distribution of  $T_{max}$  and  $V_{max}$  variables were estimated using three methods: MOM, LMOM and MLE, and the quality of matching the gamma distribution to the data was tested using the Anderson–Darling test.

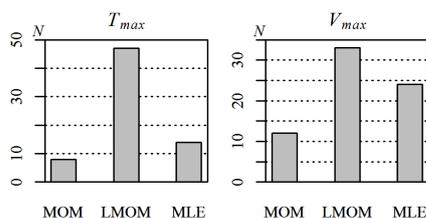
In all analysed cases, regardless of how the gamma distribution parameters are estimated, the  $p_v$  value of the Anderson–Darling test is above 5% (Fig. 6) and there is no reason to reject the hypothesis of theoretical and empirical agreement, which means that  $T_{max}$  and  $V_{max}$  can be determined, at the level of 5%, with the gamma distribution with parameters estimated by each of the methods used.

The minimum  $p_v$  values at the tested 69 water gauging stations are respectively: for the  $T_{max}$  variable: MOM — 18%, LMOM — 33%, MLE — 26%, and the  $V_{max}$  variable: MOM — 7%, LMOM — 26%, MLE — 37%. Furthermore, the median of  $p_v$  values at the 69 water gauging stations is the lowest for both low flow characteristics when estimating distribution parameters using the method of moments (0.78 for  $T_{max}$  and 0.83 for  $V_{max}$ ). In other cases, the  $p_v$  medians are high, 0.88–0.91 (Fig. 6).



**Fig. 6.** Calculated  $p_v$  values of the Anderson–Darling test, at 69 water gauging stations in the Upper Vistula catchment.

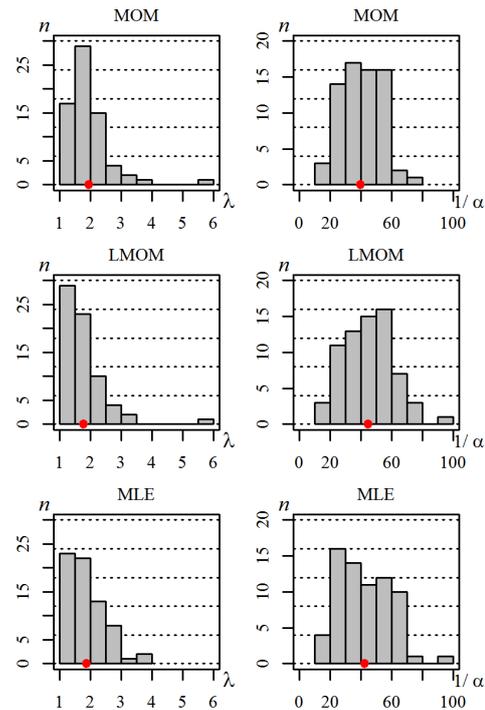
For individual gauging stations, the highest  $p_v$  of the Anderson–Darling test was usually found by estimating the distribution parameters using the LMOM method (68% of cases for  $T_{max}$  and 48% for  $V_{max}$ ), and the the least frequently — MOM (in 20% of cases for  $T_{max}$  and 35% for  $V_{max}$ ) (Fig. 7)



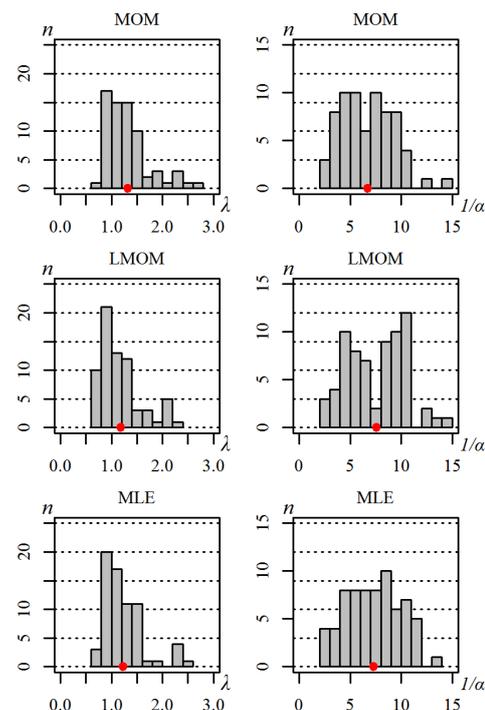
**Fig. 7.** Number of  $N$  water gauging stations with the highest  $p_v$  of the Anderson–Darling test which was found by estimating the distribution parameters using the three methods.

Figures 8 and 9 graphically illustrate the results of

parameter estimation ( $\lambda$ ,  $1/\alpha$ ) for the gamma probability distribution, respectively, for duration  $T_{max}$  and deficit  $V_{max}$  of the maximum low flow at all gauging stations.

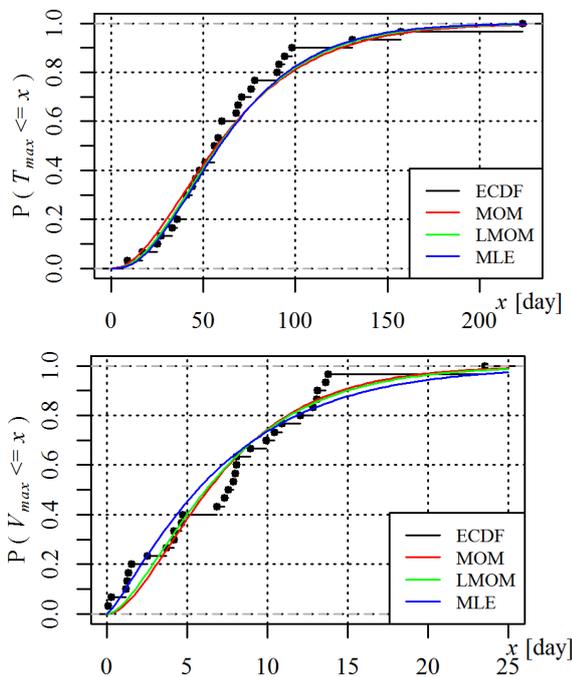


**Fig. 8.** Histograms (estimated with the methods: MOM, LMOM and MLE) of parameter values ( $\lambda$ ,  $1/\alpha$ ) for the gamma probability distribution of the  $T_{max}$  duration of the maximum low flow at the 69 water gauging stations. The average value of the parameter is marked with the red dot.



**Fig. 9.** Histograms (estimated with the methods: MOM, LMOM and MLE) of parameter values ( $\lambda$ ,  $1/\alpha$ ) for the gamma probability distribution of the  $V_{max}$  deficit of the maximum low flow at the 69 water gauging stations. The average value of the parameter is marked with the red dot.

The estimation of the gamma distribution parameters for duration  $T_{max}$  and deficit  $V_{max}$  of the maximum annual low flow results in obtaining (in relation to the use of LMOM or MLE) the average lowest values of the shape parameter  $\lambda$  ( $T_{max}$ : MOM — 1.94, LMOM — 1.76, MLE — 1.87,  $V_{max}$ : MOM — 1.31, LMOM — 1.17, MLE — 1.21), and the average highest values of the scale parameter  $1/\alpha$  ( $T_{max}$ : MOM — 39.77, LMOM — 44.54, MLE — 42.32,  $V_{max}$ : MOM — 6.67, LMOM — 7.57, MLE — 7.29).

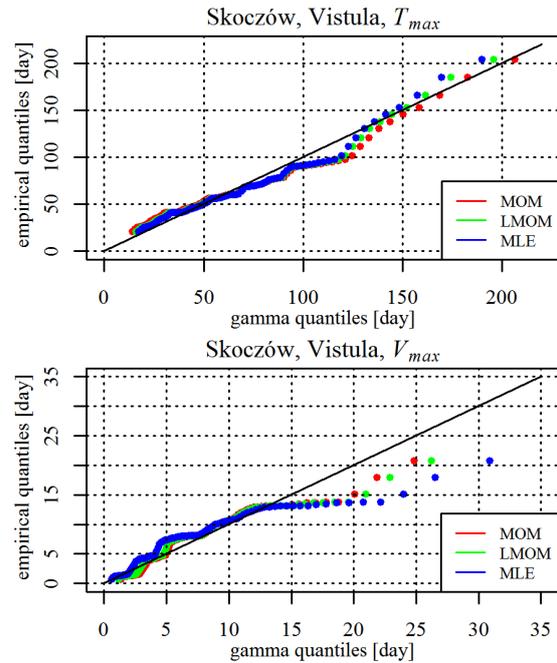


**Fig. 10.** Empirical distributions ECDF and the gamma distributions (with parameters determined by the methods MOM, LMOM and MLE) of variables  $T_{max}$  and  $V_{max}$ , at the Skoczów gauging station on the Vistula River.

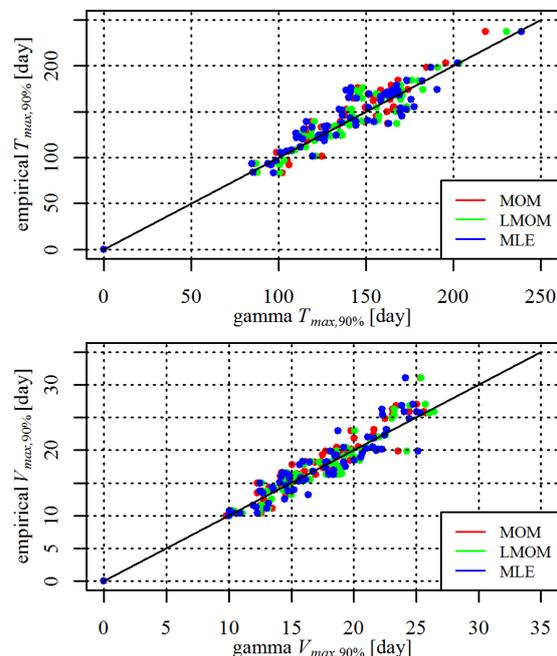
The quality of the fit of the gamma distribution with parameters determined in three ways (MOM, LMOM and MLE) is shown, as an example, for the Skoczów section, in two ways: Fig. 10 compares the empirical distribution of  $T_{max}$  and  $V_{max}$  variables and the gamma distribution of these variables with parameters determined in three ways (MOM, LMOM and MLE), and in Fig. 11, the quantiles of these distributions are compared in the Q-Q plot charts. The fit is better for  $T_{max}$  (values  $p_v$  of the Anderson–Darling test are respectively:  $p_{v,MOM} = 0.77$ ,  $p_{v,LMOM} = 0.89$ ,  $p_{v,MLE} = 0.92$ ) than for  $V_{max}$  ( $p_{v,MOM} = 0.24$ ,  $p_{v,LMOM} = 0.33$ ,  $p_{v,MLE} = 0.41$ ), especially for larger values.

**Table 1.** The Pearson correlation between empirical quantiles of the variables  $T_{max}$  and  $V_{max}$  and the gamma distribution with the estimated parameters of MOM, LMOM and MLE.

	MOM	LMOM	MLE
$T_{max}$	0.920	0.923	0.904
$V_{max}$	0.946	0.946	0.916

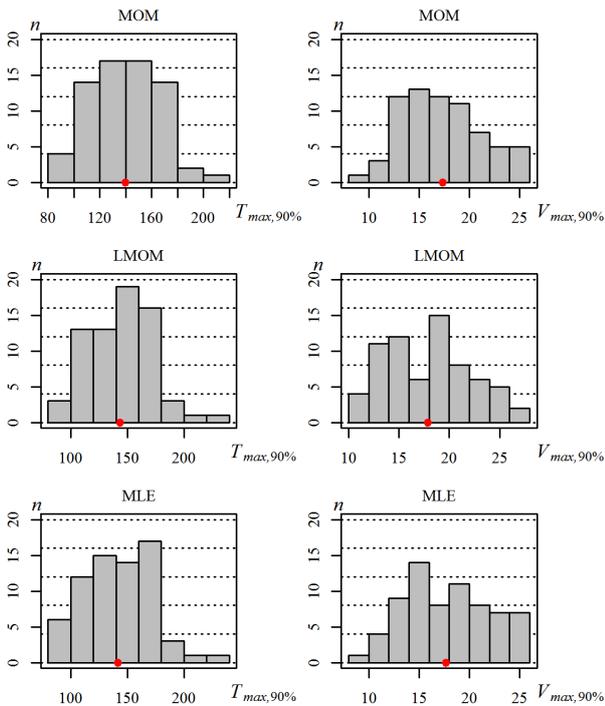


**Fig. 11.** Q-Q plots of a sample of data versus the gamma distribution (with parameters determined by the following methods: MOM, LMOM and MLE), for variables  $T_{max}$  and  $V_{max}$ , at the gauging station Skoczów on the Vistula River.



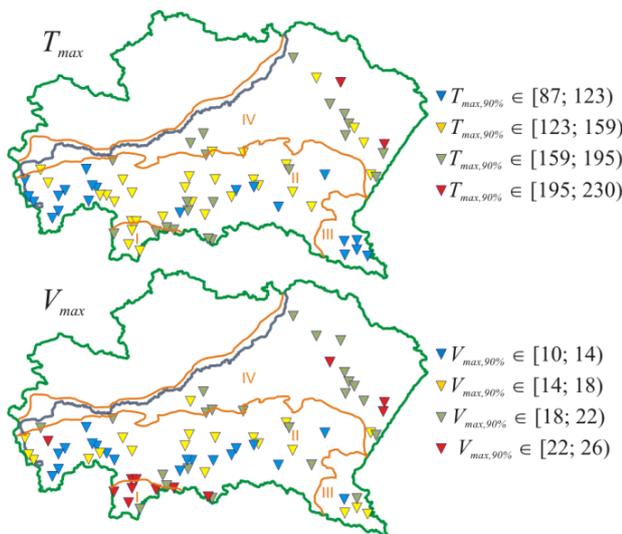
**Fig. 12.** Q-Q plots of a sample of data versus the gamma distributions (with parameters determined by the following methods: MOM, LMOM and MLE), for 90% quantiles, for variables  $T_{max}$  and  $V_{max}$ , at the 69 gauging stations.

The highest 90% gamma distribution quantiles were obtained for both the duration  $T_{max}$  and the deficit  $V_{max}$  of the maximum low flow using the LMOM method for estimation ( $T_{max,90\%}$ : mean 143.3 days, max 230.5 days,  $V_{max,90\%}$ : mean 17.8 days, max 26.4 days) and the lowest using the MOM method ( $T_{max,90\%}$ : mean 139.5 days, max 218.4 days,  $V_{max,90\%}$ : mean 17.3 days, max 25.3 days) (Fig. 13).



**Fig. 13.** Histograms of quantiles  $T_{max,90\%}$  and  $V_{max,90\%}$  of the gamma distribution with parameters estimated by: MOM, LMOM and MLE at the 69 water gauging stations. The average value of the quantiles is marked with the red dot.

In Fig. 14, the 90% the quantiles gamma distribution with LMOM parameters are shown for  $T_{max}$  and  $V_{max}$ . The highest values of the 90% quantiles were observed in the highest areas (the Tatra Mountains, Area I), the lowest in the Beskid Mountains (III) and Outer Carpathians (II).



**Fig. 14.** Spatial distribution of the 10-year duration  $T_{max}$  and deficit  $V_{max}$  of the maximum low flow in the Upper Vistula catchment (gamma distribution with LMOM-estimated parameters). Boundaries of physical and geographical units acc. to [32]: I - Central Western Carpathians, II - Outer Western Carpathians, III - Outer Eastern Carpathians, IV - Eastern Sub-Carpathian region.

## 4 Final remarks

The probability distribution of the best choice is important for the description of low flow statistics for studies related to low flow analysis.

Stationarity study of 30-year time series duration  $T_{max}$  and deficit  $V_{max}$  of the maximum annual low flow using the Mann-Kendall test showed that at all 69 water gauging stations there were no grounds for rejecting the hypothesis of lack of trend.

In all analysed cases, irrespective of how the gamma distribution parameters are estimated, the variables of duration  $T_{max}$  and deficit  $V_{max}$  can be described by a one-dimensional gamma distribution at the significance level of 5%. The quality of the fit (which was tested with the Anderson–Darling goodness of fit test) is high, the  $p_v$  probability values are the highest when using the LMOM method to estimate the distribution parameters ( $T_{max}$ :  $p_{v,min} = 32\%$ ,  $p_{v,mean} = 83\%$ ,  $V_{max}$ :  $p_{v,min} = 26\%$ ,  $p_{v,mean} = 83\%$ ).

Confirmation of a good fit of the gamma distribution is a high correlation, above 90%, between the quantiles of empirical distributions and the quantiles of the gamma distribution, regardless of the method of estimating the parameters of this distribution.

The highest 90% quantiles of the gamma distribution were obtained, both for the duration  $T_{max}$  and the deficit  $V_{max}$  of the maximum low flow, using the LMOM method for estimation. Depending on the gauging station,  $T_{max,90\%}$  is on average around 142 days and  $V_{max,90\%}$  around 18 days.

## 5 Conclusions

Based on the analysis carried out on 30-year long daily flows at the 69 water gauging stations, we can draw the following conclusions:

1. The gamma distribution can be used to describe the duration and deficit of the maximum low flow in the Upper Vistula catchment.
2. In most catchments, the best fit of the gamma distribution according to the Anderson–Darling test was obtained by estimating the parameters using the method of linear moments.
3. The largest 90% quantiles were obtained by estimating the gamma distribution parameters using the method of linear moments.
4. In the Carpathian part of the Upper Vistula catchment, the area which faces the highest risk of low flows with the longer duration (over 195 days) and the largest deficit (over 22 days) is the area of the Central Western Carpathians.

The Carpathian catchment of the Upper Vistula is a very diverse region both due to its shape, as well as to its geological structure and climate. The Central Western Carpathians, with their highest part, the Tatras, are built of granitic and metamorphic rocks, the Outer Carpathians and Sub-Carpathian region are made of flysch, mainly sandstone and shale [33].

The Carpathian rivers are characterized by high variability of water levels. This is the result of heavy precipitation, which cause a rapid increase in flow in rivers. The low permeable or impermeable flysch substrate combined with reduced evaporation in the mountain area (by lower air temperatures) causes rapid water runoff. However, the heavy snowfall and a snow cover which stays in the field for a long time contribute to the formation of deep, long-lasting winter low flows, characteristic especially of the Sudety, the Carpathian mountains [34–36].

## References

1. H. Hisdal, K. Stahl, L.M. Tallaksen, S. Demuth, *Inter. J. Climat.* **21**, 317 (2001)
2. H. Hisdal, L.M. Tallaksen (ed.), *Drought Event Definition* (ARIDE Technical Report No. 6., Assessment of the Regional Impact of Droughts in Europe, University of Oslo, 2000)
3. V.U. Smakhtin, *J. Hydrol.* **240**, 147 (2001)
4. E. Tomaszewski, *Wieloletnia i sezonowa dynamika niszówek w rzekach środkowej Polski (Multiannual and seasonal dynamics of low flows in rivers of central Poland)* (University of Lodz, Łódź, 2012)
5. S. Węglarczyk (ed.), *Słownik terminów związanych z gospodarowaniem zasobami wodnymi (Dictionary of terms related to the management of water resources)* (Cracow University of Technology, Kraków, 2006)
6. V. Yevjevich, *Hydrology Paper 23* (Colorado State University, Fort Collins, Colorado, 1967)
7. M. Zielińska, *Prz. Geofiz.* **VIII (XVI)**, 1–2, 75 (1963)
8. L. Gottschalk, K. Yu, E. Leblois, L. Xiong, *J. Hydrol.* **481**, 204 (2013)
9. H.A.J. van Lanen, Z. Kundzewicz, L.M. Tallaksen, H. Hisdal, M. Fendeková, C. Prudhomme, *Indices for different types of droughts and floods at different scales* (WATCH Water and Global Change, 2008)
10. T. Tokarczyk, *Niszówka jako wskaźnik suszy hydrologicznej (Low flow as an indicator of hydrological drought)* (Institute of Meteorology and Water Management, Warszawa, 2010)
11. E. Zelenhasić, A. Salvai, *Water Resour. Res.* **23**, 156 (1987)
12. W. Jakubowski, *Prz. Geofiz.* **3–4**, 229 (2005)
13. W. Jakubowski, *Rozkłady prawdopodobieństwa w ocenie suszy hydrologicznej*, HABIL. Diss. (*Probability distributions in hydrological drought estimation*) (Wrocław University of Environmental and Life Sciences, Wrocław, 2011)
14. L.M. Tallaksen, H.A.J. van Lanen (ed.), *Hydrological Drought, Processes and Estimation Methods for Streamflow and Groundwater*, *Dev. Water Sci.* **48**. (Els. Sc., B.V.12., Amsterdam, 2004)
15. A.K. Fleig, L.M. Tallaksen, H. Hisdal, S. Demuth, *Hydrol. Earth Syst. Sc.* **10**, 535 (2006)
16. E. Tomaszewski, *E. Acta Univ. Lodz., Folia Geogr. Phys.* **8**, 89 (2007)
17. J.H. Sung, E.S. Chung, *Hydrol. Earth Syst. Sc.* **18**, 3341 (2014)
18. T.C. Sharma, U.S. Panu, *Hydrolog. Sci. J.* **60**, 1783 (2015)
19. S. Węglarczyk, *Statystyka w inżynierii środowiska (Statistics of environmental engineering)* (Cracow University of Technology Publisher, Kraków, 2010)
20. J. Chowaniec, *Hydrogeology study of the western part of the Polish Carpathians* (Biul. Państw. Inst. Geol. **434**, Kraków, 2009)
21. I. Dynowska, M. Maciejewski (ed.), *Dorzecze górnej Wisły (The Upper Vistula basin)* (PWN, Warszawa-Kraków, 1991)
22. J.R. Stedinger, R.M. Vogel, E. Foufoula-Georgiou, in: D.R. Maidment (ed.), *Handbook of Hydrology* (McGraw-Hill, New York, 1993)
23. R Core Team, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Austria, <http://www.R-project.org/>, 2017)
24. H.B. Mann, *Econometrica* **13**, 245 (1945)
25. M.G. Kendall, *Biometrika* **30**, 81 (1938)
26. K.H. Hamed, A.R. Rao, *J. Hydrol.* **204**, 182 (1998)
27. D.R. Helsel, R.M. Hirsch, *Statistical methods in water resources* (U. S. Geological Survey, Book 4, 2002), <https://pubs.usgs.gov/twri/twri4a3/> (accessed 26 March 2018)
28. *NIST/SEMATECH e-Handbook of Statistical Methods*, <http://www.itl.nist.gov/div898/handbook/> (accessed: 26 March 2018)
29. K. Baran-Gurgul, *E3S Web of Conferences* **17**, 00004 (2017)
30. J.R.M. Hosking, *J. R. Stat. Soc. B* **52**, 105 (1990)
31. M.A. Stephens, *Biometrika* **64**, 583 (1977)
32. J. Kondracki, *Geografia regionalna Polski (Regional geography of Poland)* (PWN, Warszawa, 2000)
33. B. Ratomska, *Ocena suszy w aspekcie hydrologicznym (Evaluation of the drought in the hydrological aspect)*, PhD Th. (Cracow University of Technology, Kraków, 1993)
34. Z. Ziemońska, *Pr. Geogr.* **103** (1973)
35. A. Tlalka, *Przestrzenne zróżnicowanie niszówek letnich w dorzeczu górnej Wisły (Spatial variability of summer low flows in the Upper Vistula River catchment)*, HABIL. Diss. (Jagiellonian University, Kraków, 1982)
36. R. Soja, *Pr. Geogr.* **186** (2002)