L-MOMENTS AND THEIR USE IN MAXIMUM DISCHARGES' ANALYSIS IN CURVATURE CARPATHIANS REGION

L. ZAHARIA [1]

ABSTRACT. – L-moments and their use in maximum discharges' analysis in Curvature Carpathians region. This paper aims to present L-moments statistics and their application in analysis of annual maximum discharges series in Curvature Carpathians region, in order: 1) to determine the discordance measure and 2) to identify the regional theoretical distributions most appropriate to the data series in the studied region. The analysis performed allowed: 1) to determine two gauging stations (Mircești on Putna River and Moara Domnească on Teleajen River), whose data are discordant comparing with those from the other 41 analyzed stations and 2) to identify three theoretical distributions best fitting maximum annual discharges series: exponential, generalised Pareto and Pearson type 3 distributions.

Keywords: L-moments, probability distribution, discordance, maximum discharge, Curvature Carpathian Region.

1. INTRODUCTION

Maximum discharges' analysis is particularly important for theoretical purposes and, especially, for practical needs (river engineering works, flood risk management). Frequency analysis is among the methods widely employed to study maximum discharges. It allows estimating the probability or frequency of an event occurrence in the future, as random variable, based on data analysis from previous events and on a distribution low describing the statistical behaviour of the considered variable (Meylan et al. 2008). Frequency analysis is based on statistical treatments of data series defining various phenomena. During the last decades, among statistical parameters employed in frequency analysis, one can found L-moments. In hydrology, they have been used successfully in regional frequency analysis (Hosking and Wallis, 1997; Haddad and Rahan, 2012).

The purpose of this paper is to synthetically present L-moments (statistical parameters relatively less known and employed in hydrological practices in Romania) and to use them for the analysis of maximum discharges data series in Curvature Carpathians Region, in order: 1) to determine the discordant catchments compared with whole region they belong and 2) to identify regional theoretical distributions well explaining the variability of maximum discharge in the studied area.

[1] University of Bucharest, Faculty of Geography, 1. N. Balcescu Blvd., Sector 1, 010041 Bucharest, Romania, e-mail: zaharialili@hotmail.com

2. STUDY AREA AND DATA

The **study area** overlaps mostly Carpathians' Curvature region (especially the external side) and includes three distinct morphological units: a mountainous area (belonging to the Carpathians) to the West, a hilly area (belonging to the Subcarpathians) in the center, and a plain area (north-eastern part of the Romanian Plain) to the East (Fig. 1).

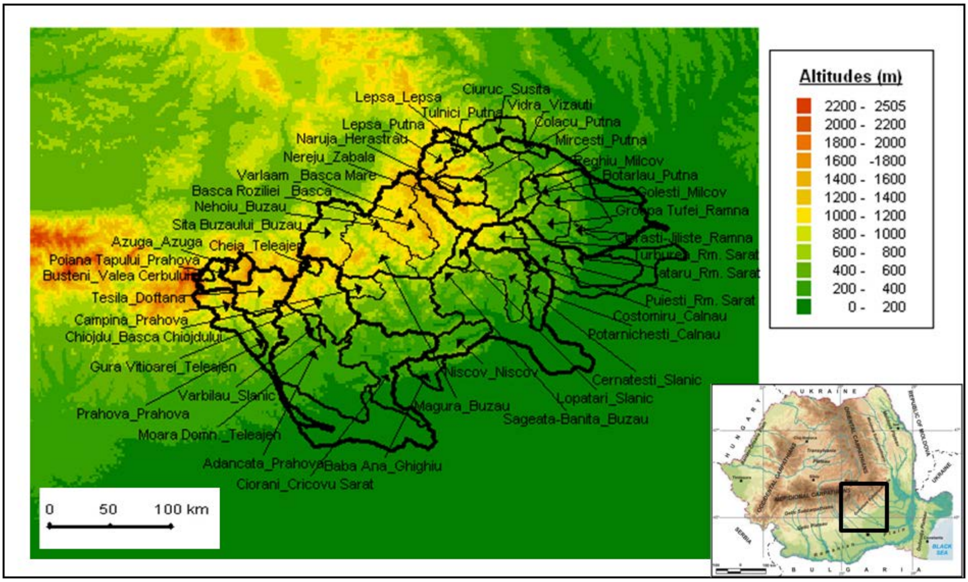

Fig. 1. ***Location of the study yrea and the analysed catchments***

The mountainous zone includes almost one third of the study area; the altitudes exceed 1600-1800 m.a.s.l. in eastern and central parts, and 2000 m.a.s.l. in the western part, reaching 2505 m (in Bucegi Mountains). The hilly area overlaps about 40% of the study area; here the altitudes are frequently of 600-800 m.a.s.l. and exceed locally 900-1000 m a.s.l. In the eastern part, a glacis connects the hills to the plain and the altitude decreases gradually from 300-350 m a.s.l. (at the contact with the Subcapathians) to less than 50 m (in the subsidence plain).

The liquid flow is relatively moderate, mean specific annual discharges varying between 10-16 l/s/km^2 in the mountains area, 3-12 l/s/km^2 in the hills and less than 7 l/s/km^2 in the plain. High flows characterise spring and summer: 41-43%, respectively 24-30% of the mean annual volume. Low flows are specific in winter (about 15% for the mountains area) and in autumn (14-15% for the hills and for plain). Maximum annual discharges and floods occur mostly in summer and spring (Zaharia, 2005a).

Data. This paper relies on maximum annual discharges from 43 gauging stations located on the main rivers crossing the study area (data from the National Institute of Hydrology and Water Management – N.I.H.W.M.). The catchments areas corresponding to these stations oscillate between 25 km^2 (Valea Cerbului

River at Buşteni) and 3992 km^2 (Buzău River at Săgeata-Baniţa), and the mean altitudes, between 93 m (Coţatcu River at Martineşti) and 1429 m (Valea Cerbului River at Buşteni) (Zaharia, 2005b).

Series length varies between 13 and 40 years. For dammed rivers, whose flow is influenced by reservoirs, only the period prior to these engineering works was considered. It is the case of Buzău and Teleajen rivers, having reservoir dams at Siriu, respectively Măneciu. The influence of the dams on the maximum discharges was proven using graphical and statistical tests (Wilcoxon's test).

3. METHODES

The study is essentially based on statistical methods. They were employed, firstly, to estimate L-moments. Afterwards, L-moments were used in other statistical analysis, which aimed at determining the discordance and at identifying regional theoretical distributions.

3.1. L-moments

L-moments are a relatively recent development within statistics. They are alternative statistics to conventional moments for describing data samples and probability distributions (Hosking and Wallis, 1997). *L*-moments were derived from Probability Weighted Moments, as defined by J. A. Greenwood et al. 1979, quoted by Hosking and Wallis (1993).

L-moments are linear combinations of Probability Weighted Moments (hence the prefix *L*) (Hosking and Wallis (1993). Conforming Hosking (1990), L-moments can be defined „for any random variable whose mean exists and form the basis of a general theory which covers the summarization and description of theoretical probability distributions, the summarization and description of observed data samples, estimation of parameters and quantiles of probability distributions, and hypothesis tests for probability distributions" (page 105).

For a probability distribution with cumulative distribution function F(x), Probability Weighted Moments (β_r) are defined by (http://researcher.watson.ibm.com/researcher/view_project.php?id=1023):

$$\beta_r = \int x \{F(x)\}^r \, dF(x), \qquad r = 0, 1, 2, \ldots .$$

The first few *L*-moments are (Hosking 1986, quoted by Zrinji and Burn, 1994) :

$$\lambda_1 = \beta_0 ,$$
$$\lambda_2 = 2\beta_1 - \beta_0 ,$$
$$\lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0 ,$$
$$\lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0 .$$

The first *L*-moment (λ_1) is the mean of the distribution (parameter of location; L-location). It can take any value. The second *L*-moment (λ_2) is a measure of the dispersion (analogue with standard deviation) and a scale parameter (L-scale). It has positive values ($\lambda_2 \geq 0$) (Hosking and Wallis, 1997).

By dividing the higher-order *L*-moments (λr) by the dispersion measure (λ_2), **L-moment ratios (τ_r)** are obtained:

$$\tau_r = \lambda_r / \lambda_2, \quad r = 3, 4,...$$

L-moment ratios are dimensionless quantities. $\boldsymbol{\tau_3}$ is a measure of skewness (*L*-skewness) and $\boldsymbol{\tau_4}$ is a measure of kurtosis (*L*-kurtosis). They have ordinarily values between -1 and +1. The ratio $\tau = \lambda_2/\lambda_1$ corresponds to ***L*-CV**, and is analogous of the conventional coefficient of variation. For distributions taking only positive values, $0 \leq \tau < 1$(Hosking, and Wallis, 1997).

Compared with the conventional moments, L –moments have the advantage of being less sensitive to the effects of sampling variability. They are most robust to outliers in the data and better adapted to small samples (Hosking, 1990).

In the study area, L-CV, τ_3 and τ_4 were computed using a Matlab routine. The Probability Weighted Moments ($\widehat{b}_j$) were determined for each hydrometric station with Hosking's estimator (Meylan and Musy, 1999) :

$$\widehat{b}_j = \frac{1}{n}\sum_{r=1}^{n} x_{[r]}\left[\frac{r-0.35}{n}\right]^j$$

where *n* is the sample size and *r* is the rank of the *i* observation in the series sorted in ascending order.

3.2. Discordance measure

Regional frequency analysis demands a preliminary examination of input data in order to identify and to eliminate those being very different from their group. For this purpose, Hosking and Wallis (1993, 1997) recommend to determine the **discordance.** This characteristic may be measured using L-moments ratio (τ, τ_3 and τ_4).

According to the above mentioned authors, the value of the discordance in a given site (Di) can be estimated with the equation:

$$D_i = \frac{1}{3}(u_i - \bar{u})^T S^{-1}(u_i - \bar{u}),$$

where $u_i = [\tau^{(i)}\ \tau_3^{(i)}\ \tau_4^{(i)}]^T$ is the vector containing the values of τ, τ_3 et τ_4 for a i site;

$$\bar{u} = N^{-1}\sum_{i=1}^{N} u_i \quad \text{(N= number of sites in the region) ;}$$

S is the covariance matrix, defined by:

$$S = (N-1)^{-1}\sum_{i=1}^{N}(u_i - \bar{u})(u_i - \bar{u})^T . \qquad (6.17.).$$

If $D_i \geq 3$, the site can be considered discordant relative to the whole group.

The discordance for all hydrometric stations was computed using a Matlab routine.

3.3. Identifying regional theoretical distributions

The performance of a method aiming at regionalising some parameters (in this study, maximum annual discharges) depends on theoretical distribution's adequate choice. In order to identify the regional distribution, one can use *L*-moment ratio diagram. It allows a visual indication of which distribution may be expected to give a good fit to data samples in a region by analyzing the different distributions and data samples plotted on the *L*-kurtosis (τ_4) - *L*-skewness (τ_3) diagram (Fig. 2).

To intuit the theoretical distribution describing the variability low of maximum annual discharges in the studied region, there were represented on **τ_4 -τ_3 diagram** the couples of values τ_4 and τ_3 (calculated based on data from 41 stations located in the study area; 2 stations, identified to be discordant, were eliminated) and 6 distributions frequently used in hydrological approaches**:** lognormal, Generalized extreme–value (GEV), Gumbel, Pearson type 3, exponential and generalized Pareto (Fig. 2).

On τ_4 -τ_3 – diagram, the distributions defined by 2 parameters are represented by a point:

- exponential distribution: $\tau_3 = 1/3$ and $\tau_4 = 1/6$;
- Gumbel distribution: $\tau_3 = 0{,}1699$ and $\tau_4 = 0{,}1504$.

For lognormal and Pearson type 3 distributions, polynomial approximations from below were used (Stedinger and Vogel, 1993, quoted by Meylan and Musy, 1999):

- log normal distribution: $\tau_4 = 0.12282 + 0.77518\,\tau_3^2 + 0.12279\,\tau_3^4 - 0.13638\,\tau_3^6 + 0.11368\,\tau_3^8$;
- Pearson type 3 distribution: $\tau_4 = 0.1224 + 0.30115\,\tau_3^2 + 0.95812\,\tau_3^4 - 0.57488\,\tau_3^6 + 0.19383\,\tau_3^8$.

For generalised Pareto distribution and GEV, there L-moments ratios were computed based on a form coefficient (c) (Meylan and Musy, 1999):

- generalized Pareto distribution: $\tau_3 = (1-c)/(3+c)$; $\tau_4 = (1-c)(2-c)/(3+c)(4+c)$;

- GEV distribution: $\tau_3 = \frac{2(1-3^{-c})}{1-2^{-c}} - 3$;

$$\tau_4 = \frac{1-5(4^{-c})+10(3^{-c})-6(2^{-c})}{1-2^{-c}}.$$

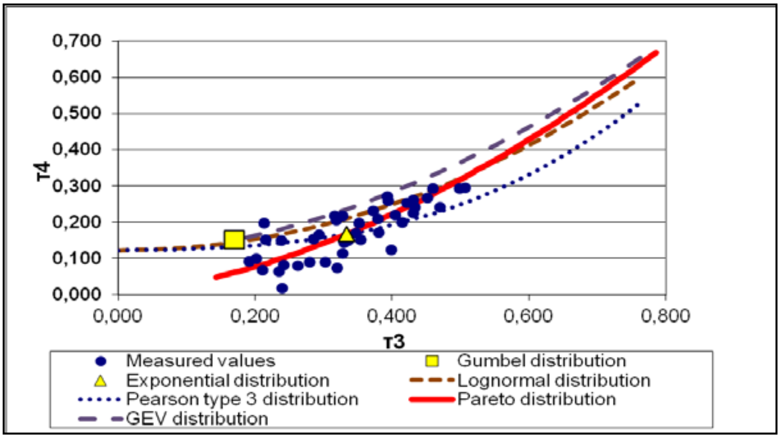

Fig. 2. ***τ_4 -τ_3 diagram with some theoretical distributions and moment ratios of the recorded data in Curvature Carpathian region***

4. RESULTS

Using the methods described above, L-moments and discordance values were determined, and the distributions with a good fit to annual maximum flow series in the study area were identified.

4.1. L-moments analysis

In studied region, L-CV (τ) oscillates from 0.32 to 0.59, τ_3, from 0.16 to 0.51, and τ_4, from 0.02 to 0.29. Diagrams indicate a direct dependence of the 2 L-moment ratios, but not an evident relation between L-moments values and morphometrical parameters of the catchment (area and mean altitude). Large catchments have low value of τ_3, contrasting to small catchment (Fig 3A).

On diagrams in figure 3 A, one can identify three catchments different from the others, by high values of τ_4 and low values of τ_3. It is the example of Mirceşti station on Putna River, Moara Domnească on Teleajen River, and Prahova on Prahova River. In the first case, Putna's flow is influenced by infiltration in the permeable bed and, thus, important water lost (Zaharia, 1999). In the other two cases, stations' deviation may be explained by human influences on the flow (water supply and irrigations).

Diagrams τ_3 – L-CV show a certain direct dependence between the 2 parameters. Large catchments have, generally, lower values of τ_3 and L-CV, than smaller catchments (Fig. B). High mean altitude catchments have lower L-CV.

4.2. Discordance analysis

Subsequently to calculus, two stations were identified as having values of discordance exceeding the critical value of 3 (the threshold of discordance): Mirceşti on Putna River (D = 3.64) and Moara Domnească on Teleajen River (D = 3.10). They have low values of τ_3 and high values of τ_4. The discordance of the two stations comparing with others from this region was shown also earlier by their position on $\tau_4 - \tau_3$ (Fig. 3A).

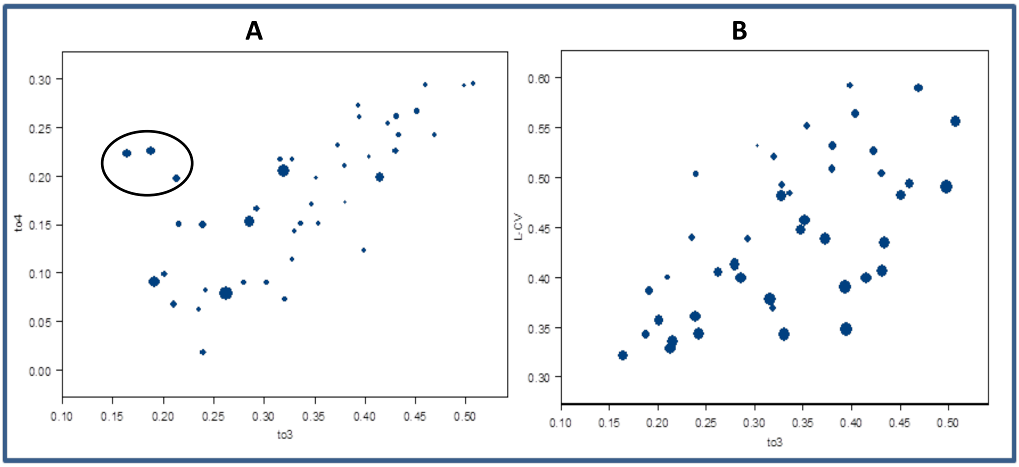

Fig. 3. ***A.*** ***τ_4 -τ_3 - diagram in relation with the catchment areas. B. L-CV - τ_3 diagram in relation with catchment areas (the size of the points is proportional with the catchment area). The circled points in diagram A detach from the rest.***

4.3. Identifying regional theoretical distributions

The analysis of the figure2 shows that exponential distributions and Generalized Pareto fit better the observed values compared to other distributions figured on the charts. A relatively good fitting has also the Pearson type 3 distribution. This distribution is generally used in Romania for estimating the discharges with different excedeence probabilities.

5. CONCLUSIONS

L-moments form the basis of a relatively new theory, which covers: the summarization and description of theoretical probability distributions and observed data samples; estimation of parameters and quartiles of probability distributions, and hypothesis tests for probability distributions (Hosking, 1990). They are alternative statistics to ordinary moments for describing data samples and probability distributions. Due to their advantages over these, L-moments are considered superior to conventional moments. They have proved very useful to facilitate the estimation process in regional frequency analysis.

After a synthetic description of L-moments, this paper presents their use to determine the discordance and identify the distributions fitting the best the series of maximum annual discharge in the Curvature Carpathian region. The identified distribution are: exponential, generalized Pareto, and Pearson type 3.The last distribution is in fact widely used in hydrological practice in Romania for estimating the maximum discharge exceeding probabilities.

Acknowledgements. The author thanks to Markus Niggli for developing the Matlab routine, and to N.I.H.W.M for providing the hydrological data.

REFERENCES

1. Haddad, K., Rahman, A. (2012), *Regional flood frequency analysis in eastern Australia: Bayesian GLS regression-based methods within fixed region and ROI framework – Quantile Regression vs. Parameter Regression Technique*, J. of Hyd. 430 – 431, 142-161.
2. Hosking, J. R. M. (1990), *L-moments: Analysis and estimation of distributions using linear combinations of order statistics*, J.R. Statist. Soc., 52, 1, 105-124.
3. Hosking J.R.M. & Wallis J.R., 1993, *Some statistics useful in regional frequency analysis,* Water Ressour. Res., 29 (2), p. 271-281.
4. Hosking, J. R. M., Wallis, J. R. (1997), *Regional frequency analysis: an approach based on L-moments.* Cambridge University Press, Cambridge, U.K.
5. Meylan, P., Musy, A., 1999, H.G.A. Bucarest.
6. Meylan, P., Favre A.- C., Musy, A. (2008), *Hydrologie fréquentielle. Une science prédictive*, Presses Polytechniques et Universitaires Romandes, Lausanne.
7. Zaharia, L. (2005a), *Studiul resurselor de apă din Carpaţii şi Subcarpaţii Curburii în vederea optimizării valorificării lor pentru alimentarea populaţiei din judeţele aferente regiunii* în „Lucrări şi rapoarte de cercetare. Centrul de cercetare „*Degradarea terenurilor şi Dinamică geomorfologică*”, I, 137 – 171.
8. Zaharia, L. (2005 b), *La matrice de corrélations entre les paramètres hydrologiques et les caractéristiques géographiques des bassins versants*, Comunicări de geografie, IX, 287-292.
9. Zaharia, L. (1999), *Resursele de apă din bazinul râului Putna. Studiu de hidrologie,* Ed. Univ. din Bucuresti.
10. Zrinji, Z., Burn, D.H. (1994), *Flood frequency analysis for ungauged sites using a region of influence approch*, J. of Hyd., 153, p. 1-21.
11. http://researcher.watson.ibm.com/researcher/view_project.php?id=1023