Revista Română de Inginerie Civilă, Volumul 15 (2024), Numărul 3

Exploring the Applicability and Insights of the Pearson Type III Distribution in Flood Frequency Analysis

Explorarea aplicabilității și perspectivelor distribuției Pearson de tip III în analiza frecvenței inundațiilor

Cristian Anghel[1], Stefan Ciprian Stanca[2] , Cornel Ilinca[3]

[1,2,3]Technical University of Civil Engineering Bucharest
Lacul Tei Boulevard Nr. 124, Bucharest 020396, Romania
[1]*E-mail:* cristian.anghel@utcb.ro
[2]*E-mail:* stefan-ciprian.stanca@phd.utcb.ro
[4]*E-mail:* cornel.ilinca@utcb.ro

DOI: 10.37789/rjce.2024.15.3.6

Abstract. *Being the parent distribution in Romania, the Pearson III distribution (PE3) is one of the statistical distributions that is most frequently used in flood frequency analysis (FFA). All the components required for the simple implementation of PE3 distribution in FFA are presented in this manuscript. The estimated methods and exact and approximate relationships for estimating the parameters and frequency factors particular to the distribution are described. All of these factors are used to identify maximum flows with various annual exceedance probabilities, using data collected at the 6 rivers with different morphometric characteristics and different lengths of data series. Five parameter estimation techniques are used in this comparative analysis, i.e. the method of ordinary moments (MOM), the method of linear moments (L-moments), the high-order linear moments method (LH-moments), the method of maximum likelihood (MLE) and the method of least squares (LSM). Given the results, it can be assumed that the L-moments approach is more reliable, stable, robust and less sensitive to variations in recorded data lengths, as well as to the presence of outliers.*

Key words: approximate form; confidence interval; estimation parameters; frequency factors; method of ordinary moments; method of linear moments; Pearson III.

1. Introduction

The flood frequency analysis (FFA) enables the computation of values with a certain likelihood of occurrence, which is crucial in the management of water resources and the design of hydrotechnical projects.

The Pearson III (PE3) distribution is one of the distributions that is most frequently employed in the statistical analysis of extreme data, along with the Log-Normal, GEV, and Log-Pearson distributions [1-7]. The PE3 distribution was applied, using different parameter estimation methods, for the frequency analysis of floods in [5, 8–10], the frequency analysis of maximum precipitation in [11–13], and the low flow frequency analysis in [4, 14]. In the investigation of flood frequency in Romania, PE3 serves as the parent distribution [7, 10].

The distribution is a special case of the four-parameter gamma distribution and a generalized version of the two-parameter gamma distribution. It is a member of the family of gamma distributions. In Romania, the PE3 distribution is used exclusively using the method of ordinary moments using the Foster-Ribkin table and linear interpolation, an approach that represents a legacy of the Soviet influence, inferior compared to modern methods of analysis.

Without closed forms, the cumulative complementary function (CDF) and the inverse function (quantile) of the PE3 distribution are represented in this article using predefined Mathcad functions that are comparable to other functions from other specialized programs (Excell).

For the method of ordinary moments (MOM), L-moments method and LH-moments, the quantile function can also be represented with the frequency factor, which is a real help considering the inverse function is defined by the gamma function.

In general, the PE3 distribution is applied using the MOM and L-moments parameter estimation methods, which are two of the most used parameter estimation methods in FFA [1, 4, 5, 10]. In comparison to other parameter estimate approaches, the L-moments method is renowned for being far more stable and less subject to bias [5,8,14,15]. The higher order linear moments (LH-moments) approach can be used to generate the "separation effect" [16]. Without explicit sample censoring, Wang proposed this method in 1997 [17], and it quickly rose to become one of the FFA's most used techniques. Only when FFA is employing the Annual Maximum Series (AMS) is its use advised. This approach lessens the impact of small maximum values in the frequency analysis by generalizing the approach of linear moments. As a result, high maximum values—always reflecting floods—are given more significance.

Regarding parameter estimation with MOM, it is generally preferred because the estimation relationships are simple and easy to use. However, it presents the disadvantage of the fact that higher-order statistical indicators (skewness and kurtosis) require correction. A solution to correct the skewness coefficient is represented by the approach of Bobee and Robitaille [1,2,5]. In Romania, this impediment was partially solved by choosing the skewness depending on the genesis of the maximum flows. Thus, according to Romanian regulations [18], a coefficient of 2 is chosen if the maximum flows come exclusively from snow melt, a coefficient of 3 if the origin is mixed (snow melt and rain), respectively a coefficient of 4 if the maximum flows come exclusively from rains. Unfortunately, this approach is often used incorrectly and excessively, because it is only valid for hydrographic basins with a surface of less than 100 km^2, because it uses an approximation considering the coefficient of variation having the value 1 and the maximum flows having the genesis exclusively from rains.

To estimate the parameters with L-moments (and also for MLE and LSM), it is necessary to solve a system of nonlinear equations, which leads to some difficulties. Thus, for the ease use of it, parameters approximation relations (for L-moments) are presented, using polynomial, exponential or rational functions. An important contribution, was made by Hosking [8] who for the first time presented relations for the estimation of the shape parameter within the complete domain of L-skewness, relations improved by Anghel and Ilinca [7].

For the LH-moments method, important contributions were made by [32-41], in which a significant number of distributions were analyzed, among which the most important are Wakeby, Lambda, Pearson V, the CHI, the inverse CHI, the Wilson–Hilferty, the Pseudo-Weibull, the Log-normal, the generalized Pareto Type I and the Fréchet distributions.

Regarding the least squares method, the manuscript presents a comparative analysis that identifies the best empirical probability that fits the PE3 distribution, so that by sampling the errors on the three levels (statistical indicators, estimation of parameters and quantities) to be as small as possible.

For the maximum likelihood method, the parameters of the PE3 distribution have a valid solution only if the skewness coefficient is lower than 2, an aspect noted both in the present analysis and in the observations of other researchers [5].

Given that uncertainties are a part of all statistical analyses, the relationships for calculating the confidence interval for the Pearson III distribution—which is required to quantify uncertainties—are described using both the Chow [19] (for MOM, L- and LH-moments) and Kite approximations (for MOM) [5]. The Chow's relation, for a 95% confidence level, is based on a Gaussian assumption, being a simplified approach. In general, all the quantiles that exceed the probabilities of the recorded values are characterized by a significant degree of uncertainty since the observed data is relatively short in duration. There are three levels of uncertainty due to the bias introduced by the fluctuation in the length of the recorded data, which must be considered when estimating the statistical indicators unique to the used method as well as when determining the values of the parameters and quantiles. These levels of uncertainty, which are distinctive to the Pearson III distribution, along with the estimate techniques and the range of data lengths, are presented in the text.

Thus, in order to highlight all these particular aspects of the PE3 distribution using these 5 parameter estimation methods, a comparative analysis is presented on 6 case studies, with different morphometric and statistical characteristics.

2. Methods

2.1. Probability Density Function and Cumulative Distribution Function

The probability density function, $f(x)$ and the complementary cumulative distribution function $F(x)$, for PE3 are [5,7,20]:

$$f(x) = \frac{(x-\gamma)^{\alpha-1}}{\beta^{\alpha}\cdot\Gamma(\alpha)}\cdot exp\left(-\frac{x-\gamma}{\beta}\right) = \frac{1}{\beta}\cdot dgamma\left(\frac{x-\gamma}{\beta},\alpha\right) \quad 1)$$

$$F(x) = 1 - \frac{1}{\beta \cdot \Gamma(\alpha)} \cdot \int_{\gamma}^{x} \left(\frac{x-\gamma}{\beta}\right)^{\alpha-1} \cdot exp\left(-\frac{x-\gamma}{\beta}\right) dx = \frac{\Gamma\left(\alpha, \frac{x-\gamma}{\beta}\right)}{\Gamma(\alpha)} \quad 2)$$

where α, β, γ are the shape, the scale and the position parameters and x can take any values of range $\gamma < x < \infty$ if $\beta > 0$ or $-\infty < x < \gamma$ if $\beta < 0$ and $\alpha > 0$.

2.2. Quantile function

The PE3 distribution does not have a closed form for the inverse function $x(p)$. This can be expressed using predefined functions from dedicated programs, such as: $qgamma$ function (Mathcad), $gamma.inv$ function (Excell), etc. In this article, the relationships are defined using predefined functions in Mathcad.

The quantile of the PE3 distribution has the following expression:

$$x(p) = x\big(F(x)\big) = F(x)^{-1} = \gamma + \beta \cdot qgamma(1-p, \alpha) \quad 3)$$

where p is the probability of exceedance. If $\beta < 0$ (negative skewness) then the first argument of the inverse of the distribution function Gamma, $\Gamma^{-1}(1-p;\alpha)$ becomes $\Gamma^{-1}(p;\alpha)$.

The built-in function from Mathcad $qgamma(1-p,\alpha) = \gamma^{-1}\big((1-p)\cdot\Gamma(\alpha),\alpha\big)$, returns the inverse cumulative probability distribution for probability p, for the Gamma distribution, where γ^{-1} is the inverse of the lower incomplete gamma function.

Based on the frequency factor, the inverse functions for MOM, L-moments, and LH-moments can be written as follows:

$$x(p) = \mu + \sigma \cdot K_{MOM}(p,\alpha) \quad (4)$$

$$x(p) = L_1 + L_2 \cdot K_L(p,\alpha) \quad (5)$$

$$x(p) = L_{H1} + L_{H2} \cdot K_{LH}(p,\alpha) \quad (6)$$

where μ is the expectation and σ $^{\sigma}$ is the standard deviation; L_1 and L_2 are the first two L-moments; L_{H1} and L_{H2} are the first two LH-moments; $K_{MOM}(p,\alpha), K_L(p,\alpha)$ and $K_{LH}(p,\alpha)$ represent the frequency factors for estimating the parameters with MOM L-moments and LH-moments.

The exact relationships for frequency factors, are:

$$K_{MOM}(p,\alpha) = \frac{qgamma(1-p,\alpha) - \alpha}{\sqrt{\alpha}} \quad (7)$$

$$K_L(p,\alpha) = \frac{\sqrt{\pi}\cdot\Gamma(\alpha)(qgamma(1-p,\alpha) - \alpha)}{\Gamma(\alpha+0.5)} \quad (8)$$

$$K_{LH}(p,\alpha)=\frac{\frac{2}{3}\cdot\left(qgamma(1-p,\alpha)-2\cdot z_2\right)}{z_1} \tag{9}$$

where the expressions for z_1 and z_2 can be found in section 2.3.

Numerous approximation relations of the frequency factor are published in the literature [5] for MOM estimation, the most significant being the Kite approximation, for $|C_s|<2$, the Cornish-Fisher approximation, for $|C_s|<2$, the Wilson-Hilferty approximation, for $|C_s|<2$, the modified Wilson-Hilferty approximation, for $0.25\leq C_s\leq 9.75$.

The frequency factor with MOM can also be approximated using a polynomial development in skewness (C_s):

$$K_{MOM}(p,C_s)=a+b\cdot C_s+c\cdot C_s^2+d\cdot C_s^3+e\cdot C_s^4+f\cdot C_s^5+g\cdot C_s^6+h\cdot C_s^7 \tag{10}$$

Table 1 lists the polynomial function coefficients for the annual exceedance probability that are utilized the most in technical hydrology.

Table 1

Coefficients of the approximation function with MOM

P [%]	a	b	c	d	e	f	g	h
0.01	3.71828	2.1462	0.15579	-0.0769315	0.0150378	-0.0017271	0.0001106	-0.00000303
0.1	3.09014	1.42629	0.049631	-0.0421189	0.00794983	-0.00083309	0.00004794	-0.00000118
0.5	2.57601	0.937811	-0.00485114	-0.024367	0.00459158	-0.0004292	0.00002047	-0.00000038
1	2.32661	0.733146	-0.0218707	-0.0185502	0.00358677	-0.00031539	0.00001302	-0.00000017
2	2.05408	0.533496	-0.034201	-0.0138703	0.00286305	-0.00023957	0.00000831	-0.000000042
3	1.88115	0.419782	-0.0389303	-0.0116643	0.00257668	-0.00021375	0.00000687	-0.000000006
5	1.64524	0.280836	-0.0418754	-0.0094549	0.00237315	-0.00020267	0.00000657	0.0000000005
10	1.28196	0.103328	-0.0395043	-0.0074825	0.00241382	-0.00023132	0.00000899	-0.00000008
20	0.842052	-0.052671	-0.027535	-0.0068667	0.002969	-0.00033372	0.00001445	-0.00000016
40	0.254237	-0.164334	0.0070463	-0.015678	0.0078439	-0.0013773	0.00010621	-0.00000308
50	0.0006921	-0.174131	0.019451	-0.018001	0.010156	-0.002096	0.00018921	-0.00000639
80	-0.845883	-0.010892	-0.041893	0.064938	-0.022096	0.0033839	-0.0002494	0.0000072

Given that the frequency factor is stated using the Gamma function, which can be challenging to calculate, an approximation relation based on the L-skewness and annual exceedance probability is presented. The approximation is the following polynomial relation:

$$K_{TL}(p, \tau_3) = a + b \cdot \tau_3 + c \cdot \tau_3^2 + d \cdot \tau_3^3 \quad (11)$$

In Table 2, the coefficients of the polynomial function for the most popular annual exceedance probabilities are shown.

Table 2

Coefficients of the approximation function with L-moments

P [%]	a	b	c	d
0.01	6.590	23.38	17.214	-3.7117
0.1	5.4765	15.559	8.986	0.47591
0.5	4.5651	10.245	4.4167	1.5525
1	4.1231	8.0174	2.8187	1.5366
2	3.6401	5.8441	1.4754	1.2797
3	3.3336	4.6063	0.81958	1.042
5	2.9154	3.094	0.14699	0.66702
10	2.2715	1.1625	-0.45319	0.08242
20	1.4918	-0.53214	-0.63128	-0.39305
40	0.44907	-1.699	-0.25238	-0.49031
50	0.0000044	-1.814	0.00423	-0.28014
80	-1.4918	-0.52533	0.62038	0.92798
90	-2.2715	1.1681	0.44733	1.14

The three parameters' values vary depending on the estimating technique employed. The relationships for estimating the parameters of the PE3 distribution described in the section that follows.

2.3. Parameter estimation

The parameter estimation is presented for MOM, L-moments and LH-moments, common methods in flood frequency analysis. The advantage of these methods is that they are characterized by statistical indicators (expected value, coefficient of variation, skewness, L-skewness, and LH-kurtosis) that can be determined regionally. Regarding the LH-moments method, only the relationships for the 1st order level are analyzed, because an alternative to the analysis using the Annual Exceedance Series (AES) is desired.

The distribution parameters' expressions for MOM estimation can be found in [5, 10]. Regarding the parameter estimation with L-moments, using the quantile function, the exact parameter estimate for the L-moment technique is carried out numerically (definite integrals). Using the Gaussian Quadrature method, the integrals are numerically determined. But, an approximate form of parameter estimation can be adopted, because the third L-moments (L_3) and L-skewness ($\tau_3 = L_3/L_2$), depends only on the shape coefficient [7,8].

Like the L-moments method, exact equations for estimation with LH-moments are obtained from solving a system of nonlinear equations using defined integrals. For this method, LH-skewness is also characterized only by the shape parameter. Thus, it was possible to obtain approximate relations for estimating this parameter, giving values for LH-skewness. Thus, for the LH–moments method, α can be approximate with the next relation:

If $0.12 < |\tau_{H3}| \leq 0.34$:

$$\alpha = exp\begin{pmatrix} 7757.0921831+40914.6033757 \cdot ln(|\tau_{H3}|) + \\ 93713.9484593 \cdot ln(|\tau_{H3}|)^2 + 121792.0331514 \cdot ln(|\tau_{H3}|)^3 + \\ 98255.1222272 \cdot ln(|\tau_{H3}|)^4 + 50397.8680523 \cdot ln(|\tau_{H3}|)^5 + \\ 16054.8135102 \cdot ln(|\tau_{H3}|)^6 + 2904.9945626 \cdot ln(|\tau_{H3}|)^7 + \\ 228.664592 \cdot ln(|\tau_{H3}|)^8 \end{pmatrix} \quad (12)$$

If $0.34 < |\tau_{H3}| \leq 0.85$:

$$\alpha = exp\begin{pmatrix} -13.4247904 - 121.5293664 \cdot ln(|\tau_{H3}|) - \\ 649.9763722 \cdot ln(|\tau_{H3}|)^2 - 2075.3170378 \cdot ln(|\tau_{H3}|)^3 - \\ 4110.4652507 \cdot ln(|\tau_{H3}|)^4 - 5114.9286399 \cdot ln(|\tau_{H3}|)^5 - \\ 3890.8525714 \cdot ln(|\tau_{H3}|)^6 - 1653.2523283 \cdot ln(|\tau_{H3}|)^7 - \\ 300.612615 \cdot ln(|\tau_{H3}|)^8 \end{pmatrix} \quad (13)$$

$$\beta = \frac{2 \cdot L_{H2}}{3 \cdot z_1} \quad (14)$$

$$\gamma = L_{H1} - 2 \cdot \beta \cdot z_2 \quad (15)$$

where, $z_1 = \int_0^1 qgamma(p, \alpha) \cdot (3 \cdot p^2 - 2 \cdot p) \cdot dp$, which can be approximated with the following equation:

$$z_1 = \frac{-0.00315255+0.87292281 \cdot \alpha + 0.18314623 \cdot \alpha^2}{\begin{pmatrix} 1 + 2.01526823 \cdot \alpha + 0.07089912 \cdot \alpha^2 - \\ 0.00034641 \cdot \alpha^3 + 0.00000094 \cdot \alpha^4 \end{pmatrix}} \quad (16)$$

and, $z_2 = \int_0^1 qgamma(p, \alpha) \cdot p \cdot dp$, which can be approximated with the following equation:

$$z_2 = \frac{\begin{pmatrix} 0.01180195 + 0.87724953 \cdot \alpha + 0.46798927 \cdot \alpha^2 + \\ 0.01808637 \cdot \alpha^3 + 0.00004649 \cdot \alpha^4 \end{pmatrix}}{1 + 0.80457526 \cdot \alpha + 0.03470298 \cdot \alpha^2 + 0.0000921 \cdot \alpha^3} \quad (17)$$

The MLE method is an easy method for estimating the parameters of a theoretical distribution, which consists in the logarithm and the derivation of the objective function, the latter being the product of the probability density function. By deriving the logarithmic objective function depending on the parameters of the theoretical distribution and minimizing them, the following system of a nonlinear equation results that leads to the determination of the position parameter [5,20].

The least squares method is a less used method because the estimation of the parameters is not robust, it can be used for an initial estimation of the parameters used as kernels for methods using the gradient method. It is a method that uses the cumulative function of the theoretical distribution [5,10,20].

3. Case Studies

As case studies, the frequency analysis of the maximum flows on 6 rivers (Ialomita, Siret, Bahna, Jijia, Nicolina and Viseu) with different morphometric and statistical characteristics is carried out.

Figure 1 shows the locations of the six monitoring stations, for the six studied rivers.

The analysis period varies, each monitoring station being characterized by a length of measurements greater than 20 years.

In the case of the analysis using the method of ordinary moments, the skewness coefficient is chosen according to the genesis of the flows, by multiplying the coefficient of variation of the analyzed data by a coefficientt (ξ), in order to reflect this genesis. The mathematical meaning of statistical indicators can be found in reference materials such as [7,10].

Fig. 1. Location of the six analyzed rivers.

The most important morphometric information regarding the rivers analyzed are highlighted in Table 3 [21].

Table 3

The morphometric elements for the analyzed rivers.

River	Length [km]	Average Stream Slope [‰]	Sinuosity Coefficient [−]	Average Altitude, [m]	Catchments Area, [km^2]
Ialomita	417	1.5	1.88	327	10350
Siret	559	1.7	1.86	515	42890
Jijia	275	1.0	1.45	152	5757
Bahna	35	28	1.45	559	137
Nicolina	20	16	1.37	138	177
Viseu	82	15	1.31	1011	1581

Tables 4 and 5 list the statistical indicator values for the examined data sets.

Table 4

The statistical indicators for the analyzed rivers: MOM

River	Number of records (n)	Hydrometric Station	MOM			
			ξ	μ	C_v	C_s
			[−]	[m^3/s]	[−]	[−]
Ialomita	33	Tandarei	2	224	0.527	0.33
Siret	39	Lungoci	2	1443	0.634	1.41
Jijia	35	Vladeni	3	56.1	0.824	1.85
Bahna	30	Bahna	3	13.3	1.519	3.11
Nicolina	39	Iasi	3	14.1	1.193	2.80
Viseu	20	Bistra	3	392	0.694	2.66

Table 5

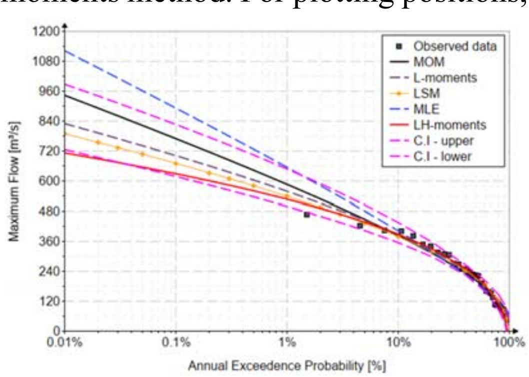
The statistical indicators for the analyzed rivers: L-and LH-moments method

River	L-moments method							LH-moments method						
	L_1	L_2	L_3	L_4	τ_2	τ_3	τ_4	L_{H1}	L_{H2}	L_{H3}	L_{H4}	τ_{H2}	τ_{H3}	τ_{H4}
	[m^3/s]	[m^3/s]	[m^3/s]	[m^3/s]	[−]	[−]	[−]	[m^3/s]	[m^3/s]	[m^3/s]	[m^3/s]	[−]	[−]	[−]
Ialomita	224	68.6	6.13	1.69	0.306	0.089	0.025	293	56.1	5.22	2.30	0.191	0.093	0.041
Siret	1443	490	112	90.6	0.339	0.228	0.185	1932	451	135	89.9	0.233	0.299	0.199
Jijia	56.1	23.2	7.86	6.01	0.414	0.338	0.259	79.4	23.3	9.25	6.13	0.294	0.396	0.263
Bahna	13.3	8.10	4.91	3.52	0.608	0.608	0.436	21.3	9.73	5.62	3.68	0.456	0.577	0.378
Nicolina	14.1	7.55	3.60	2.22	0.536	0.477	0.294	21.6	8.36	3.88	2.34	0.386	0.464	0.280
Viseu	392	121	63.5	49.3	0.309	0.525	0.407	513	138	75.2	53.0	0.270	0.543	0.383

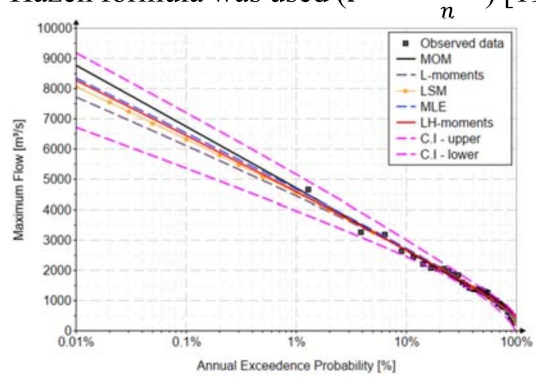
Results and Discussions

The analysis's goal is to assess how well the methods provided here perform in order to forecast the values of the quantiles corresponding to rare and very rare events.

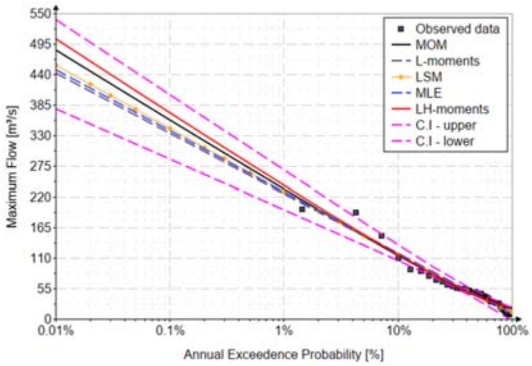
Considering that the quantile values are the ones of interest, the presented results will be based on this aspect. Figure 2 shows the results obtained on the 6 case studies, using the 5 analyzed estimation methods, as well as the confidence interval for the L-moments method. For plotting positions, the Hazen formula was used ($P = \frac{(i-0.5)}{n}$) [19].



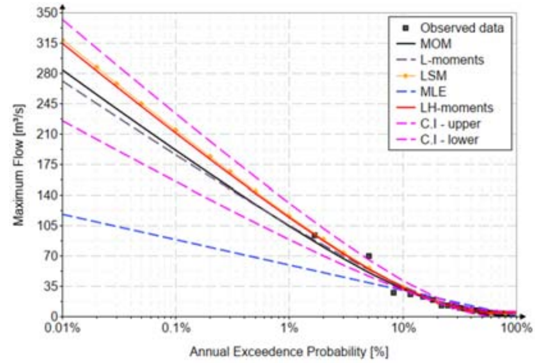
(a) Ialomita



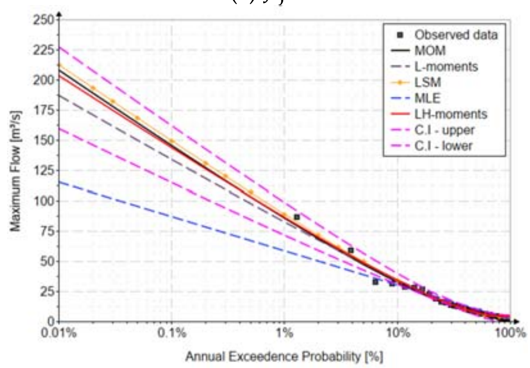
(b) Siret



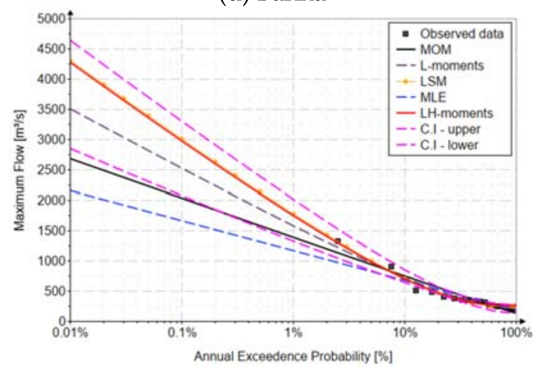
(c) Jijia



(d) Bahna



(e) Nicolina



(f) Viseu

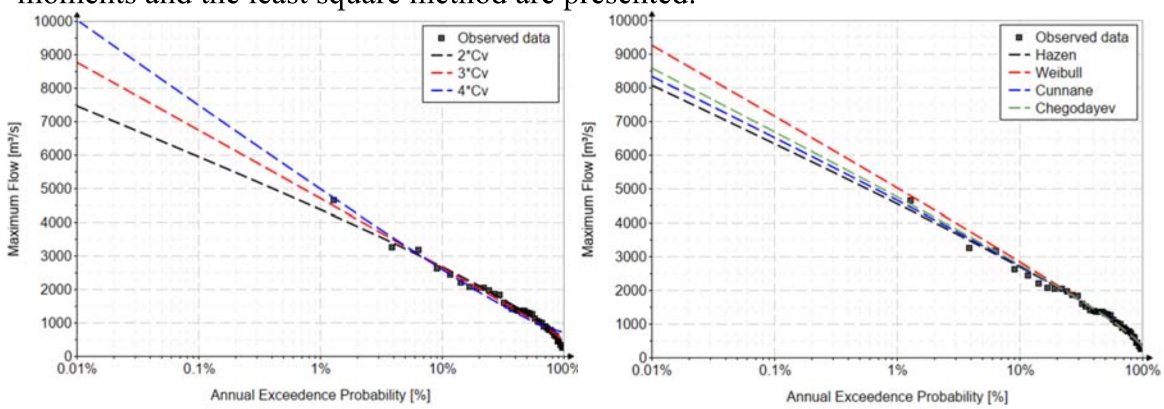
Fig. 2. Evaluations of the quantile function for the five methods of parameter estimation.

The confidence interval is built for L-moments, based on Chow's approximation [1, 5, 10] defined for a statistical distribution for 90% confidence level (10% significance level). This assumes that the confidence interval is a variable function of the probability and standard error specific to each statistical distribution [5, 19].

Analyzing the obtained results, it can be observed that for the data series with skewness greater than 2 (Nicolina, Viseu and Bahna), the maximum likelihood method has no solution, the generated quantile values being characterized by very large errors, the errors increasing togheter with the increase of the skewness value.

In the case of the method of ordinary moments, the resulting values are significantly influenced by the small length of the data series, an aspect that can also be observed in the case of the Viseu river (n=20 values). This aspect is due to the particularities of the method, the errors being directly proportional to the increase in the degree of the statistical indicators that need to be calibrated (variance and skewness). Also, the establishment of skewness based on the genesis of the flows (without a complex analysis regarding this genesis), implies a subjective nature of the analysis which is a disadvantage. These errors are accentuated when, for basins larger than 100 km^2, the simplified approach is usually used, considering the variation coefficient as 1 and the multiplication coefficient as 4. For comparison, the graph of the curves corresponding to the three values of the multiplication coefficient for the Siret river is presented. This subjective character also appears in the case of the least squares method, many researchers choosing the Weibull probability as the predefined empirical probability, which otherwise leads to significant deviations. The empirical probability must be established depending on the estimation method and the nature of the distribution used. Important contributions regarding this aspect have been made in recent materials [22], the Hazen empirical probability being the one that generates the smallest deviations from the theoretical values of the Pearson III distribution.

In the graphs of Figure 3, these particular aspects of the method of ordinary moments and the least square method are presented.



(a) MOM (b) LSM

Fig. 3. Comparative analysis for the Siret river.

Among the five analyzed methods, the L-moments method is the recommended one, being more stable and robust for short data lengths. It also represents, along with the LH-moments method, the only parameter estimation methods that present clear selection criteria, namely the calibration of higher-order linear moments, with approximate relationships and graphs of variation of these two indicators. The LH-moments method is also a method that, although it uses the series of maximum annual flows, it can be used as an alternative to the frequency analysis with partial series, fulfilling the same role, namely assigning a smaller weight to the lower extreme values (these are not always flood flows, but only maximum values corresponding to each year of analysis, representing the main disadvantage of using the block maximum/annual maximum method).

4. Conclusions

The Pearson III distribution is frequently used in hydrology, in Romania it is the parent distribution in flood frequency analysis.

In Romania, in most cases, the Pearson III distribution is used using tabular calculation developed only for MOM, with inadequate linear interpolation. Moreover, it is used inappropriately (for basins larger than 100 km^2). In recent materials [7,10,23], important contributions were made regarding the applicability of the Pearson III distribution, using the MOM and L-moments estimation methods.

Considering the results obtained on the 6 case studies, the following conclusions can be drawn:

- the method of linear moments is a much more stable, robust method and less sensitive to the variability of recorded data lengths, as well as to the presence of outliers. It is the only method that has clear criteria for selecting the best distribution, namely the calibration of higher order indicators, thus being able to make a pre-selection of the necessary distributions taking into account the values of the two indicators specific to the analyzed set. The same observations are also valid for the LH-moments method, which also has the advantage of the fact that it partially fulfills the so-called "separation effect" of the maximum flows from the annual series of maximum flows (comparable to the analysis of partial series);

- the Pearson III distribution cannot be applied, for skewness values greater than 2, using the maximum likelihood method of parameter estimation.

- the method of ordinary moments is recommended to be used only in the case of large data series (n>100 values), so that the correction of the skewness of the analyzed data set is minimal. The simplified approach of choosing skewness as 4 cannot be used for watershed larger than 100 km^2;

- the application of the Pearson III distribution using the least squares method is recommended to be performed only using the Hazen empirical probability, because after the comparative analysis it was observed that the biases compared to the theoretical values (population) are the smallest for this empirical probability.

These elements presented in the article are part of a wider research carried out in the Faculty of Hydrotechnics, in the elaboration of some proposals for the implementation of the L-moments method, in the future regulations regarding the analysis of extreme events in Romania and abandoning the old Soviet practices of use of the Pearson III distribution and some non-tehnical concepts without a mathematical basis such as the uncertainty interval [24].

References

[1] Bulletin 17B Guidelines for determining Flood Flow Frequency; Hydrology Subcommittee, Interagency Advisory Committee on Water Data, U.S. Department of the Interior, U.S. Geological Survey, office of Water Data Coordination, Reston, Virginia: 1981.
[2] Bulletin 17C Guidelines for determining Flood Flow Frequency; U.S. Department of the Interior, U.S. Geological Survey, Reston, Virginia: 2017.
[3] World Meteorological Organization. (WMO-No.100) 2018 Guide to Climatological Practices; WMO: Geneva, Switzerland, 2018.
[4] EM 1110-2-1415 Hydrologic Frequency Analysis, Engineering and Design; Department of the Army U.S. Army Corps of Engineers: 1993.
[5] Rao, A.R.; Hamed, K.H. Flood Frequency Analysis; CRC Press LLC: Boca Raton, FL, USA, 2000.
[6] Gubareva, T.S.; Gartsman, B.I. Estimating Distribution Parameters of Extreme Hydrometeorological Characteristics by L-Moment Method. Water Resour. 2010, 37, 437–445.
[7] Anghel, C.G.; Ilinca, C. Parameter Estimation for Some Probability Distributions Used in Hydrology. Appl. Sci. 2022, 12, 12588. https://doi.org/10.3390/app122412588.
[8] Hosking, J.R.M and Wallis, J.R. (1997) Regional Frequency Analysis: An Approach Based on L−moments. Cambridge University Press, UK, http://dx.doi.org/10.1017/cbo9780511529443.
[9] Hosking, J.R.M. L-moments: Analysis and Estimation of Distributions using Linear, Combinations of Order Statistics. J. R. Statist. Soc. 1990, 52, 105–124.

[10] Ilinca, C.; Anghel, C.G. Flood-Frequency Analysis for Dams in Romania. Water 2022, 14, 2884. https://doi.org/10.3390/w14182884
[11] Ciupak, M.; Ozga-Zieliński, B.; Tokarczyk, T.; Adamowski, J. A Probabilistic Model for Maximum Rainfall Frequency Analysis. Water 2021, 13, 2688. https://doi.org/10.3390/w13192688.
[12] Nwaogazie IL, Sam MG. Probability and non-probability rainfall intensity-duration-frequency modeling for port-harcourt metropolis, Nigeria. Int J Hydro. 2019;3(1):66-75. DOI: 10.15406/ijh.2019.03.00164.
[13] Ye, L., Hanson, L. S., Ding, P., Wang, D., and Vogel, R. M.: The probability distribution of daily precipitation at the point and catchment scales in the United States, Hydrol. Earth Syst. Sci., 22, 6519–6531, https://doi.org/10.5194/hess-22-6519-2018, 2018.
[14] Matalas N.C. Probability Distribution of Low Flows. Statistical Studies in Hydrology. Geological Survey, United States Covernment Printing Office, USA, Washington:1963.
[15] Greenwood, J.A.; Landwehr, J.M.; Matalas, N.C.; Wallis, J.R. Probability Weighted Moments: Definition and Relation to Parameters of Several Distributions Expressable in Inverse Form. Water Resour. Res. 1979, 15, 1049–1054, https://doi.org/10.1029/wr015i005p01049.
[16] Houghton, J.C. Birth of a parent: The Wakeby distribution for modeling flood flows. Water Resour. Res. 1978, 14, 1105–1109.
[17] Wang, Q.J. LH moments for statistical analysis of extreme events. Water Resour. Res. 1997, 33, 2841–2848.
[18] STAS 4068/1962; Maximum Water Discharges and Volumes, Determination of maximum Water Discharges and Volumes of Watercourses, the Romanian standardization institute, Bucharest, Romania.
[19] V.T.Chow, D.R.Maidment, L.W.Mays Applied Hydrology, 1988, MCGraw-Hill, Inc., ISBN 007-010810-2.
[20] Singh, V.P. Entropy-Based Parameter Estimation in Hydrology; Springer Science + Business Media: Dordrecht, The Netherlands.
[21] The Romanian Water Classification Atlas, Part I—Morpho−Hydrographic Data on the Surface Hydrographic Network; Ministry of the Environment: Bucharest, Romania, 1992.
[22] Anghel, C.G.; Stanca, S.C.; Ilinca, C. Two-Parameter Probability Distributions: Methods, Techniques and Comparative Analysis. Water 2023, 15, 3435. https://doi.org/10.3390/w15193435.
[23] Ilinca, C.; Anghel, C.G. Flood Frequency Analysis Using the Gamma Family Probability Distributions. Water 2023, 15, 1389. doi: 10.3390/w15071389.
[24] Drobot, R.; Draghia, A.F.; Chendes, V.; Sirbu, N.; Dinu, C. Consideratii privind viiturile sintetice pe Dunare. Hidrotehnica 2023, 68. (In Romanian).