COINCIDENCE OF LOW FLOWS AT NEIGHBOURING HYDROLOGIC STATIONS

Vojislav Vukmirović, Jasna Petrović and Dragana Jovičić
Faculty of Civil Engineering, University of Belgrade, P.O. Box 42, 11120 Belgrade, Yugoslavia, e-mail: jasnap@hikom.grf.bg.ac.yu

***Abstract*:** The aim of this paper is to investigate the possibility of estimating probability of simultaneous occurrence (coincidence) of low flows in two rivers. Three aspects are considered: (1) joint distribution of low flows of two rivers before junction using bivariate normal distribution after logarithmic transformation of flows; (2) mutual dependance of flows at two rivers by means of the correlation field; (3) distribution of the sum of flows at two rivers before junction to describe flows downstream of the junction. The results of analysis with 30-days minimum flows from two hydrologic stations in Serbia is used to demonstrate the correlation ellipses, probability of simultaneous occurrence of low flows, and the distribution function of the sum of flow rates at two stations compared to the downstream control station.
***Keywords*:** low flows, simultaneous occurrence, bivariate distributions.

1. Introduction

Simultaneous occurrence of extreme low flows in two rivers is an interesting issue when hydrologic stations are situated near the junction of two rivers, and there is no hydrologic station immediately downstream the junction. The aim of this paper is to investigate the possibility to estimate probability of simultaneous occurrence (coincidence) of low flows in two rivers and corresponding probability of low flows downstream of the junction.

When considering flow rates in two rivers near their junction, it is virtually certain that they are mutually dependent with a certain degree of correlation. Dependance of two flow rate series can be investigated by using the regression analysis and/or multivariate distributions. The latter approach is applied in this study.

Among various multivariate distributions in the statistical theory, bivariate normal distribution has been studied far more extensively than any other multivariate distribution and the mathematical apparatus for its calculation is more easily applicable then for other bivariate distributions. However, hydrological variables, such as low flows, are generally skewed and it is therefore necessary to perform their normalization, i.e. to transform the random variables so that the resulting series are normally distributed. The most common transformation is the logarithmic function.

In this paper we shall review the bivariate normal distribution and present the procedure for obtaining the probability distribution function of low flows downstream the junction of two rivers.

2. Bivariate normal distribution

2.1. Definition and calculation

Simultaneous occurrence of two random variables, X_1 and X_2, is described with a bivariate distribution function defined as:

$$F(x_1,x_2)=P\{X_1 \le x_1, X_2 \le x_2\}=\int_{-\infty}^{x_1}\int_{-\infty}^{x_2} f(x_1,x_2)\,dx_1 dx_2 \tag{1}$$

where $f(x_1,x_2)$ is the joint probability density function of random variables X_1 and X_2. Bivariate normal distribution assumes that the two random variables X_1 and X_2 are normally distributed, with means m_1 and m_2 and standard deviations s_1 and s_2, respectively. The normal joint density function is defined with:

$$f(x_1,x_2)=\frac{1}{2\pi s_1 s_2\sqrt{1-r^2}}\cdot$$
$$\cdot\exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{(x_1-m_1)^2}{s_1^2}-2r\frac{(x_1-m_1)(x_2-m_2)}{s_1 s_2}+\frac{(x_2-m_2)^2}{s_2^2}\right]\right\} \tag{2}$$

where r is the correlation coefficient between X_1 and X_2. By substituting:

$$t_1=\frac{x_1-m_1}{s_1},\quad t_2=\frac{x_2-m_2}{s_2} \tag{3}$$

the bivariate normal distribution function can be standardized:

$$\Phi(h,k)=P\{t_1\le h,t_2\le k\}=\frac{1}{2\pi\sqrt{1-r^2}}\int_{-\infty}^{h}\int_{-\infty}^{k}\exp\left\{-\frac{t_1^2-2r\,t_1t_2+t_2^2}{2(1-r^2)}\right\}dt_1dt_2 \tag{4}$$

To enable calculations of $\Phi(h,k)$, the above expression is expanded into a series (Abramowitz and Stegun, 1970):

$$\Phi(h,k)=\Phi(h)\,\Phi(k)+\sum_{n=0}^{\infty}\frac{\varphi^{(n)}(h)\,\varphi^{(n)}(k)}{(n+1)!}r^{n+1} \tag{5}$$

where:

$$\Phi(u)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u}e^{-\frac{t^2}{2}}dt\,,\quad \varphi(t)=\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\,,\quad \varphi^{(n)}(t)=\frac{d^n\varphi(t)}{dt^n}$$

The derivatives $\varphi^{(n)}(t)$ can be expressed using the Hermite polynomials:

$$\varphi^{(n)}(t)=(-1)^n H_n(t)\varphi(t)$$

The Hermite polynomials are defined by the following recurrent formula:

$$H_n(t)=tH_{n-1}(t)-(n-1)H_{n-2}(t)\quad n\ge 2$$
$$H_0(t)=1$$
$$H_1(t)=t$$

2.2. Correlation field

The normal joint density of two variables X_1 and X_2 (eqn. 2) forms a 3D bell shape, whose cross-sections for $f(x_1,x_2)=const$ are elliptical contours, called correlation ellipses. The same is true for the standardized normal joint density of t_1 and t_2. These contours define ranges of the variables within which specified proportion of the probability distribution lies. If α is the probability that the variables t_1 and t_2 are inside the correlation ellipse (domain E):

$$\alpha=\frac{1}{2\pi\sqrt{1-r^2}}\iint_E\exp\left\{-\frac{t_1^2-2r\,t_1t_2+t_2^2}{2(1-r^2)}\right\}dt_1dt_2$$

it can be shown (Johnson and Kotz, 1972) that the corresponding ellipse has the equation:

$$t_1^2-2rt_1t_2+t_2^2=-2(1-r^2)\ln(1-\alpha) \tag{6}$$

For a fixed value $t_2 = k$, the above equation gives pairs of values h_1 and h_2 for the variable t_1. and the ellipse can be constructed for specified probability α.

2.3. Transformation of variables

As indicated in introduction, hydrological variables are seldom normally distributed. They are generally skewed, and therefore it is necessary to transform them so that the resulting series are normally distributed.

The logarithms of flow rates can usually be fitted to the normal distribution. If the variable $\log Q$ is normally distributed, then Q follows the two-parameter log-normal distribution. Similarly, if $\log(Q - Q_0)$ is normally distributed, then Q follows the three-parameter log-normal distribution. Coincidence of low flows in two rivers is then analyzed using the bivariate normal distribution of two random variables X_1 and X_2 defined as:

$$\begin{aligned} X_1 &= \log(Q_1 - Q_{01}) \\ X_2 &= \log(Q_2 - Q_{02}) \end{aligned} \tag{7}$$

where Q_1 and Q_2 are flow rates at two rivers, and Q_{01} and Q_{02} are location parameters of three-parameter log-normal distributions fitted to the series of Q_1 and Q_2. Statistics of the series X_1 and X_2 are means m_1 and m_2, standard deviations s_1 and s_2, and skews $g_1 = g_2 = 0$.

3. Flow rates downstream of the junction

If some characteristic flow rates Q_1 and Q_2 in two streams are known (e.g. flow rates of specified probability), it is of practical importance to determine the same characteristic flow rate Q_s after the junction of two streams.

The transformed random variable Z, representing the flow downstream the junction, is defined as the sum:

$$Z = X_1 + X_2 \tag{8}$$

The domain of the variable Z, represented as area D in Figure 1, is the area below the line $x_1 + x_2 = z$. Distribution function $F(z)$ of the variable Z is then:

$$F(z) = P\{Z \le z\} = \iint\limits_{x_1 + x_2 \le z} f(x_1, x_2)\, dx_1 dx_2 = \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{z - x_1} f(x_1, x_2)\, dx_2 \tag{9}$$

where $f(x_1,x_2)$ is the bivariate normal density (2). Differentiating the above equation in respect to z, we obtain:

$$f(z) = \int_{-\infty}^{\infty} f(x_1, z - x_1) dx_1 \tag{10}$$

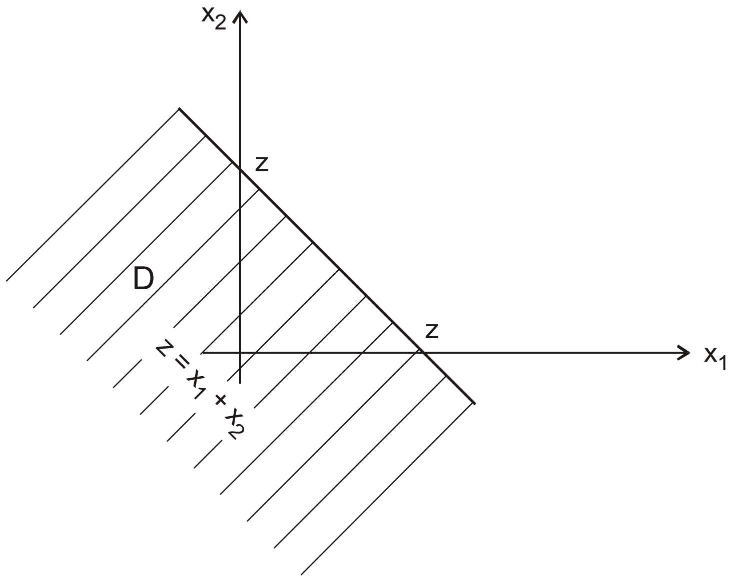

Figure 1. Domain of integration of bivariate normal density function for the sum of transformed variables representing flow rates at two streams.

Integrating the right-hand side of (10) it can be shown that Z is normally distributed:

$$f(z)=\frac{1}{\sqrt{2\pi}\sqrt{s_1^2+2rs_1s_2+s_2^2}}\cdot\exp\left\{-\frac{(z-m_1-m_2)^2}{2(s_1^2+2rs_1s_2+s_2^2)}\right\} \tag{11}$$

From the above equation it is clear that the mean and the variance of Z are:

$$\begin{aligned} m_z &= m_1+m_2 \\ s_z^2 &= s_1^2+2rs_1s_2+s_2^2 \end{aligned} \tag{12}$$

Distribution function of Z is then:

$$F(z)=\frac{1}{s_z\sqrt{2\pi}}\int_{-\infty}^{z}\exp\left\{-\frac{(z-m_z)^2}{2s_z^2}\right\} \tag{13}$$

However, if the transformation of the original variables is logarithmic (such as in (7)), then the sum of the transformed variables (8) leads to the product of the original variables, which is not what are we looking for. Therefore, if variables Q_1 and Q_2 follow log-normal distribution, the above procedure is not suitable for finding the distribution function of their sum.

In case that variables Q_1 and Q_2 follow the Pearson type III distribution (or three-parameter gamma), random variable $Z = Q_1 + Q_2$ will also be Pearson III (or three-parameter gamma) distributed with the following parameters:

– mean: $m_z = m_1+m_2$

– variance: $s_z^2 = s_1^2+2rs_1s_2+s_2^2$ (14)

– skew: $g_z = \frac{1}{s_z^3}\left[g_1s_1^3+g_2s_2^3+3(m_{12}+m_{21})\right]$

where m_1, s_1, g_1 and m_2, s_2, g_2 are means, standard deviations and skews of Q_1 and Q_2 respectively, and m_{12} and m_{21} are second-order mixed moments of two variables.

4. Case study

The subject of this study was coincidence of low flows at station Mojsinje on the Juzna Morava river and station Jasika on the Zapadna Morava river (Figure 2). The downstream control station was station Varvarin on the Velika Morava river.

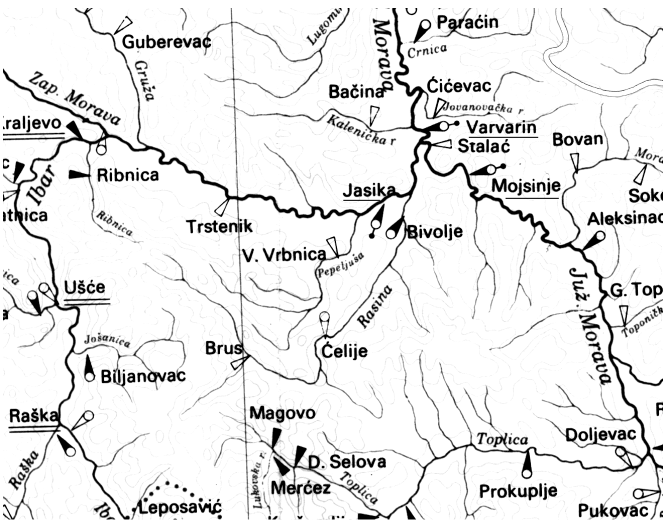

Figure 2. Location of the hydrologic stations in the study.

The analysis was based on 30-days minimum flows. Data covered the 1959-1994 period. In order to apply the bivariate normal distribution of flows at the two stations, it was necessary to transform the flows according to equation (7). Location parameters Q_{01} and Q_{02} are chosen so that the resulting series $X_j = \log(Q_j - Q_{0j})$, j = 1,2, have skewness coeffcients equal to 0. Table 1 presents statistics of the series. The correlation coefficient between transformed flow rates at stations Jasika and Mojsinje is 0.726.

Table 1. Statistics of flow rates at stations Jasika and Mojsinje.

	Flow rates Q		$\log Q$		$\log(Q - Q_0)$	
	Jasika	Mojsinje	Jasika	Mojsinje	Jasika	Mojsinje
location parameter Q_0					4.219	0.697
mean	29.03	20.91	1.432	1.285	1.352	1.267
standard deviation	11.63	9.24	0.162	0.176	0.195	0.184
coef. of variation	0.401	0.442	0.113	0.137	0.144	0.145
coef. of skewness	1.382	1.781	0.212	0.064	0.000	0.000

The correlation ellipses are constructed according to equation (6) for different probabilities α and they are shown in Figure 3. However, correlation ellipses are suitable for checking if the bivariate normal distribution is an appropriate model, i.e. for checking whether a specified percentage of observed pairs of values (Q_1, Q_2) fall into a corresponding ellipse.

For analyzing the probability of simultaneous occurence of low flows at two stations smaller then a specified value, it is more convenient to construct lines of equal probabilities of non-exceedance. Such lines are presented in Figure 4.

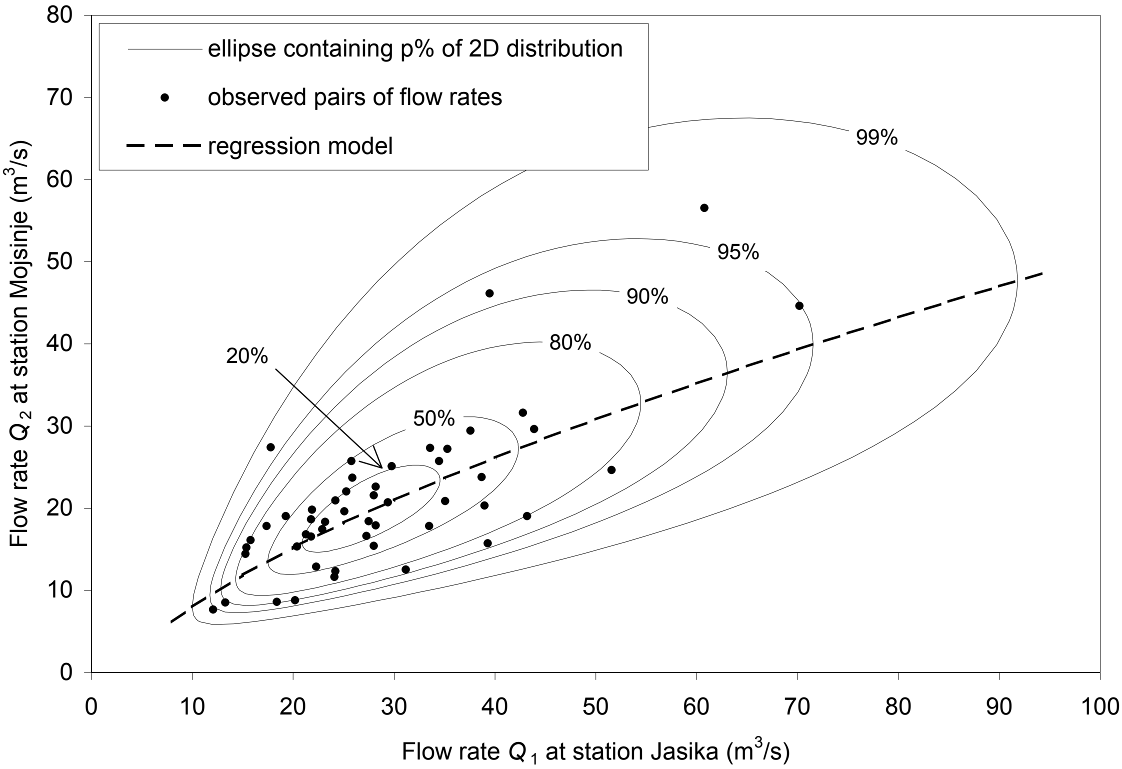

Figure 3. Correlation ellipses for stations Mojsinje and Jasika.

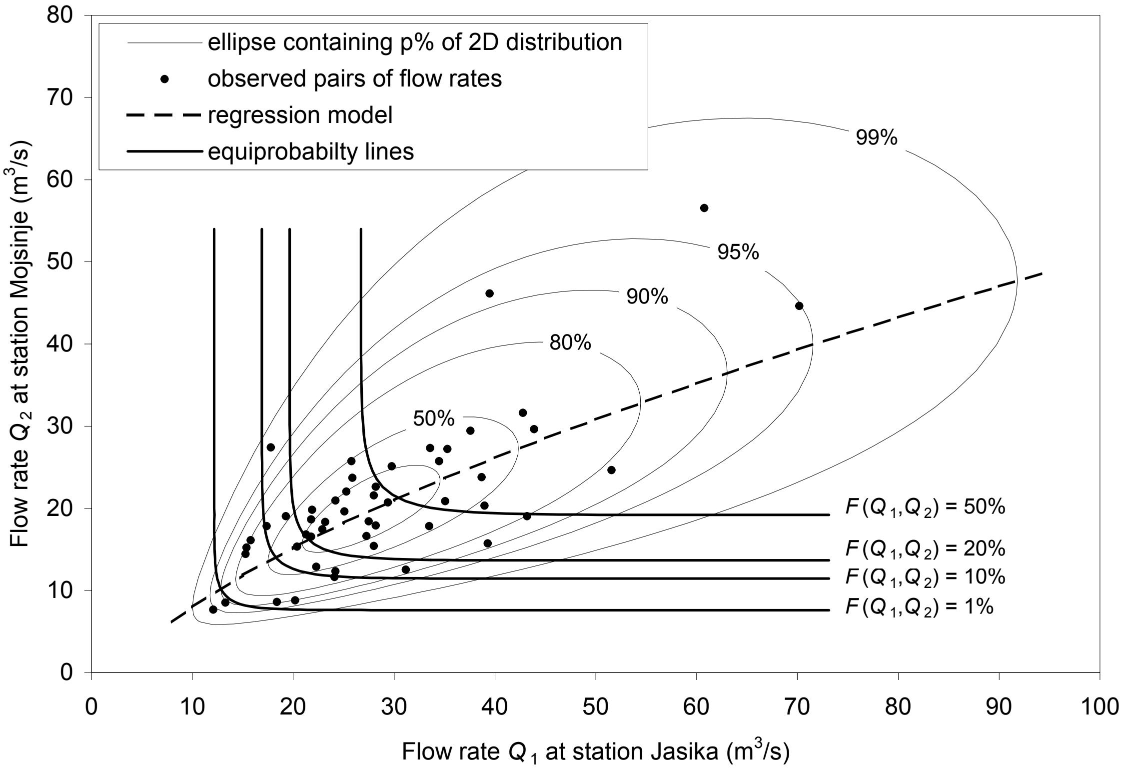

Figure 4. Lines of equal probability of simulatneaous non-exceedance of flow rates $F(Q_1,Q_2) = P\{Q_1 \le q_1, Q_2 \le q_2\}$ at stations Mojsinje and Jasika.

Distribution function of flow rates downstream of the junction of two rivers was obtained using the Pearson type III distribution, as explained in section 3. Relevant statistics and parameters are given in Table 2.

Figure 5 presents distribution functions of 30-days low flows at stations Mojsinje and Jasika, distribution function of the sum of these flows, as well as the distribution function of the station Varvarin as the control station.

It can be seen that the distribution of the sum agrees very well with the distribution of the control station in the domain of low waters. Since the other domain (below the probability of 80%) is not very interesting for the low flow analysis, disagreement of these two distributions is not essential.

Table 2. Statistics of flow rates for stations Jasika and Mojsinje, their sum and flow rates at the downstream control station Varvarin.

	Jasika Q_1	Mojsinje Q_2	sum of Jasika and Mojsinje $Q_s = Q_1 + Q_2$	Varvarin Q_3
mean	29.03	20.91	49.94	54.59
standard deviation	11.63	9.24	19.58	22.57
coef. of skewness	1.382	1.781	1.609	1.517
mixed moment m_{12}	1248.16			
mixed moment m_{21}	1416.85			
correlation coefficient $r(Q_1,Q_2)$	0.758			

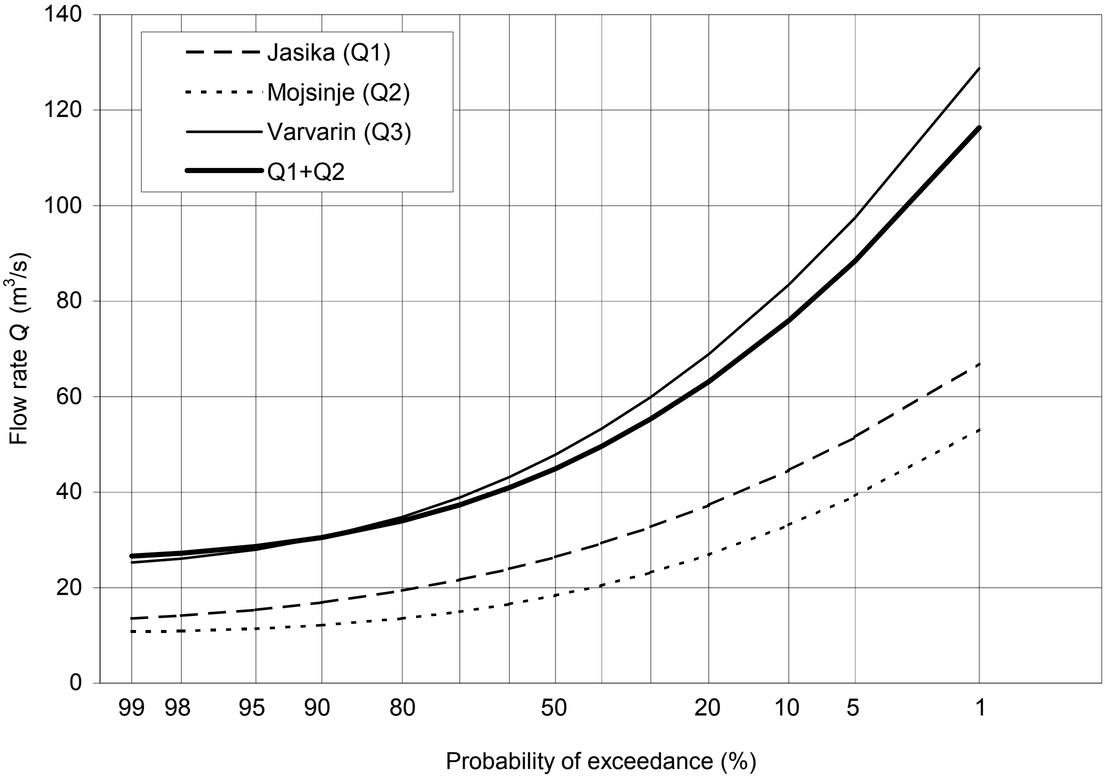

Figure 5. Distribution functions of individual stations, their sum and of the control station.

5. References

Abramowitz, M., and Stegun, A. (1990) *Handbook of Mathematical Functions*, Dover Publications, New York.

Johnson, N.L., and Kotz, S. (1972) *Distribution in Statistics: Continuous Multivariate Distributions*, John Wiley & Sons, New York.