The influence of the analyzed data lengths variability on the behavior of the GEV and Pearson III distributions

Influența variabilității lungimilor datelor analizate asupra comportamentului distribuțiilor GEV și Pearson III

Cristian Gabriel Anghel¹, Cosmin Craciun², Constantin Albert Titus³, Cornel Ilinca⁴

 ^{1,2,4}Technical University of Civil Engineering Bucharest Lacul Tei Boulevard Nr. 124, Bucharest 020396, Romania
³Politehnica University Timisoara, Romania
Piata Victoriei Nr. 2, Timisoara, Romania

¹E-mail: cristian.anghel@utcb.ro ²E-mail: cosmin.craciun@phd.utcb.ro ³E-mail: albert.constantin@upt.ro ⁴E-mail: cornel.ilinca@utcb.ro

DOI: 10.37789/rjce.2024.15.3.5

Abstract. This manuscript presents the influence of the variability of the recorded data series on the behavior and generation of quantile values for two of the most used statistical distributions in the frequency analysis of extreme events in hydrology, namely the generalized extreme value (GEV) distribution and the Pearson III distribution. The methods for estimating the analyzed parameters are the method of ordinary moments and the method of linear moments, which represent two of the most used methods for estimating the parameters of statistical distributions. According to the results obtained, the L-moments method represents a more stable and robust method characterized by much smaller biases than the ordinary moments method, for quantile values in the field of rare and very rare events (low and very low annual exceedance probabilities).

Key words: estimation parameters; method of ordinary moments; method of linear moments; Pearson III; GEV.

1. Introduction

The variability of the lengths of recorded data has an important role in determining the maximum extreme values in hydrology, especially in the frequency analysis of maximum flows and precipitation.

In general, the direct determination of these maximum values corresponding to the annual exceedance probabilities of interest in hydrology is done through frequency analysis [1]. This implies the use of certain statistical distributions, respectively certain parameter estimation methods [2-4].

Regarding the probability distributions, two of the most used distributions are the generalized extreme value (GEV) distribution and the Pearson III distribution [2-5]. In recent materials [4,6-12], important contributions have been made to these distributions as well as to a significantly large number of other distributions and families of distributions.

As parameter estimation methods, the method of ordinary moments (MOM) and L- moments methods are two of the most analyzed methods, having the advantage of being based on statistical indicators that can be determined regionally [5]. Otherwise, the L-moments method is the most used method in the processes of regionalization of extreme events.

Regarding these two distributions and parameter estimation methods, important contributions were made by Anghel and Ilinca [4,6-12] who made contributions regarding the approximate estimation of the parameters; relations and variation diagrams of higher order indicators for the L-moment method; expression of the inverse function using predefined functions in Excell and Mathcad; the expression of the inverse function with the frequency factor estimated with the L-moments method, as well as the approximation relations of these frequency factors (depending on τ_3) for the most common annual exceedance probabilities in the FFA.

Considering that in the frequency analysis it is desirable to obtain results characterized by a low degree of uncertainty, a determining role is played by the influence of the variability of the analyzed data lengths, knowing that small and medium data lengths can be characterized by relative errors that increase with the decrease of the annual probability of exceeding. The rarer the event, the greater the relative errors (bias). They also depend on the intrinsic characteristics of the analyzed distribution as well as on the parameter estimation method.

2. Methods and Materials

The analyzed estimation methods are MOM and L-moments. The analysis consists in highlighting the deviations due to the influence of the sizes of the analyzed data sets. Considering that the values of rare and very rare quantiles are of interest in the FFA, the analysis presents the maximum flow with the probability of exceeding equal to 0.01%, the determination of this value being mandatory for the verification of Dams type retention works of importance class 1 [13].

In general, this stage of highlighting the deviations from the theoretical curve is a subsequent stage of establishing the best distribution, thus we have the certainty that the data set analyzed comes from the respective distribution.

In the case of MOM, these deviations are presented for usual values of the coefficient of variation encountered in the analysis of extreme events in hydrology. The skewness coefficient is established by choosing the 3 multiplication coefficients depending on the genesis of the maximum flows according to Romanian practice [14].

The influence of the analyzed data lengths variability on the behavior of the GEV and Pearson III distributions

In the case of the L-moments method, the biases are presented for the entire matrix of the theoretical values of the τ_2 and τ_3 indicators (τ_3 being considered positive). The coefficient of L-variation (τ_2) always takes positive values between 0 and 1. The limits of L-skewness vary between $2 \cdot \tau_2 - 1 \le \tau_3 < 1$, and those of L-kurtosis between $\frac{1}{4} \cdot (5 \cdot \tau_3^2 - 1) \le \tau_4 < 1$ [5].

In general, L-skewness is considered positive, the same approach being present in this manuscript.

The method of determining the biases consists in choosing the theoretical values of the indicators specific to the methods, estimating the parameters of the distributions, and recalculating all these values through sampling. For simplicity, the arithmetic mean (expected value) is chosen as 1.

The Table 1 presents the most important relationships that characterize these two statistical distributions necessary for their use in FFA, such as the density function, the cumulative function, the inverse function (quantile function) [1-5].

The analyzed distributions

Table 1

| Distri. | Density function $f(x)$ | Complementary cumulative distribution function, $F(x)$ | Quantile function $x(p)$ |
|---------|--|---|--|
| PE3 | $\frac{(x-\gamma)^{\alpha-1}}{\beta^{\alpha}\cdot\Gamma(\alpha)}\cdot exp\left(-\frac{x-\gamma}{\beta}\right)$ $=\frac{1}{\beta}\cdot dgamma\left(\frac{x-\gamma}{\beta},\alpha\right)$ | $1 - \frac{1}{\beta \cdot \Gamma(\alpha)} \cdot \int_{\gamma}^{x} \left(\frac{x - \gamma}{\beta}\right)^{\alpha - 1} \cdot exp\left(-\frac{x - \gamma}{\beta}\right) dx$ $= \frac{\Gamma\left(\alpha, \frac{x - \gamma}{\beta}\right)}{\Gamma(\alpha)}$ | $\gamma + \beta \cdot qgamma(1 - p, \alpha)$ |
| GEV | $\frac{1}{\Gamma(\alpha)} \cdot \frac{ \beta }{ \theta } \cdot \left(\frac{x-\gamma}{\theta}\right)^{\alpha \cdot \beta - 1} \\ \cdot exp\left(-\left(\frac{x-\gamma}{\theta}\right)^{\beta}\right)$ | $1 - exp\left(-\left(1 - \frac{\alpha}{\beta} \cdot (x - \gamma)\right)^{\frac{1}{\alpha}}\right)$ | $\gamma + \frac{\beta}{\alpha} \cdot (1 - (-\ln(1-p))^{\alpha})$ |

The exact and approximate parameter estimation relationships, respectively the $\tau 3-\tau 4$ variation relationships are presented in [2,5,6,8]. The predefined functions in Mathcad are also equivalent in Excel, as was presented in other materials [8].

3. Results and Disscutions

The analysis is presented for the most common values of the higher order statistical indicators found in the FFA.

In the sampling process (after determining the parameters and implicitly the theoretical values of the inverse function of the distribution) the Hazen empirical probability is used, because it has been observed that it is claimed for these two

distributions and parameter estimation methods, on the grounds that in the sampling process the values on the three levels of bias (indicators, parameters and quantiles) should be characterized by the smallest deviations from the theoretical values. Deviations are expressed in percentages. A positive bias means that the calculated values (for the sample) are lower than the theoretical values (population) and require an increase with the resulting percentage. In the opposite case, for a negative bias, the values need to be reduced by the resulting percentage.

Thus, Table 2 shows the results obtained by applying the most used empirical probabilities in FFA, for the least interested events, namely the maximum value with the annual probability of exceeding 0.01%, 0.1%, 0.5% and 1%.

Table 2

| n=25 values; Cv=1; Cs=3*Cv | | | | | | | | | | | | |
|----------------------------|---------|-------|-------|---------|----------------------------------|-------|------------------|-------|--|--|--|--|
| Bias | Weibull | Hazen | Blom | Cunnane | Cunnane Adamowski Cegodaev Hirsh | | Landwehr/ APL | | | | | |
| Q _{0.01%} | 39.1% | 21.3% | 27.3% | 26.2% | 32.0% | 30.2% | 49.0% | 28.5% | | | | |
| Q _{0.1%} | 35.4% | 18.4% | 24.1% | 23.0% | 28.5% | 26.9% | 45.3% | 25.3% | | | | |
| Q _{0.5%} | 31.2% | 15.2% | 20.4% | 19.5% | 24.7% | 23.1% | 41.0% | 21.6% | | | | |
| Q _{1%} | 28.6% | 13.2% | 18.2% | 17.3% | 22.3% | 20.8% | 38.4% | 19.4% | | | | |

Empirical probability choice. Results for 25 values. Pearson III distribution.

Thus, after identifying the corresponding empirical probability, sampling is done (n=80, 50 and 25 values) recalculating each time the statistical indicators of the series, the distribution parameters and the quantile values.

For example, tables 3 and 4 show these biases on the 3 levels, for the usual values of C_V and τ_2 , for the Pearson III distribution.

The biases for Pearson III distribution: MOM.

Table 3

| | Cv=0.5; Cs | s=3*Cv | | $C_{v}=0.5; C_{s}=4*C_{v}$ | | | | |
|--------------------|------------|--------------|--------|----------------------------|--------|--------|--|--|
| Indicator | R | lecord lengt | h | Record length | | | | |
| mulcator | 80 | 50 | 25 | 80 | 50 | 25 | | |
| Cv | 1.00% | 1.40% | 1.10% | 1.60% | 2.20% | 1.90% | | |
| Cs | 10.7% | 14.1% | 20.5% | 12.2% | 15.7% | 22.2% | | |
| μ | 0.20% | 0.30% | 0.60% | 0.60% 0.20% | | 0.70% | | |
| a | -25.4% | -35.4% | -58.0% | -29.5% | -40.6% | -65% | | |
| b | 11.7% | 15.5% | 22.7% | 13.8% | 17.8% | 25.6% | | |
| γ -28.1% -42 | | -42.3% | -85.0% | -13.9% | -19.3% | -31.9% | | |
| Q _{0.01%} | 4.88% | 6.53% | 2.55% | 6.70% | 8.83% | 3.53% | | |

The influence of the analyzed data lengths variability on the behavior of the GEV and Pearson III distributions

| | Cv=0.5; Cs | s=3*Cv | Cv=0.5; Cs=4*Cv | | | | | |
|-------------------|------------|--------------|-----------------|---------------|-------|-------|--|--|
| Indicator | R | Record lengt | h | Record length | | | | |
| Inucator | 80 | 50 | 25 | 80 | 50 | 25 | | |
| Q _{0.1%} | 3.93% | 5.28% | 3.45% | 5.44% | 7.21% | 4.74% | | |
| Q _{0.5%} | 3.05% | 4.09% | 3.45% | 4.19% | 5.59% | 4.74% | | |
| Q _{1%} | 2.55% | 3.45% | 5.33% | 3.53% | 4.74% | 7.21% | | |

Table 4

| τ | ₂ =0.182; τ ₃ = | $0.1; \tau_4=0.12$ | .6 | $\tau_2=0.67; \tau_3=0.5; \tau_4=0.25$ | | | | |
|--------------------|---------------------------------------|--------------------|--------|--|--------|--------|--|--|
| Indicator | F | Record lengt | h | Record length | | | | |
| | 80 | 50 | 25 | 80 | 50 | 25 | | |
| $	au_2$ | 4.95% | 4.40% | 0.60% | -0.75% | -1.19% | -1.70% | | |
| $	au_3$ | -0.09% | -0.14% | -0.27% | -1.05% | -1.65% | -3.22% | | |
| $	au_4$ | -2.38% | -3.97% | -6.35% | -1.60% | -2.80% | -5.20% | | |
| а | 0.19% | 0.28% | 0.56% | 2.84% | 2.84% | 7.58% | | |
| b | -1.06% | -1.06% | -3.19% | -1.60% | -2.53% | -5.01% | | |
| γ | 43.8% | 55.0% | 70.0% | 50.0% | 50.0% | 75.0% | | |
| Q _{0.01%} | -0.43% | -0.70% | -1.33% | -0.99% | -1.58% | -3.10% | | |
| Q _{0.1%} | -0.40% | -0.63% | -1.21% | -0.81% | -1.30% | -2.55% | | |
| Q _{0.5%} | -0.36% | -0.56% | -1.02% | -0.62% | -0.97% | -1.90% | | |
| Q _{1%} | -0.32% | -0.49% | -0.92% | -0.48% | -0.77% | -1.50% | | |

The biases for Pearson III distribution: L-moments.

Taking into account that in practice we can meet various regimes, with different variability, Tables 5 and 6 show the biases of the Pearson III distribution for MOM, respectively for the L-moments method for the entire matrix of statistical indicators (for the rare event $Q_{0.01\%}$). It can be seen that in the case of MOM, for medium and large data variabilities and skewness, the biases are significant. In the case of the L-moments method, the biases are very small. In table 6, the highlighted areas represent the usual range of values recorded in the frequency analysis of extreme events in hydrology.

Cristian Gabriel Anghel, Cosmin Craciun, Constantin Albert Titus, Cornel Ilinca

Table 5

| | 0.01% annual exceedance probability | | | | | | | | | | | | |
|---|-------------------------------------|------|------|-------|--------|----------|-------|-------|-------|-------|--|--|--|
| e | | | | | | Cv | | | | | | | |
| ς | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 | 1.1 | 1.3 | 1.5 | 1.7 | 2 | | | |
| | Records, n=80 | | | | | | | | | | | | |
| 2 | 0.33 | 1.67 | 3.22 | 4.87 | 6.56 | 8.31 | 10.1 | 11.9 | 13.71 | 16.4 | | | |
| 3 | 0.42 | 2.5 | 4.88 | 7.48 | 10.2 | 12.9 | 15.6 | 18.3 | 20.9 | 24.8 | | | |
| 4 | 0.62 | 3.35 | 6.7 | 10.3 | 13.9 | 17.6 | 21.1 | 24.5 | 27.8 | 32.5 | | | |
| | | | | | Record | ds, n=50 | | | | | | | |
| 2 | 0.46 | 2.31 | 4.37 | 6.51 | 8.69 | 11.0 | 13.2 | 15.4 | 17.7 | 21.0 | | | |
| 3 | 0.63 | 3.42 | 6.53 | 9.84 | 13.3 | 16.7 | 20.0 | 23.3 | 26.5 | 31.1 | | | |
| 4 | 0.89 | 4.51 | 8.83 | 13.4 | 17.9 | 22.4 | 26.7 | 30.8 | 34.7 | 40.3 | | | |
| | | | | | Record | ds, n=25 | | | | | | | |
| 2 | 0.71 | 3.58 | 6.69 | 9.76 | 12.8 | 15.85 | 18.98 | 21.97 | 24.98 | 29.38 | | | |
| 3 | 1.04 | 5.22 | 9.74 | 14.4 | 19.1 | 23.6 | 28.1 | 32.4 | 36.6 | 42.4 | | | |
| 4 | 1.37 | 6.84 | 12.9 | 19.17 | 25.3 | 31.2 | 36.8 | 42.0 | 47.0 | 53.8 | | | |

The biases for Pearson III distribution: MOM. Extended fields of statistical indicators.

Table 6

The biases for Pearson III distribution: L-moments. Extended fields of statistical indicators.

| | | 0. | 01% ann | ual excee | dance pro | obability | y | | | | | | |
|-----|---------------|-------|---------|-----------|-----------|-----------|-------|-------|-------|--|--|--|--|
| | | | | | τ_2 | | | | | | | | |
| τ3 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | | | | |
| | Records, n=80 | | | | | | | | | | | | |
| 0 | -0.35 | -0.49 | -0.53 | -0.58 | -0.61 | -0.62 | -0.63 | -0.64 | -0.65 | | | | |
| 0.1 | -0.352 | -0.5 | -0.54 | -0.56 | -0.61 | -0.63 | -0.64 | -0.65 | -0.66 | | | | |
| 0.2 | -0.41 | -0.53 | -0.59 | -0.62 | -0.63 | -0.66 | -0.67 | -0.67 | -0.69 | | | | |
| 0.3 | -0.5 | -0.62 | -0.68 | -0.72 | -0.72 | -0.74 | -0.76 | -0.77 | -0.77 | | | | |
| 0.4 | -0.61 | -0.75 | -0.8 | -0.82 | -0.85 | -0.86 | -0.87 | -0.88 | -0.88 | | | | |
| 0.5 | -0.75 | -0.87 | -0.93 | -0.96 | -0.97 | -0.99 | -1.0 | -1.01 | -1.01 | | | | |
| 0.6 | -0.92 | -1.05 | -1.1 | -1.13 | -1.15 | -1.16 | -1.17 | | | | | | |
| 0.7 | -1.2 | -1.33 | -1.38 | -1.4 | -1.42 | -1.42 | -1.43 | -1.44 | -1.45 | | | | |
| 0.8 | -1.72 | -1.85 | -1.89 | -1.92 | -1.94 | -1.95 | -1.96 | -1.96 | -1.96 | | | | |
| 0.9 | -3.28 | -3.42 | -3.46 | -3.49 | -3.5 | -3.52 | -3.52 | -3.53 | -3.53 | | | | |
| | | | | Records | s, n=50 | | | | | | | | |
| 0 | -0.55 | -0.74 | -0.85 | -0.9 | -0.94 | -0.96 | -0.98 | -1.0 | -1.01 | | | | |
| 0.1 | -0.55 | -0.74 | -0.83 | -0.91 | -0.96 | -0.98 | -0.99 | -1.01 | -1.02 | | | | |
| 0.2 | -0.64 | -0.83 | -0.92 | -0.97 | -0.99 | -1.03 | -1.05 | -1.05 | -1.07 | | | | |
| 0.3 | -0.77 | -0.97 | -1.06 | -1.12 | -1.13 | -1.16 | -1.18 | -1.19 | -1.2 | | | | |
| 0.4 | -1.0 | -1.16 | -1.26 | -1.29 | -1.33 | -1.35 | -1.37 | -1.37 | -1.39 | | | | |
| 0.5 | -1.2 | -1.39 | -1.48 | -1.52 | -1.55 | -1.57 | -1.58 | -1.6 | -1.61 | | | | |

| | 0.01% annual exceedance probability | | | | | | | | | | | | | | |
|---------------|-------------------------------------|----------------|-------|-------|-------|-------|-------|-------|-------|--|--|--|--|--|--|
| _ | | τ ₂ | | | | | | | | | | | | | |
| τ3 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | | | | | | |
| 0.6 | -1.5 | -1.68 | -1.77 | -1.82 | -1.85 | -1.86 | -1.88 | -1.89 | -1.9 | | | | | | |
| 0.7 | -1.96 | -2.17 | -2.25 | -2.28 | -2.31 | -2.33 | -2.34 | -2.35 | -2.36 | | | | | | |
| 0.8 | -2.86 | -3.08 | -3.16 | -3.2 | -3.22 | -3.24 | -3.25 | -3.26 | -3.27 | | | | | | |
| 0.9 | -5.66 | -5.89 | -5.98 | -6.02 | -6.05 | -6.07 | -6.08 | -6.09 | -6.09 | | | | | | |
| Records, n=25 | | | | | | | | | | | | | | | |
| 0 | -1.05 | -1.41 | -1.59 | -1.7 | -1.78 | -1.84 | -1.88 | -1.91 | -1.94 | | | | | | |
| 0.1 | -1.04 | -1.42 | -1.61 | -1.71 | -1.8 | -1.85 | -1.89 | -1.92 | -1.95 | | | | | | |
| 0.2 | -1.2 | -1.56 | 1.72 | -1.82 | -1.88 | -1.94 | -1.97 | -1.99 | -2.02 | | | | | | |
| 0.3 | -1.46 | -1.83 | -2.0 | -2.09 | -2.14 | -2.18 | -2.22 | -2.24 | -2.26 | | | | | | |
| 0.4 | -1.84 | -2.22 | -2.4 | -2.48 | -2.55 | -2.58 | -2.61 | -2.63 | -2.66 | | | | | | |
| 0.5 | -2.33 | -2.74 | -2.89 | -2.98 | -3.04 | -3.08 | -3.11 | -3.13 | -3.15 | | | | | | |
| 0.6 | -2.99 | -3.4 | -3.57 | -3.65 | -3.71 | -3.75 | -3.78 | -3.8 | -3.82 | | | | | | |
| 0.7 | -4.05 | -4.48 | -4.65 | -7.73 | -4.79 | -4.82 | -4.85 | -4.87 | -4.88 | | | | | | |
| 0.8 | -6.18 | -6.65 | -6.82 | -6.91 | -6.96 | -7.0 | -7.03 | -7.05 | -7.06 | | | | | | |
| 0.9 | -13.4 | -13.9 | -14.1 | -14.2 | -14.3 | -14.3 | -14.4 | -14.4 | -14.4 | | | | | | |

The influence of the analyzed data lengths variability on the behavior of the GEV and Pearson III distributions

Tables 7 and 8 show the results obtained with the GEV distribution. It can be seen that the influence of the variability of the recorded data is much more pronounced than in the case of the Pearson III distribution, due to the fact that this is a heavy tail distribution. Even in this case, the L-moments method gives better results, the biases being much smaller than MOM.

Table 7

The biases for GEV distribution: MOM. Extended fields of statistical indicators.

| | 0.01% annual exceedance probability | | | | | | | | | | | | | |
|---|-------------------------------------|------|------|------|----------|------|------|------|------|------|--|--|--|--|
| u | | Cv | | | | | | | | | | | | |
| 5 | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 | 1.1 | 1.3 | 1.5 | 1.7 | 2 | | | | |
| | Records, n=80 | | | | | | | | | | | | | |
| 2 | -0.012 | 1.74 | 5.71 | 10.9 | 16.2 | 21.0 | 25 | 28.4 | 31.1 | 34.4 | | | | |
| 3 | 0.15 | 4 | 11.5 | 19.0 | 25.2 | 29.9 | 33.4 | 36.0 | 38.1 | 40.5 | | | | |
| 4 | 0.36 | 7 | 17.1 | 25.4 | 31.1 | 35.1 | 38 | 40.1 | 41.8 | 43.6 | | | | |
| | | | | | Records, | n=50 | | | | | | | | |
| 2 | 0.02 | 2.5 | 7.7 | 13.9 | 20.0 | 25.3 | 29.6 | 33.2 | 36.1 | 39.4 | | | | |
| 3 | 0.3 | 5.4 | 14.5 | 23.1 | 29.7 | 34.6 | 38.3 | 41.0 | 43.2 | 45.6 | | | | |
| 4 | 0.6 | 9.0 | 20.9 | 29.9 | 35.9 | 40.1 | 43.0 | 45.1 | 46.8 | 48.6 | | | | |
| | | | | | Records, | n=25 | | | | | | | | |
| 2 | 0.1 | 4.1 | 11.4 | 19.4 | 26.7 | 32.6 | 37.3 | 41.1 | 44.1 | 47.5 | | | | |
| 3 | 0.5 | 8.3 | 20.0 | 30.0 | 37.3 | 42.5 | 46.2 | 49.0 | 51.2 | 53.6 | | | | |
| 4 | 1.0 | 13.0 | 27.3 | 37.3 | 43.7 | 47.9 | 50.9 | 53.0 | 54.7 | 56.5 | | | | |

Table 8

| The biases for GEV distribution: L-moments. Extended fields of sta | tatistical indicators. |
|--|------------------------|
|--|------------------------|

| | | | 0.01% a | nual av | coodona | nrohoh | ility | | | | | |
|---------------|-------|-------|------------|-----------|------------------|----------|-------|-------|-------|--|--|--|
| | | | J.UI 70 al | iniual ex | +2 | e probab | шцу | | | | | |
| t3 | 0.1 | 0.2 | 0.2 | 0.4 | 0.5 | 0.6 | 07 | 0.0 | 0.0 | | | |
| | 0.1 | 0.2 | 0.5 | 0.4 D | 0.5 | | 0.7 | 0.0 | 0.9 | | | |
| 0 | 0.44 | 0.66 | 0.70 | Recoi | $\frac{1}{0.04}$ | 0.08 | 1.02 | 1.05 | 1.07 | | | |
| 0 1 | -0.44 | -0.00 | -0.79 | -0.87 | -0.94 | -0.98 | -1.02 | -1.03 | -1.07 | | | |
| 0.1 | -0.57 | -0.78 | -0.89 | -1.01 | -1.01 | -1.04 | -1.07 | -1.09 | -1.1 | | | |
| 0.2 | 0.09 | 0.11 | 0.12 | 0.12 | 0.13 | 0.13 | 0.13 | 0.13 | 0.14 | | | |
| 0.3 | 3.02 | 3.48 | 3.66 | 3.76 | 3.82 | 3.86 | 3.89 | 3.91 | 3.93 | | | |
| 0.4 | 9.62 | 10.41 | 10.7 | 10.85 | 10.95 | 11.01 | 11.05 | 11.09 | 11.12 | | | |
| 0.5 | 19.97 | 20.85 | 21.17 | 21.33 | 21.43 | 21.49 | 21.5 | 21.57 | 21.6 | | | |
| 0.6 | 32.9 | 33.72 | 34 | 34.15 | 34.23 | 34.3 | 34.33 | 34.36 | 34.39 | | | |
| 0.7 | 46.7 | 47.4 | 47.6 | 47.8 | 47.8 | 47.9 | 47.9 | 47.9 | 48.0 | | | |
| 0.8 | 59.6 | 60.3 | 60.5 | 60.6 | 60.6 | 60.7 | 60.7 | 60.7 | 60.8 | | | |
| 0.9 | 70.4 | 71.1 | 71.4 | 71.5 | 71.6 | 71.6 | 71.7 | 71.7 | 71.7 | | | |
| Records, n=50 | | | | | | | | | | | | |
| 0 | -0.68 | -1.01 | -1.21 | -1.35 | -1.44 | -1.51 | -1.57 | -1.6 | -1.64 | | | |
| 0.1 | -0.86 | -1.18 | -1.35 | -1.45 | -1.52 | -1.57 | -1.61 | -1.64 | -1.66 | | | |
| 0.2 | 0.05 | 0.07 | 0.07 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | | | |
| 0.3 | 3.87 | 4.48 | 4.7 | 4.83 | 4.91 | 4.96 | 5.0 | 5.03 | 5.05 | | | |
| 0.4 | 12.0 | 12.97 | 13.34 | 13.53 | 13.64 | 13.72 | 13.78 | 13.82 | 13.85 | | | |
| 0.5 | 23.9 | 24.96 | 25.34 | 25.53 | 25.64 | 25.72 | 25.78 | 25.82 | 25.85 | | | |
| 0.6 | 37.85 | 38.79 | 39.12 | 39.28 | 39.38 | 39.44 | 39.49 | 39.53 | 39.56 | | | |
| 0.7 | 51.8 | 52.6 | 52.9 | 53.0 | 53.1 | 53.2 | 53.2 | 53.2 | 53.2 | | | |
| 0.8 | 64.3 | 65.0 | 65.2 | 65.3 | 65.4 | 65.4 | 65.5 | 65.5 | 65.5 | | | |
| 0.9 | 74.2 | 75.0 | 75.2 | 75.4 | 75.4 | 75.5 | 75.5 | 75.6 | 75.6 | | | |
| | | 1 | 1 | Recor | ds, n=25 | 5 | 1 | | | | | |
| 0 | -1.27 | -1.89 | -2.26 | -2.51 | -2.69 | -2.82 | -2.92 | -3 | -3.07 | | | |
| 0.1 | -1.54 | -2.12 | -2.42 | -2.61 | -2.73 | -2.82 | -2.89 | -2.95 | -2.99 | | | |
| 0.2 | -0.09 | -0.12 | -0.13 | -0.13 | -0.14 | -0.14 | -0.14 | -0.14 | -0.15 | | | |
| 0.3 | 5.45 | 6.27 | 6.6 | 6.77 | 6.88 | 6.95 | 7.01 | 7.05 | 7.08 | | | |
| 0.4 | 16.14 | 17.45 | 17.9 | 18.2 | 18.35 | 18.46 | 18.54 | 18.59 | 18.64 | | | |
| 0.5 | 30.4 | 31.75 | 32.22 | 32.46 | 32.61 | 32.71 | 32.78 | 32.84 | 32.88 | | | |
| 0.6 | 45.58 | 46.72 | 47.11 | 47.31 | 47.43 | 47.51 | 47.56 | 47.61 | 47.64 | | | |
| 0.7 | 59.5 | 60.4 | 60.7 | 60.9 | 61.0 | 61.0 | 61.1 | 61.1 | 61.1 | | | |
| 0.8 | 71.0 | 71.8 | 72.0 | 72.2 | 72.2 | 72.3 | 72.3 | 72.3 | 72.4 | | | |
| 0.9 | 79.6 | 80.4 | 80.6 | 80.8 | 80.9 | 80.9 | 81.0 | 81.0 | 81.0 | | | |

For usual values of the coefficient of variation, skewness, coefficient of L-variation and L-skewness, Figure 1 shows the curves of the inverse functions, at different sampling values (n=25,50,80) compared to the theoretical values (n>1000),

The influence of the analyzed data lengths variability on the behavior of the GEV and Pearson III distributions for both distributions and both parameter estimation methods.



(c) PE3, MOM (μ =1; C_v=1.5; C_s=3) (d) PE3, L-moments (L₁=1; τ_2 =0.6; τ_3 =0.5) Fig. 1. Evaluations of the quantile function.

4. Conclusions

The frequency analysis has an important role in the direct determination of the maximum flows, corresponding to the low annual probabilities of exceeding.

Considering that in many cases a small number of records are available, the probability distributions are, depending on the parameter estimation method, more or less influenced by this variability of the recorded data lengths.

Considering that the Pearson III and generalized extreme value (GEV) distributions are two of the most used statistical distributions in FFA, the main objective of the manuscript was to highlight the behavior of these distributions and the biases of the inverse function at different river regimes, highlighted by the theoretical values of the higher order statistical indicators.

According to the obtained results, of the two parameter estimation methods (for the two distributions), the least affected by this variability is the L-moments method, the resulting biases on the entire theoretical definition matrix of the statistical indicators being much smaller than in the case of the method of ordinary moments. Of the two analyzed distributions, the GEV distribution is more affected, due to its intrinsic characteristic of being a heavy-tail distribution.

Thus, taking into account this deviation from the theoretical values, it is very important that the confidence interval of the distribution chosen as the best model is presented in the FFA. An accessible solution is the representation of the confidence interval using Chow's approximate relationship [6,15], an advantage being also represented by the fact that in recent materials it has also been adapted to the L-moments method [7-12]. It is thus desired to avoid the use of non-technical concepts such as the uncertainty interval [16].

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] Rao, A.R.; Hamed, K.H. Flood Frequency Analysis; CRC Press LLC: Boca Raton, FL, USA, 2000.

[3] Gubareva, T.S.; Gartsman, B.I. Estimating Distribution Parameters of Extreme Hydrometeorological Characteristics by L-Moment Method. Water Resour. 2010, 37, 437–445.

[4] 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.

[5] 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.

[6] Ilinca, C.; Anghel, C.G. Flood-Frequency Analysis for Dams in Romania. Water 2022, 14, 2884. https://doi.org/10.3390/w14182884.

[7] 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.

[8] Ilinca, C.; Anghel, C.G. Flood Frequency Analysis Using the Gamma Family Probability Distributions. Water 2023, 15, 1389. doi: 10.3390/w15071389.

[9] Ilinca, C.; Anghel, C.G. Frequency Analysis of Extreme Events Using the Univariate Beta Family Probability Distributions. Appl. Sci. 2023, 13, 4640. doi: 10.3390/app13074640.

[10] Anghel, C.G.; Ilinca, C. Evaluation of Various Generalized Pareto Probability Distributions for Flood Frequency Analysis. Water 2023, 15, 1557. doi: 10.3390/w15081557.

[11] Anghel, C.G.; Ilinca, C. Predicting Flood Frequency with the LH-Moments Method: A Case Study of Prigor River, Romania. Water 2023, 15, 2077. doi: 10.3390/w15112077.

[12] Anghel, C.G.; Stanca, S.C.; Ilinca, C. Extreme Events Analysis Using LH-Moments Method and Quantile Function Family. Hydrology 2023, 10, 159. doi: 10.3390/hydrology10080159.

[13] STAS 4068/2-87; Annual probabilities of maximum flows and volumes under normal and special operating conditions. The Romanian Standardization Institute: Bucharest, Romania, 1987.

[14] STAS 4068/1962; Maximum Water Discharges and Volumes, Determination of maximum Water Discharges and Volumes of Watercourses, the Romanian standardization institute, Bucharest, Romania.

[15] V.T.Chow, D.R.Maidment, L.W.Mays Applied Hydrology, 1988, MCGraw-Hill, Inc., ISBN 007-010810-2.

[16] Drobot, R.; Draghia, A.F.; Chendes, V.; Sirbu, N.; Dinu, C. Consideratii privind viiturile sintetice pe Dunare. Hidrotehnica 2023, 68. (In Romanian).