Flow Rate Determination as a Function of Rainfall for the Ungauged Suhareka River

Laura Kusari[1], Lavdim Osmanaj[1*], Hana Shehu[1], Samir Bungu[1]

[1] Faculty of Civil Engineering, University of Pristina, Kosovo
* Corresponding author's e-mail: lavdim.osmanaj@uni-pr.edu

ABSTRACT

For ungauged rivers, when there are no hydrological measurements and there is a lack of data on perennial flow rates, the latter one to be determined based on other hydrological data. The river Suhareka catchment represents a similar case. Since there is no data on Suhareka's flow rates, the authors of this study aimed for the flow rate determination based on rainfall measurements. From the available data on annual precipitation (monthly sums) provided by the Kosovo Hydrometeorological Institute for the Suhareka hydrometric station, the observed monthly rainfall data for 30 years were analysed. Those gaps were initially filled by connecting the hydrometric station in Suhareka with those of Prishtina, Prizren and Ferizaj, and as a result a fairly good fit was ensured. Moreover, the intensity-duration-frequency curves were formed using the expression of Sokolovsky, as a mathematical model of the dependence $I\,(T, P)$. For a transformation of rainfall into flow, the American method SCS was used. As a result, the equation for the Suhareka River basin was derived, which enabled the determination of maximum inflows, for different return periods. The results obtained through this paper, indicates that even for ungauged river basins the peak flows can be determined from available rainfall data.

Keywords: catchment area, regression coefficient, CN parameter, rainfall intensity, flow curves, rainfall floods.

INTRODUCTION

For the water management process, the data on river's flow rates it of utmost importance. Therefore, the river flow rate determination is a focus of every hydrological research. Runoff plots are very important means to monitor runoff and soil loss (Baoyan et al., 2017). Unfortunately, many catchments are ungauged, and thus there are limitations for flood calculation using rainfall-runoff models (Nam & Shin, 2018). As stated, hydrological modeling is instrumental for both scientific application and for providing public services (Kolbjorn & Alfredsen, 2020).

When there is a lack of data on the river flow rates and many parameters of catchment properties are missing, then another aproach must be considered. The measurement model states in the form of equations, the relationship between the measurements and the true value of interest (McMillan et al., 2018; Nearing et al., 2016). Due to the lack of flow rate measurements, the

relationship between rainfall and flow can help solve this defficiency. The flow forecasting as investigated in this paper, can rely solely on using the available rainfall data.

The Suhareka muncipality is on the sothern part of Kosovo and is characterised with the continental climate with the Mediterenean impact. The Suhareka river, flowing through this region is a small river but with higher flow oscilations through out the year, meaning that the Q_{max}/Q_{min} flow relation is quite high. The Suhareka River is also known as a dry river, but since 2015 it has been flooding the downstream areas about 3 times and damaging the agriculture and industrial activity in its vicinity. Therefore, the integrated flood management is a priority in this area and it requries the knowledge of the probabilistic flow duration curves. On the basis of the hydrologic data availibility, there are some methods that can be used in this regard. The rain intensity (i) ore rain height (H) dependence on duration time (tk) and return period (P), can be determined in

a pluviograph station that has a long observation period, as a final product of statistical analyses of rain (KHMI, 1984.). As for the pluvial floods, these are dynamic processes influenced by rain intensity and its distribution, catchments area, river flow density, catchment land use, soil geologic structure, soil moisture content, soil infiltration rate, etc. For the rainfall-runoff relationship identification, the most common method is the SCS method that is based on the dynamic interaction between rain intensity, soil infiltration and surface runoff.

FLOW TREND DYNAMICS AND DISTRIBUTION FUNCTIONS FOR EXTREME RAINFALL

The available rainfall data, for a 30 year time period, have some data deficiency during this time; therefore, the statistical population parameters will be evaluated on the basis of the representative group. As it is known, the larger the representative group, the smaller will be the errors in the population parametrs estimation and vice versa. From the population, we have the yearly rainfall series for 80 years, including those years when the data are missing. The representative group chosen is a 30 year series (Table 1).

In order to evaluate the arithmetic mean of population, of the representative group, the

standard deviation of the population (SDP) (Bektesh B, 2005) has to be found first, as follows:

$$SDP = \sqrt{\sigma^2 \times \frac{n}{n-1}} \qquad (1)$$

where: σ – standard deviation of the representative group; n – number of cases in group.

On the basis of the equation (1) it can be said that:

$$SDP = \sqrt{97.16218^2 \times \frac{30}{30-1}} = 98.82$$

While standard deviation of the arithmetic mean (*SDXavg*) will be:

$$SDXavg = \frac{SDP}{\sqrt{n}} \qquad (2)$$

respectively,

$$SDXavg = \frac{98.82}{\sqrt{30}} = 18.04$$

The results enable the finding of the boundary interval of the arithmetic mean of the population. As per the obtained results, for the probability coefficients 95%, the boundary interval is calculated as follows:

$$794.833 \pm (1.96 \times 18.04)$$

$$794.833 - (1.96 \times 18.04) = 759.47$$
$$\text{and}$$
$$794.833 + (1.96 \times 18.04) = 830.19$$

With the probability as high as 95%, it was found out that the mean arithmetic of population is within the range 759.47 mm and 830.19 mm.

Standard error of the standard deviation (*SEoSD*) (Bektesh, 2005) is:

$$SEoSD = \frac{SD}{\sqrt{2n}} \qquad (3)$$

and

$$SEoSD = \frac{98.82}{\sqrt{2 \times 30}} = \frac{98.82}{7.746} = 12.75$$

Boundary intervals for the 95% probability is as following:

$$98.82 \pm (1.96 \times 12.75)$$

$$98.82 - (1.96 \times 12.75) = 73.83$$
$$\text{and}$$
$$98.82 + (1.96 \times 12.75) = 123.81$$

In theory and scientific research, trends are calculated as is regression, with the following formula:

Table 1. Representative group of the rainfall population (KHMI, 1984)

| Year	P_{year} (mm)	Year	P_{year} (mm)
1954/55	931	1970/71	831
1955/56	778	1971/72	616
1956/57	648	1972/73	960
1957/58	786	1973/74	782
1958/59	719	1974/75	714
1959/60	745	1975/76	709
1960/61	781	1976/77	849
1961/62	760	1977/78	999
1962/63	969	1978/79	876
1963/64	933	1979/80	804
1964/65	716	1980/81	894
1965/66	818	1981/82	712
1966/67	773	1982/83	720
1967/68	669	1983/84	823
1968/69	751	Xmes. =	794.8333
1969/70	779	σ=	97.16218

Annual Rainfall Trend
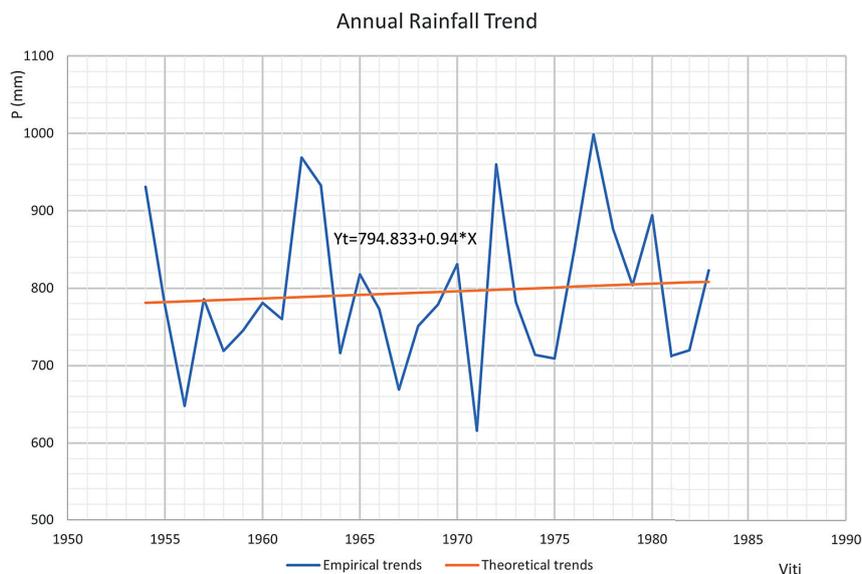
Figure 1. Annual series trend of the representative group

$$y = a + bx \qquad (4)$$

Parameter (*a*) shows the mean value of the time series, while parameter (*b*) shows the mean of the phenomena that is a study object, while (*x*) are stanadard values given for each year (Figure 1).

$$a = \frac{\sum Y}{n} \qquad (5)$$

$$b = \frac{\sum XY}{\sum X^2} \qquad (6)$$

Regression and correlation analyses

In scientific research, the regression analyses is a very useful method of identifying the correlation between two or more variables or phenomena. The correlation between two variables is a bivariate correlation, while the correlation between three or more is known as a complex or multivariate correlation.

When ther is no linear function with the needed correlation coefficient of the (*y*) as dependent variable, from a single (x) as an independend variable, then a possible correlation could be tested through multi linear regression. At double linear regression, the results can be stated as a line $y = b_0 + b_1 x$, where b_0 is dependent from (z) (Maniak, 2010). Thus, a relationship *y* = *y*(*x_i*; *z_i*) is obtained:

$$y(xi; zi) = b_0 + b_{1x} + b_{2z} \qquad (7)$$

To determine the b_o, b_1 and b_2 coefficients, the sum of all small quadrates S should be kept in minimum:

$$S = \sum_{i=1}^{N} [yi - y(xi; zi)]^2 =$$

$$= \sum_{i=1}^{N} [yi - \binom{bo + b1 \times xi +}{+ b2 \times zi}]^2 \rightarrow min \qquad (8)$$

By partaill derivation of this equation, through b_o, b_1 and b_2 coefficients, the 3x3 equation system is obtained:

$$\frac{\partial S}{\partial bo} = 0 \rightarrow \sum_{i=1}^{N} \begin{pmatrix} yi - bo - \\ -b1 \times xi - \\ -b2 \times zi \end{pmatrix} (-1) = 0$$

$$\frac{\partial S}{\partial b1} = 0 \rightarrow \sum_{i=1}^{N} \begin{pmatrix} yi - bo - \\ -b1 \times xi - \\ -b2 \times zi \end{pmatrix} (-xi) = 0 \quad (9)$$

$$\frac{\partial S}{\partial b2} = 0 \rightarrow \sum_{i=1}^{N} \begin{pmatrix} yi - bo - \\ -b1 \times xi - \\ -b2 \times zi \end{pmatrix} (-zi) = 0$$

Knowing that and the relevant expressions for Δxi and Δzi, following some transformations, the needed coefficients are the following:

$$bo = ym - b1 \times xm - b2 \times zm \qquad (10)$$

$$b1 = \frac{\sum \Delta xi \Delta yi \times \sum \Delta zi^2 - }{\sum \Delta xi^2 \times \sum \Delta zi^2 - (\sum \Delta xi \Delta zi)^2} \qquad (11)$$

$$b2 = \frac{\sum \Delta yi \Delta zi \times \sum \Delta xi^2 - }{\sum \Delta xi^2 \times \sum \Delta zi^2 - (\sum \Delta xi \Delta zi)^2} \quad (12)$$

The degree of determination is $B = r_{yxz}^2$ and is defined as follows:

$$r_{yxz}^2 = B = \frac{s_y^2(xi, zi)}{s_y^2} = $$
$$= \frac{\frac{\sum_{i=1}^{N}(y(xi, zi) - ym.)^2}{N-1}}{\frac{\sum_{i=1}^{N}(yi - ym.)^2}{N-1}} \quad (13)$$

with:

$$s_y^2(xi, zi) = \frac{\sum_{i=1}^{N}\left(\begin{matrix} b1 \times xi - \\ -b2 \times zi \end{matrix}\right)^2}{N-1} = $$
$$= \left[\begin{matrix} b1 \times \sum \Delta xi \Delta yi + \\ +b2 \times \sum \Delta yi \Delta zi \end{matrix}\right] /(N-1)$$

Since $-1 \le r \le 1$, it implies that $0 \le B \le 1$.

In the considered case, the triple regression was obtained, connecting the hydrometric station of Suhareka (y) with those of Prizren (x), Ferizaj (z) and Prishtina (t) (Maniak, 2010). From these analyses of variance, an important, but not sufficient correlation was obtained (Table 2, Figure 2).

According to this analysis, the regression equation will be:

$$y = 455.2839 + 0.3215x + \\ + 0.2394z - 0.1634t$$

The inadequacy of this relationship is also noticed by the high values of t-test and low ones of *F*-test as well as *P*-value which should be below the 5% probability level (<0.05). However, the cause of the failure of a linear regression is usually the nonlinearity between the variables.

Non linear regression and transformations

The cause of failure of a linear regression is mainly nonlinearity between variables (Maniak, 2010; Husno, 2007). When using a nonlinear regression, the curve function is often unknown (Karakuş, 2020). If the multiple nonlinear regressions are taken into account, then:

$$y = a_0 * x_1^{a1} x_2^{a2} x_3^{a3} \dots x_n^{an} \quad (14)$$

Figure 2. Normal probability plot
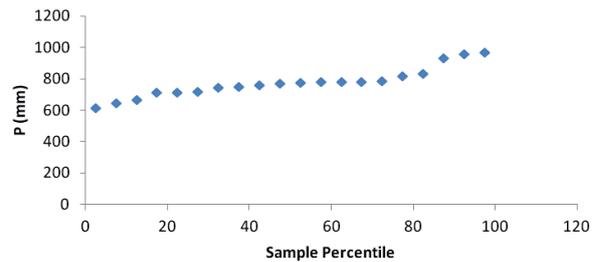
Table 2. Summary output of the multiple regression analyses

Summary output							
Regression statistics							
Multiple *R*				0.449555			
R Square				0.2021			
Adjusted *R* Square				0.052494			
Standard Error				91.1738			
Observations				20			
ANOVA							
Parameters		*df*	SS		MS	F	Significance *F*
Regression		3	33688.216		11229.41	1.35088	0.293297
Residual		16	133002.58		8312.661		
Total		19	166690.8				
Parameters	Coefficients	Standard Error	*t* Stat	*P*-value	Loëer 95%	Upper 95%	Loëer 95.0%	Upper 95.0%
Intercept	455.2839	177.71444	2.561884	0.020895	78.54608	832.0217	78.54608	832.0217
690	0.321528	0.2162926	1.486541	0.156575	-0.13699	0.780048	-0.136992	0.780048
757	0.239476	0.2685362	0.891782	0.385725	-0.3298	0.808747	-0.329796	0.808747
695	-0.16346	0.3472157	-0.47076	0.644163	-0.89952	0.572608	-0.899521	0.572608

which can be linearized using the following transformations:

$$\log y = \log a_0 + a_1 \log x_1 + $$
$$+ a_2 \log x_2 + \cdots + a_n \log x_n$$

On the basis of what was said above, similarly as in the case of linear regression, the Suhareka hydrometric station was connected with all three other stations. The nonlinear triple regression equation has the form $y = a_0 * x_1 a^1 * x_2 a^2 * x_3 a^3$ or $\log y = \log a_0 + a_1 * \log x_1 + a_2 * \log x_2 + a_3 * \log x_3$, where after determining the coefficients a_0, a_1, a_2 and a_3 the same takes the form:

$$\log y = 0.3057 + 0.384 * \log x_1 + $$
$$+0.358 * \log x_2 + 0.172 * \log x_3$$

or after anti-logarithm:

$$y = 2.02 * x_1^{0.384} * x_2^{0.358} * x_3^{0.172}$$

The results of these calculations are shown in the Figure 3.

Distribution functions for extreme rainfall

At a first glance, it seems that between individual cases that produce mass phenomena, there is an irregularity and chaos, but when they are studied scientifically, it is seen that there are genuine regularity, principles and laws. These rules, principles and laws are best revealed by the law of large numbers of Laplace and Gauss.

The normal distribution is a symmetric two-parametric distribution with density function:

$$f(x) = \frac{1}{\sigma * \sqrt{2\pi}} * exp - \frac{1}{2} * \left[\frac{x - xm.}{\sigma}\right]^2,$$
$$-\infty < x < +\infty \tag{15}$$

For standardization, take the standard variable $k = (x - x_m) / \Sigma$ and for $x_m = 0$, $\sigma = 1$ we have:

$$f(k) = \frac{1}{\sqrt{2\pi}} * \exp - \left[\frac{k^2}{2}\right] \approx 0.4 * e^{-\left(\frac{k^2}{2}\right)},$$
$$-\infty < k < +\infty \tag{16}$$

The normal distribution function for calculating the stagnation probability is:

$$P(X \le x) = \frac{1}{\sigma * \sqrt{2\pi}} *$$
$$* \int_{-\infty}^{x} exp - \frac{1}{2}\left[\frac{x - xm.}{\sigma}\right]^2 dx \tag{17}$$

The normal distribution is represented by the surface under the density function, which can be formed by the area $x_m \pm k\sigma$ arranged symmetrically with the center. This area $x_m \pm \sigma$ contains 68.26% of all cases.

Normal distribution symmetry is used to represent the distribution as right in a suitable probability diagram. The straight line is easily determined by points, i.e the average x_m in P_u = 50% and the values $x_m \pm \sigma$, which are 84.13% and 15.87% (or 1/6 of all values). With a linear division of the axes, the distribution function P (x)
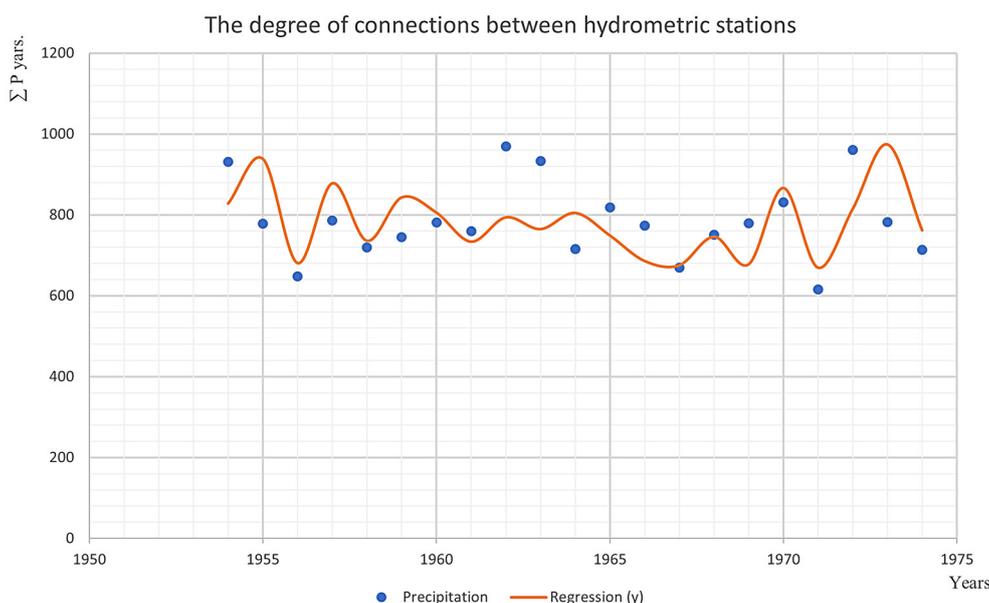

Figure 3. Degree of connections between hydrometric stations – triple nonlinear regression

is an S-shaped curve. Taking an annual series of 38-year rainfall and dividing it into classes with width $b = 50$ mm, the statistical parameters of normal distribution are obtained:

$$xm = \frac{\sum_{i=1}^{N} ni * xi}{N} =$$

$$= \frac{29050}{38} = 764.47 \text{ mm}$$

and

$$\sigma = \sqrt{\frac{\sum ni * (xi - xm.)^2}{N-1}} =$$

$$= \sqrt{\frac{630789.47}{38-1}} = 130.569 \text{ mm}$$

From here, $Cv = \dfrac{\sigma}{xm} = \dfrac{130.569}{764.47} = 0.1707$ while due to symmetry $Cs = 0$.

The normal distribution function takes the following values:

For $P_u = 50\% \rightarrow x = x_m = 764.47$ mm; for $P_u = 84.13\% \rightarrow x = x_m + \sigma = 895$ mm; and for $P_u = 15.87\% \rightarrow x = x_m - \sigma = 633.9$ mm. The Table 3 presents these calculations for different probability factors (Figure 4).

Whereas, the density function takes these values: for $x = x_m \rightarrow fo = 1 / (\sigma * \sqrt{2\pi}) * b * exp -1/2 * (0) = 50 / (130.569 * \sqrt{2\pi}) = 0.1527$, similarly for the cases $x = x_m \pm \sigma; x = x_m \pm 2\sigma; x = x_m \pm 3\sigma$ are obtained

Table 3. Normal distribution values of the distribution function

Pu=50% :	$x=x_{mes} =$	764.4736842	Mm
Pu=84.13%	$x=x_{mes}+s_x =$	895.0430687	Mm
Pu=15.87%	$x=x_{mes}-s_x =$	633.9042998	Mm
Pu=95%:	$x=x_{mes}+1.96s_x=$	1020.389678	Mm
Pu=99%:	$x=x_{mes}+2.58s_x=$	1101.342696	Mm
Pu=5% :	$x=x_{mes}-1.96s_x=$	508.5576907	Mm
Pu=1% :	$x=x_{mes}-2.58s_x=$	427.6046723	Mm
% Pu	Gauss (mm)	Galton (mm)	
1	427.6046723	471.7654729	
5	508.5576907	527.8134886	
15.87	633.9042998	628.0165127	
50	764.4736842	752.6742906	
84.13	895.0430687	902.0759428	
95	1020.389678	1073.331016	
99	1101.342696	1200.847922	

Table 4. Calculations

$x=x_{mes} \pm k*\sigma \ ku; \ k = \pm 1, \pm 2, \pm 3$	
$x=x_{mes.}:$	$f_0 = 0.152772476$
$x=x_{mes} \pm \sigma$	$f_1 = 0.092661191$
$x=x_{mes} \pm 2\sigma$	$f_2 = 0.020675506$
$x=x_{mes} \pm 3\sigma$	$f_3 = 0.001697149$

(Figure 5), in which case the calculations are formed in tabular form together with the Table 4.

On the surface $\pm 3\sigma$, which includes 99.73% of all cases ($0.9973 * 38 = 37.897$), almost all
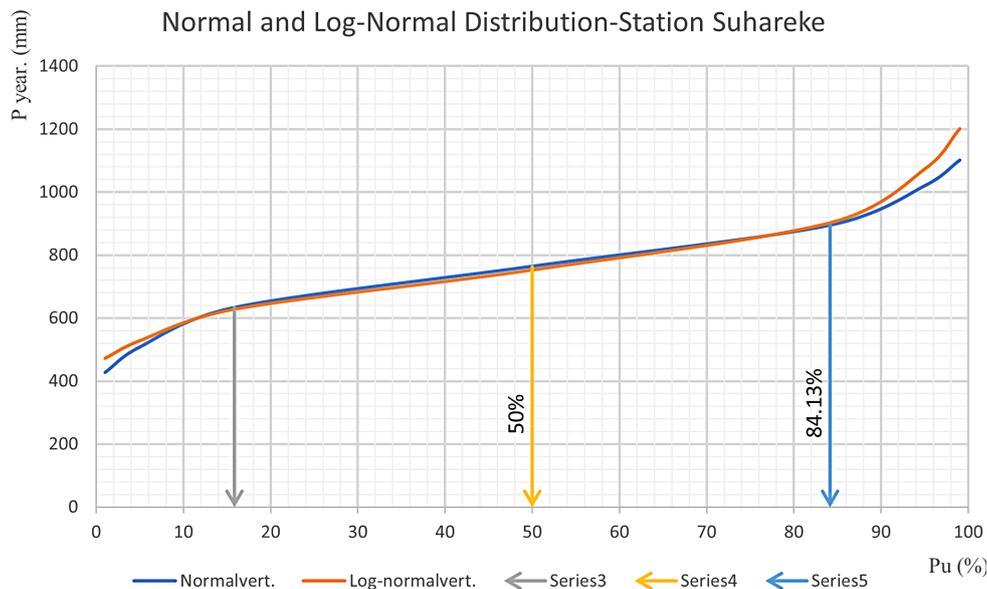

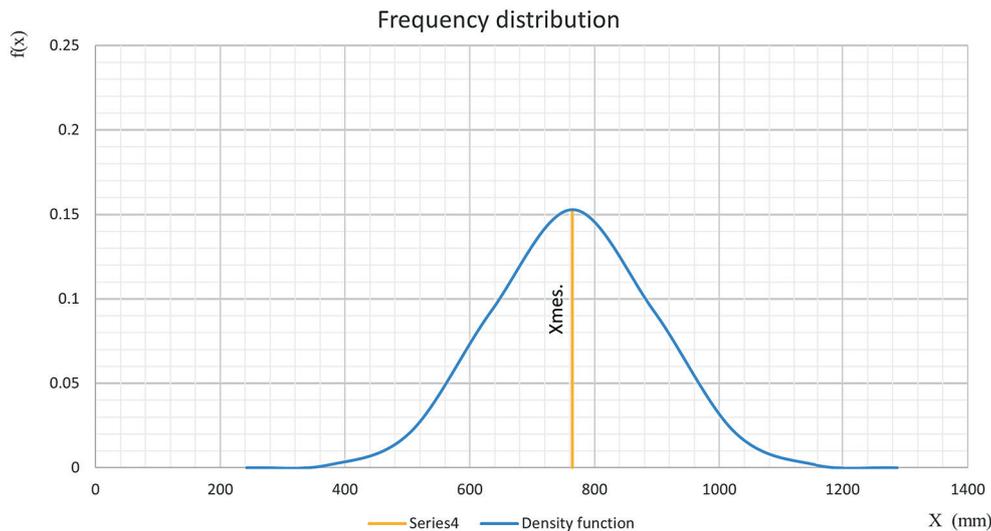**Figure 4.** Normal and log-normal distribution – Station Suhareke

Figure 5. Normal distribution density function

cases of our group fall within the curve bounded by $x = x_m \pm 3\sigma$.

Although only 0.26% of all values are out of bounds $x = x_m \pm 3\sigma$, it is still a shortcoming of the normal distribution that its smallest value $-\infty$ is not physically meaningful. To obtain only the positive range (space) of cases, instead of x_i we obtain $y_i = \log x_i$ or $y_i = \ln x_i$ (log-normal, Galton or Fechner distribution). In this case the function will be defined in the interval $(0, \infty)$.

Autoregressive models for simulating monthly precipitation

The time series models used are often based on the equation

$$x_i = \mu_x + \rho_1 * (x_{i-1} - \mu_x)$$
$$+ t_i * \sigma_x * \sqrt{(1 - \rho_1{}^2)} \tag{18}$$

known as the first-order equation of Markov models.

For the rainfall of a season or a year with average μ_x and autocorrelation coefficient ρ with time shift 1 we have:

$$q_i = \mu + \rho * (q_{i-1} - \mu) + e_i \tag{19}$$

The recursion link for the generation of synthetic time series is the Fiering model. The application of the Fiering model to a normally distributed group with mean μ, standard deviation σ and autocorrelation coefficient r_i is given by the recursion expression (Maniak, 2010; Mujumdar, 2019):

$$q_i = \mu * (1 - r_1) + r_1 q_{i-1} +$$
$$+ t_i \sigma \sqrt{(1 - r_1{}^2)} \tag{20}$$

To establish the recursion equation, the average monthly rainfall is initially calculated as:

$$\mu_j = \frac{1}{N} * \sum_{k=1}^{N} x_{k,j} \tag{21}$$

The variance of the individual time intervals (monthly) t will be:

$$\sigma_j{}^2 = \frac{1}{N} \sum_{k=1}^{N} x_{k,j} -$$
$$- \frac{1}{N(N-1)} \left(\sum_{k=1}^{N} x_{k,j} \right)^2 \tag{22}$$

While the autocorrelation coefficient is calculated with the expression:

$$r_j = \frac{\sum_{k=1}^{N} x_{k,j} x_{k,j-1} - N \mu_j \mu_{j-1}}{\sigma_j \sigma_{j-1}(N-1)} \tag{23}$$

The general form of the equation suitable for use is:

$$q_{i,j} = \mu_j + \frac{r_j \sigma_j}{\sigma_{j-1}} * \left(q_{i,j-1} - \mu_{j-1} \right) +$$
$$+ t_{i,j} \sigma_j \sqrt{(1 - r_j{}^2)} \tag{24}$$

where: $q_{(i, j)}$ – the precipitation generated by string (i) in the i-th time interval, e.g for $t = 1$ per month we have $j = 1,2,\ldots, 12$;

μ_j – monthly precipitation averages (j) ($j{\leq}12$); σ_j – standard monthly deviation (j); r_j – the correlation coefficient between qj and qj-1; t_i – random numbers normally distributed with $\mu = 0$ and $\sigma = 1$.

The ratio $\dfrac{r_j \sigma_j}{\sigma_{j-1}} = bj$ otherwise is known as the regression coefficient.

On the basis of what was said above, the 30-year series of monthly rainfall were takend and the statistical parameters were set as in the Table 5.

The model is based on equation (24), where 12 consecutive equations are described. The starting point is the average rainfall in October, 61.33 mm. To generate synthetic time series, a sequence of randomly distributed random numbers were computed $z_1, z_2, \dots$. These numbers z_i can be generated from a table of random numbers such as -0.313, -0.951, 0.590 and so on.

However, according to the model, z_i numbers should be converted to random gamma distribution numbers according to the equation:

$$tg = \frac{2}{C_{t,j}}\left(1 + + \frac{C_{t,j}t_i}{6} - \frac{C_{t,j}{}^2}{36}\right)^3 - \frac{2}{C_{t,j}} \quad (25)$$

where $C_{t,j}$ is the coefficient of asymmetry expressed by the equation

$$C_{t,j} = \frac{C_{sj} - r_{j-1}{}^3 C_{sj-1}}{\sqrt{\left(1 - r_j{}^2\right)^3}} \quad (26)$$

where: t_i – random numbers normally distributed (0; 1), t_g – random gamma distribution numbers (0; 1; C_g), C_{sj} – coefficient of asymmetry for months.

After converting these numbers, new numbers are obtained according to equation (25); where the 12 equations were then laid out for each month and the values were simulated:
First year:

$$q_N = 86.7 + 0.0057 * (q_{i,o} - 61.33) + + t_i * 39.22 * \sqrt{(1 - 0.006^2)}$$

Table 5. Statistical parameters of monthly precipitation in (mm) for 30-year series

| Month	X_{mes}	S_j	Cs	r_j	b_j
N	86.7	39.22187	-0.00594	0.006158	0.005767
D	80	39.41906	0.291129	0.13133	0.13199
J	70.33333	46.03097	1.160096	-0.17544	-0.20487
F	57	38.62017	0.937803	0.1177	0.098751
A	65.5	46.08968	1.380183	-0.24722	-0.29504
M	60.56667	25.39303	0.453322	-0.08208	-0.04522
M	75.7	39.49784	0.768659	-0.0487	-0.07574
J	65.86667	41.95789	1.671599	-0.02247	-0.02387
J	62.63333	49.11737	2.635784	-0.0126	-0.01475
A	45.63333	33.25501	0.658228	-0.21308	-0.14426
S	63.56667	50.52939	1.741598	0.012554	0.019075
O	61.33333	41.87776	0.600436	-0.19055	-0.15792
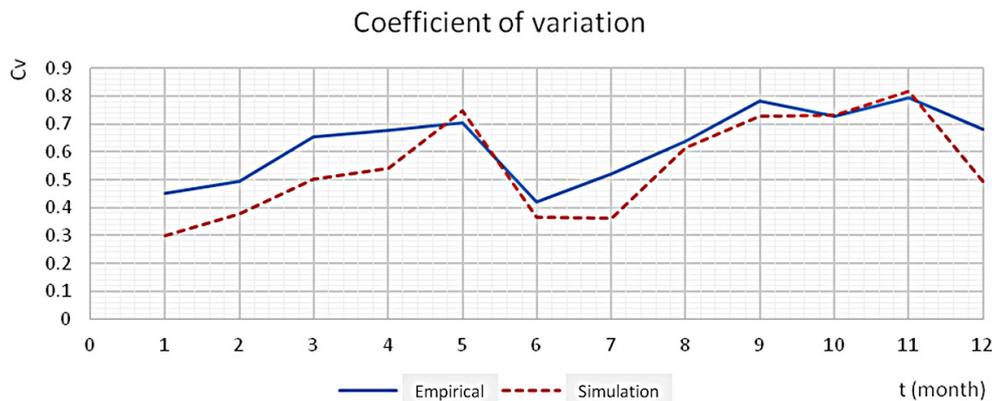
Figure 6. Model summary according to equation (24), for a 20-year simulated monthly rainfall hydrograph

$$q_D = 80 + 0.132 * \left(q_{i,N} - 86.7\right) + \\ + t_i * 39.41 * \sqrt{(1 - 0.313^2)}$$

Second year:

$$q_N = 86.7 + 0.0057 * \left(q_{i-1,o} - 61.33\right) + \\ + t_i * 39.22 * \sqrt{(1 - 0.006^2)}$$

The similarity of the empirical values with the simulated ones give a coefficient of variation which expresses the ratio of standard deviations and their arithmetic means (Figure 6).

RESULTS AND DISCUSSION

Processing of IDF curves

Assuming that the rainfall series follow the Pearson III distribution, and based on the data on the maximum daily rainfall for the Suhareka region, the probability distribution of heavy 24h precipitation (according to the Pearson-III and log-Pearson-III distributions) for the combined series of maximum rainfall is determined.

Pearson-III distribution for a 30-year series of maximal daily precipitation has the following statistical parameters: $X_m = 27.116$ mm; $C_{vx} = 0.338$; $C_{sx} = 0.845$, while log-Pearson-III: $Y_m = 1.409$; $C_{vy} = 0.104$; $C_{sy} = -0.129$.

As a mathematical model of the dependence $I\,(T, P)$ the following expression is taken (Babac, 2006).

$$I(T, P) = \frac{I_0(P)}{(A * T + 1)^B} \quad (27)$$

where: A and B are dimensionless parameters, $Io\,(P)$ is the rainfall intensity limit, i.e.

$\lim_{T \to 0} I(T, P)$. According to the Russian author Sokolovsy, this model was used for the territory of the European part of the former Soviet Union. If it is assumed that parameters *A* and *B* of equation (27) can be presented in the form of a map for the observed territory, then the intensity values of the boundary *Io* (*P*) for a point can be determined through the corresponding values of the maximum daily rainfall *Hd* (*P*). According to this model for rain duration *T* = 1440 min. and probability *P*:

$$H(1440, P) = \frac{I_0(P)}{(1440 * A + 1)^B} * 1440 \quad (28)$$

Relationship $H(1440, P) = a * H_d(P)$ is quite logical, and that coefficient (*a*) is close to unity because these are magnitudes determined by a series of annual maximum daily rainfall. Then, the following is derived from equation (28):

$$I(T, P) = \frac{a}{1440} * \\ * \left(\frac{1440 * A + 1}{A * T + 1}\right)^B * H_d(P) \quad (29)$$

Using the last equation and the coefficients for Suhareka (*A* = 0.3; *a* = 1.0; *B* = 0.79) the following expression is obtained:

$$I(T, P) = \frac{1}{1440} * \\ \left(\frac{1440 * 0.3 + 1}{0.3 * T + 1}\right)^{0.79} * H_d(P) \quad (30)$$

It is worth noting that the parameter *A* in the model in most cases is constant, while *B* is the
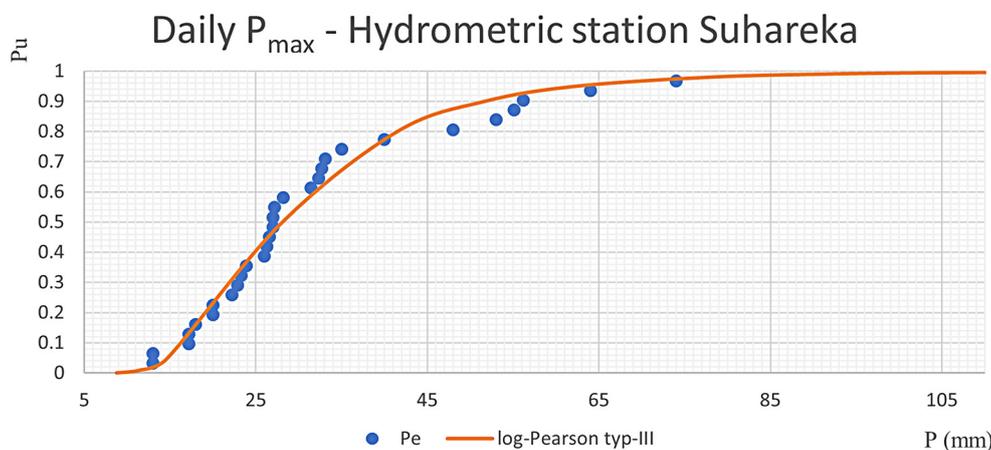
Figure 7. Probability of 24h heavy rain distribution according to Pearson type-III
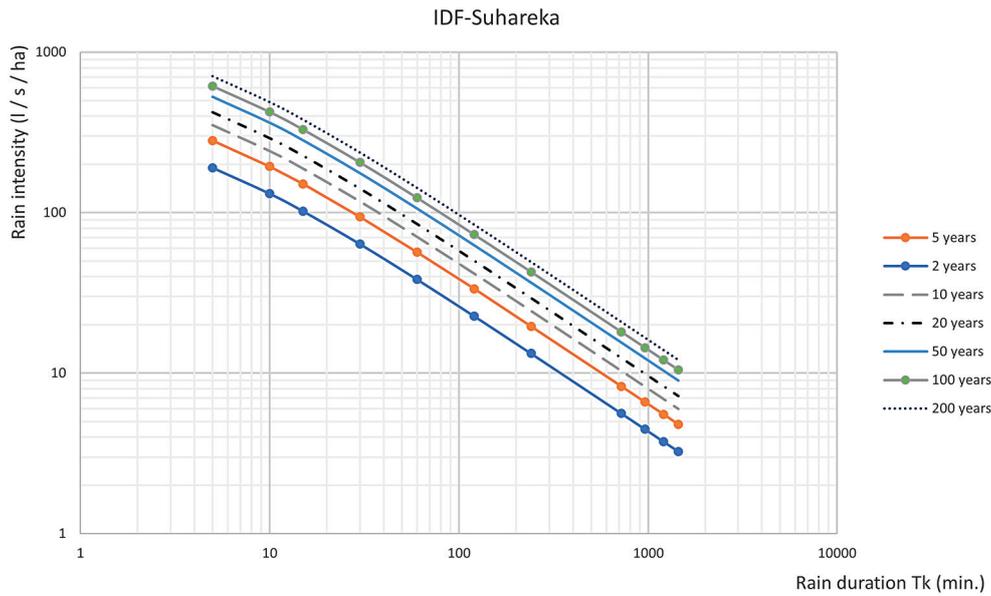
Figure 8. Precipitation intensity curves for different return periods

'coefficient of reduction of rainfall intensity over time' or as it is otherwise known the 'continental coefficient' which seems more acceptable because it has a physical interpretation.

According to the Pearson type-III distribution (Figure 7), the maximum daily rainfall heights for different return periods are calculated according to the expression $x_T = x_m + \sigma * k(C_{sx}; T)$, where k- represents the frequency factor.

To assess the suitability of the theoretical distribution with the empirical one, statistical testing was performed according to the test χ_2. The same resulted in the value c2 = 2.64 < c = 7.81, which means that the hypothesis is accepted below the 1-α = 5% probability level or statistical certainty is 95%.

By substituting the rainfall heights for the respective return periods in eq. (30), rain intensities are found (Figure 8).

Flow curves

Since no flow measurements have been performed in the Suhareka region, then it is necessary for high waters to be indirectly determined by the transformation of rainfall into flow. For this purpose, the American SCS method is used, which uses the following characteristics of the basin: the length of the flow (*L*), the length to the center of gravity of the river basin (*Lc*), the slope of the flow (i_{ur}) and the surface of the river basin (*F*).

Numerous analyses have shown that in the triangular-shaped synthetic hydrograph, the relationship between rain duration (t_k) and delay time (t_p) is linear. The latency time of the basin is defined by the expression:

$$t_p = T_0 + a * t_k \tag{31}$$

where: *a* – represents the regression coefficient which is a function of the surface *a = f(F)*, while T_0 depends on the topographic characteristics of the basin and is determined by the empirical expression:

$$T_o = 0.4 * L^{0.63} * \left(\frac{L * L_c}{\sqrt{i_{ur}}}\right)^{0.086} \tag{32}$$

By recognizing precipitation with different durations and different return periods the effective precipitation can be determined according to the SCS method:

$$P_e = \frac{(P - 0.2 * d)^2}{P + 0.8 * d}, [mm] \tag{33}$$

where: *d* – soil moisture deficit, which is determined by the formula:

$$d = \left(\frac{1000}{CN} - 10\right) * 25.4, [mm] \tag{34}$$

Whereas, the CN number expresses the characteristics of the soil in the catchment, which in the considered case is more complicated, since there are two groups of soils (*B* and *C*) with

119

different composition: group B with about 40%, and group *C* with 60% participation. Thus, it turns out that the number $CN \cong 73$,

while $d = \left(\dfrac{1000}{73} - 10\right) * 25.4 = 93.945$ mm.

Precipitation from IDF curves, due to the climatic effect and non-uniform distribution of precipitation is increased by 8%, and the effective precipitation is determined according to equation (33).

For the Suhareka river basin the following characteristics can be used: $L = 15$ km; $Lc = 7.5$ km; $i_{ur} \cong 5\%$ and $F = 80.2$ km², of which according to equation (32) and (31) results: $T_o = 3.761$h dhe $t_p = 3.761 + a * t_k$. For $F = 80.2$ km² the regression coefficient is $a = 0.44$. The setting time of the synthetic hydrograph is defined as: $T_p = t_k /2 + t_p = T_o + 0.94 * t_k = 3.761 + 0.94 * t_k$.

Due to the small surface area of the basin, it can be assumed that in the synthetic hydrograph it is $Tp = Tr$ (rise time = fall time). Eventually, by condition:

$$P_e * F = \frac{1}{2} * Q_{max.} * (T_p + T_r)$$

Derives,

$$Q_{max.} = \frac{P_e * F}{T_p} \tag{35}$$

By adjusting the units in the last equation, the maximum inflows for different return periods are obtained, which are presented in the Figure 9.

Finally, the dependence of the flow factor (X) of the area for different durations (h) and different return periods (years) is given. In this case, the effective rainfall is expressed according to the formula:
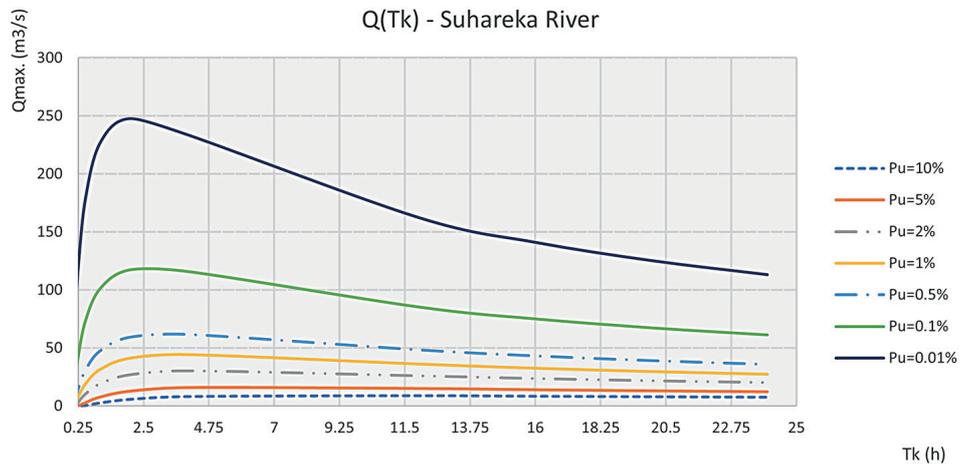

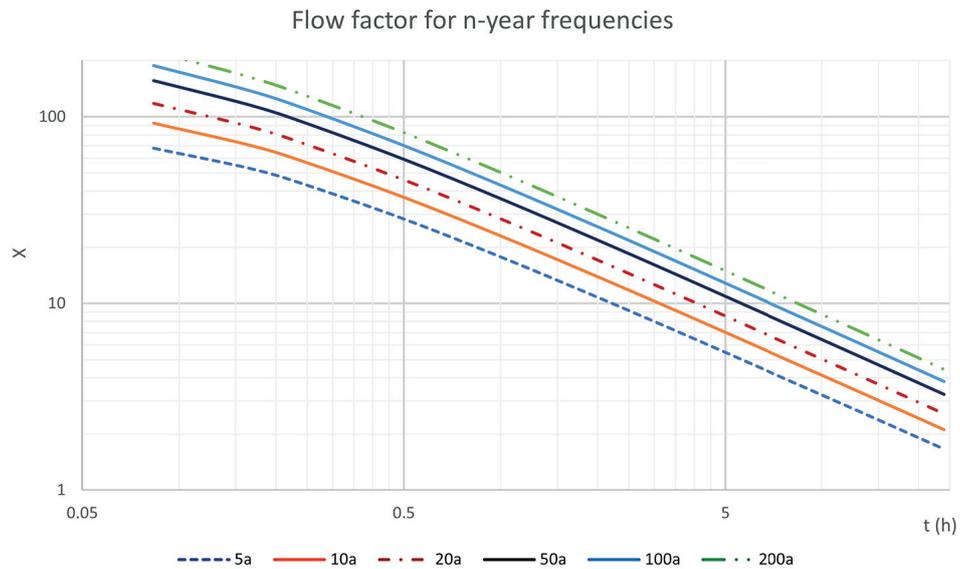**Figure 9.** Maximum flows for different probabilities

Figure 10. Flow factor for different return periods

$$P_e = \frac{\left(P - \frac{200}{CN} + 2\right)^2}{P + \frac{800}{CN} - 8} \qquad (36)$$

where: P – precipitation for different return periods and $CN = 73$.

The flow factor is determined according to the expression (Ven, 1962):

$$X = \frac{P_e}{t} \qquad (37)$$

where: t – time in hours (h).

CONCLUSIONS

The aim of this research was to estimate the flow for ungauged Suhareka River, based on the rainfall data available. In the considered case, the triple regression was obtained, connecting the hydrometric station of Suhareka (y) with those of Prizren (x), Ferizaj (z) and Prishtina (t). From these analyses of variance an important, but not sufficient correlation was obtained. The inadequacy of this relationship is noticed by the high values of t-test and low ones of F-test as well as P-value. However, the cause of the insufficient correlationn is the nonlinearity between the variables, so after using transformation the Suhareka hydrometric station was connected with three other stations, using non linear triple regression analyses. The correlation coefficient in this case is higher than that of linear regression. Thus, $r_{yx1x2x3} \cong 0.62$, which represents a very important correlation of these stations.

The next step was the calculation of the distribution functions for extreme rainfall, with the use of Laplace and Gaussian law of large numbers. For a series of 30 years monthly rainfall data for the Hydrometric station in Suhareke, the Normal and Log Normal Distribution as well as Normal Distribution density function were calculated. Assuming that the rainfall series follow the Pearson III distribution, and based on the data on the maximum daily rainfall for the Suhareka region, the probability distribution of heavy 24h precipitation (according to the Pearson-III and log-Pearson-III distributions) for the combined series of maximum rainfall is determined.

According to the Pearson type-III distribution, the heights maximum daily rainfall heights for different return periods, in Hydrometric Station Suhareks, are calculated.

Statistical testing was performed (according to the test $\chi_{2)}$ to assess the suitability of the theoretical distribution with the empirical one. The results showed the acceptance of the hypothesis below the $1 - \alpha = 5\%$ probability level or with statistical certainty about 95%.

Aa a result, the Precipitation intensity curves were given for different return periods. Since no flow measurements have been performed in the Suhareka region, the SCS method was used, by which the high waters were indirectly determined by the transformation of rainfall into flow. The equation for the Suhareka Basin charcteristics is derived $\frac{P_e * F}{T_p}$ for this purpose and the maximum inflows for different return periods were obtained.

The results obtained through this paper, indicates that even for ungauged river basins the peak flows can be determined from the available rainfall data. This will be of great help to the water engineers that are facing many data deficiencies while managing water resources.

REFERENCES

1. Baoyan L., Wang D., Fu S., Cao W. 2017. Estimation of peak flow rates for small drainage areas. Water Resources Management, 1635–1647.

2. Kolbjorn E., Alfredsen K. 2020. Hydrology and water resources management in a changing world. Hydrology Research, 143–145.

3. McMillan H., Westerberg I.K., Krueger T. 2018. Hydrological data uncertainty and its implications. Wiley Interdisciplinary Reviews Water. https://doi.org/10.3390/w10111669

4. Nam K.W., Shin M.J. 2018. Estimation of peak flow in ungauged catchments using relationship between runoff coefficient and curve number. Water.

5. Nearing G.S., Gupta H.V., Clark M.P., Tian Y., Harrison K.W., Weijs S.V. 2016. A philosophical basis for hydrological uncertainty. Hydrological Sciences Journal, 1666–1678.

6. Maniak U. 2010. Hydrologie und Wasserwirtschaft'; Eine einführung für Ingenieure, 6. neu bearbeitete Auflage. TU Braunschweig.

7. Ven T.C. 1962. Hydrologic determination of waterway areas for the design of drainage structures in small drainage basins. University of Illinois at Urbana-Champaign.

8. Husno H. 2007. Inżenjerska Hidrologija. Sarajevo.

9. Babac P. 2006. Osnovi Hidrotehnike u šumarstvu primeri iz teorije i prakse. Beograd.

10. KHMI. 1984. Hydromeorological yearbooks of Kosovo 1954–1983.

11. Bektesh B. 2005. Statistika elementare Prishtinë.

12. Karakuş C. 2020. Istatistiksel analiz, olasılık ve rassal değişkenler.