

**BIVARIATE JOINT PROBABILITY ANALYSIS OF FLOOD  
HAZARD AT RIVER CONFLUENCE**

Veronika Bačová Mitková\*

Assessing the statistical significance of floods in the complex hydrological conditions that exist at the confluence of the main river and its tributaries, as well as the choice of hydrological design parameters for flood protection in these areas is one of the main tasks in the current hydrology. The main aim of this paper is a joint bivariate frequency analysis of annual peak discharges and synchronously occurred maximum discharges at the main river and its tributary. The annual maximum and daily maximum discharges of the Morava River, as main river, and its tributary Myjava River were analysed. We selected the most appropriate copula function for our bivariate analysis. Selected copulas were used to illustrate the joint occurrence probabilities and joint return periods of the discharge pairs and consequently to determine the joint probabilities of measured discharges. Results of such analyses provide comprehensive information about flood situations where a devastating effect may be increased in the case were floods occur at the same time on the main river and its tributary. And at the same time, the results obtained by the bivariate analysis of the variables which characterize the hydrological regime can contribute to a more reliable assessment of the flood risks.

KEY WORDS: maximal mean daily discharge, annual peak discharges, corresponding discharge, joint probability distribution, copula function.

**Introduction**

Flood wave is the result of the numerous natural processes, such as rainfall distribution and intensity, catchment characteristics and its area, land use, water reservoirs, etc. The determining of the flood wave significance for design of the water management project is based on adequate mathematical technique such as statistical theory of probability. The basic mathematical technique is based on the evaluation of the selected one dimensional variable, which characterizes some extreme event. It is established that, during the flood wave occurrence there is correlation between some hydrological characteristics (discharge, duration, volume) which may have an impact on the design of flood control measures. Therefore, a multidimensional approach to analysing hydrological characteristics is increasingly preferred in many studies (e.g. Yue, 2000; Hawkes et. al., 2002; Favre et al., 2004; De Michele et al., 2005; Szolgay et al., 2016; etc.).

Regulation of rivers as well as adjustments of the river basins often brings changes in concentration of basins drainage as well as increasing of the speed of flood wave. One of a result of such interventions, it may be coincidence of the waves on the main river and its tributary. Therefore, the one dimensional approach gives satisfactory results in the case of simple systems, for example, where the main river does not capture major

tributaries. One dimensional approach may not give satisfactory results for the evaluation of flood risk in situations where floods occur on two or more rivers and join together at the same time. The bivariate statistical approach to analysis of flood events should be further developed and defined at neighbouring profiles on the main river and its tributaries.

Prohaska et al. (1999) dealt with synchronously occurring flood waves on the Danube and its tributaries. Their analysis was based on the theory of bivariate variable statistics and results confirmed that flood wave genesis is very complex within the Danube basin. In Slovak territory the coincidence of multiple flood waves caused the flood with time return period of 100-year on Tisa and Bodrog River in year 2000. For example the flood occurred in August, 2002 in Czech Republic on the Vltava River and Dyje River showed an increase in return period of the discharges with an increase in area of the basin. It was caused by coincidence of the flood waves in profiles of the river network (Report of T. G. Masaryk Water Research Institute, 2002). Chen et al. (2012) dealt with analysing the flood risk due to the correspondence of flood discharges at the main river and its tributaries using selected marginal probability distributions and multidimensional (4D) copula functions. Authors evaluated X-Gumbel copula function as appropriate for joint conditional distribution function and

return periods of joint discharges. Espinoza et al. (2013) analysed formation of the floods on the Amazonia River and its tributary. They focused on the flood occurred in 2012, when coincidence of two large flood waves has occurred. The coincidence of flood peaks from the Changjiang River and four other rivers, and the highest precipitation disaster drivers for Dongting Lake region flood vulnerability was study in Li et al. (2013). Joint probability of two random variables in contrast to the conditional probability reflects the probability with which the two random variables occur simultaneously. Gupta et al. (1976) dealt the using of the joint distribution function in estimating relationship of the large floods and their return period taking into account the seasonal influences. Bender et al. (2016) analysed flood peaks of two streams where was unlikely that all block maxima values occurred simultaneously. They concluded that in the majority of cases four marginal distributions and two distinct joint distribution functions are required to fully describe the stochastic behaviour of the system. Gilja et al. (2018) dealt with joint frequency analysis at the river confluences. The research conducted that flood hazards at the Sava River could be underestimated by traditional univariate analysis. Tadić et al. (2016) analysed the joint occurrence probability of floods on the Rivers Danube and Drava near Osijek. Their results showed that the probability of such situation is low (0.79%) but they reminded that such a situation occurred in 1966 and it was one of the biggest floods. Therefore, the results obtained by the bivariate analysis of the variables which characterize the hydrological regime can contribute to a more reliable assessment of the flood risks.

The main aim of this research is to analyse the floods occurred at the Morava River and its tributary Myjava and provide a practical approach for the designed flood estimation in areas of the rivers confluence. In the context of climatic extreme events, statistical techniques such as event coincidence analysis will be relevant for investigating the impacts of anthropogenic climate change on human societies and ecosystems globally. The results obtained by the bivariate (as well as threevariate) analysis of the variables which characterize the hydrological regime can contribute to a more reliable assessment of the flood risks.

In this research we focus upon an:

- description of the methodological approach;
- preparing of the data and identification of the variable pair combination for analysing the relationship between flood wave discharges on the main river and its tributary;
- identify univariate distribution function of the variables (discharges) of the main river and tributary;
- identify bivariate joint distribution function of the selected pair;
- compare joint exceedance probability with univariate flood frequency analysis based on the measured values on downstream of the confluence.

## Methodology

Copula functions were used as a mathematical tool for

determining a joint cumulative distribution of two dependent variables. According the Sklar (1959) theory for two dimensional (bivariate) distribution function  $H_{(x,y)}$ , we can write:

$$H_{(x,y)} = C(F_{(x)}, F_{(y)}) \quad (1)$$

where  $F_{(x)}=u$  and  $F_{(y)}=v$  are marginal distribution functions. If  $F_{(x)}$  and  $F_{(y)}$  are continuous, then the copula function  $C$  is unique.

The identification of the univariate (marginal) distribution is the first step of the bivariate analysis. The random variables may have different properties and thus need to be converted to variables having interval of  $[0, 1]$  by scaling the data. Knowing the marginal distribution, we are able to separate marginal behaviour and dependence structure. The dependence structure is fully described by the joint distribution of uniform variables obtained from marginal distribution. To determine univariate parametric distribution functions (marginal distributions), standard MLM (maximum likelihood method) method was used. According to the goodness-of-fit tests (Kolmogorov-Smirnov and  $\chi^2$ ) the marginal distributions can be selected.

In our study we used the Archimedean class of copula functions. Among existing types of copulas, the Archimedean one is the very popular class used in hydrological application (Zhang and Sing, 2006; Favre et al., 2004; De Michele et al., 2005; Gargouri-Ellouze and Eslamian, 2014; Dehghani, 2019 etc.). This class of copulas is popular in empirical applications for flexibility, easy construction and includes a whole suite of closed-form copulas that covers a wide range of dependency structures, including comprehensive and non-comprehensive copulas, radial symmetry and asymmetry, and asymptotic tail dependence and independence. The Clayton, Gumbel-Hougaard and Frank copulas were selected for this study (Table 1). The copula parameter  $\theta$  was estimated using a mathematical relationship between the Kendall's coefficient of rank correlation and the generating function  $\varphi(t)$  (Nelsen, 2006).

Testing how well a statistical model (copula in this case) describes a set of observations is discussed as a topic in literature (e.g. Kojadinovic and Yan, 2011; Karmakar and Simonovic, 2009; Shiao et al., 2010; Chowdhary et al., 2011; Genest et al., 2009; Bender 2016). According to Meylan et al. (2012) we can divide these tests into three groups: a) based on probability integral transformation; b) based on the kernel estimation of the copula density and c) based on the empirical process of copulas. The first implies conditioning on successive components of the random vector and has the drawback of depending on the order in which this conditioning is done. The second category of tests depends on various arbitrary choices, such as the kernel, the window size and the weight function, which make their application cumbersome. Several goodness-of-fit tests can be used for comparison of the empirical joint probability population and the probability population derived by parametric copulas (e.g. Kolmogorov-Smirnov, C-square,

**Table 1.** Probability functions, parameter space, generating function and relationship of non-parametric dependence measure with association parameter for the selected Archimedean copulas

Copula function	$C(u, v, \theta)$	parameter $\theta$	Kendall's $\tau$	Generator $\varphi(t)$
Clayton	$(u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$	$[-1, \infty) \setminus \{0\}$	$\frac{\theta}{\theta + 2}$	$\frac{1}{\theta}(t^{-\theta} - 1)$
Gumbel-Hougaard	$\exp[-((-\ln u)^\theta + (-\ln v)^\theta)^{1/\theta}]$	$[1, \infty)$	$\frac{\theta - 1}{\theta}$	$(-\ln t)^\theta$
Frank	$-\frac{1}{\theta} \ln[1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{(e^{-\theta} - 1)}]$	$(-\infty, \infty) \setminus \{0\}$	$1 + \frac{4}{\theta}[D_1(\theta^*) - 1]$	$-\ln \frac{e^{-\theta} - 1}{e^{-\theta} - 1}$

Anderson-Darling or Cramér-von-Mises).

The empirical probability (Gringorten, 1963; Cunnane, 1978; Yue et al., 1999; Zhang and Singh, 2006) represents Equation (2).

$$F(x, y) = P(X \leq x_i, Y \leq y_j) = \frac{\sum_{m=1}^i \sum_{l=1}^j n_{ml} - 0.44}{N + 0.12} \quad (2)$$

where  $N$  is the total number of the variables,  $j$  and  $i$  ascending ranks of  $x_i$  and  $y_i$ ,  $n_{ml}$  is the number of occurrence of the combinations of  $x_i$  and  $y_j$ .

In hydrological frequency analysis, the return period of the hydrological variable that occurs once in a year can be defined as:

$$T = \frac{1}{P(X \leq x)} = \frac{1}{\lambda(1 - F(x))} \quad (3)$$

where  $T$  is the return period in years and  $F(x)$  is univariate cumulative distribution function and  $\lambda$  represents average number of events per year. In frequency analysis of annual maximum value of  $\lambda$  is equal (1).

In multivariate statistical analysis, we can determine the return period of the phenomenon in two ways. The first is a joint return period, while the second is a conditional return period. The first one defines joint return periods as: the return periods using one random variable equalling or exceeding a certain magnitude and/or using another random variable equalling or exceeding another certain magnitude. The second one is conditional return period for one random variable, given that another random variable equals or exceeds a specific magnitude.

Joint return period for two variables defined by more authors (Shiau, 2003; Salvadori and De Michele, 2006) and it can be written in the form of:

$$T^{and}_{x,y} = \frac{1}{\lambda(1 - F(x) - F(y) + H(x, y))} \quad (4)$$

or

$$T^{or}_{x,y} = \frac{1}{\lambda(1 - H(x, y))} \quad (5)$$

Equation (4) represents the joint return period of  $X \geq x$

and  $Y \geq y$ . Equation (5) represents joint return period of  $X \geq x$  or  $Y \geq y$ . These relationships indicate, that different combinations of the numbers  $x$  and  $y$ , can take same return period (relation 8).  $H(x, y)$  is the joint cumulative distribution function (can be expressed as copula function).

$$T^{or}_{(x,y)} \leq \min[T_x, T_y] \leq \max[T_x, T_y] \leq T^{and}_{(x,y)} \quad (6)$$

Conditional return period for  $X$ , given  $Y \geq y$  may be expressed as (Shiau, 2003):

$$T_{(x|Y \geq y)} = \frac{1}{\lambda(1 - F(y))(1 - F(x) - F(y) + H(x, y))} \quad (7)$$

where  $x$  and  $y$  are random variables and  $H(x, y)$  is the joint cumulative distribution function. Conditional cumulative distribution function of  $X$ , given  $Y \geq y$  may be expressed as:

$$F_{(x|Y \geq y)} = \frac{F(y) - H(x, y)}{1 - F(y)} \quad (8)$$

where  $H(x, y)$  is the joint cumulative distribution function of the random variables  $X$  and  $Y$ , and  $F(y)$  is cumulative distribution function of the variable  $Y$ . An equivalent formula for conditional return period of  $Y \geq y$ , given  $X \leq x$  can be obtained.

## Study area

The paper is focused on the bivariate analysis of the floods occurred at the main river and its tributary for the designed flood estimation in areas of the rivers confluence.

The Morava is a left tributary of the Danube River in Central Europe. The length of the Morava River is 329 km and its basin area covers 26 579.7 km<sup>2</sup>. The river originates on the Králický Sněžník mountain near the border between the Czech Republic and Poland and has a vaguely southward trajectory. The lower part of the river's course forms the border between the Czech Republic and Slovakia and then between Austria and Slovakia.

The Myjava River is a river in western Slovakia and for a small part in the Czech Republic and left tributary of the Morava River. The length of the Myjava River is 79 km and its basin area is 806 km<sup>2</sup>. It rises in the White Carpathians near the village of Nová Lhota in Moravia, but crosses the Czech-Slovak border shortly afterwards and flows in a southern direction until the town of Myjava, where it enters the Myjava Hills and turns west. Near Sobotište it flows into the Záhorie Lowland and turns south until the village of Jablonica, turning north-west and from Senica it flows west, passing through Šaštín-Stráže and finally flowing into the Morava River near Kúty. Table 2 lists selected main river, tributary, gauging stations and measurement period. The first year of the analysed period is the beginning of data measurements in current stations. The scheme of the selected rivers confluence is presented in Fig.1.

### Preparing of the data

With regard to flood coincidence analysis, it is necessary to consider gauging stations immediately downstream and upstream from the tributary. In our work we investigated combination of the first pairs of  $Q_{amaxup} - Q_{amaxtr}$  and second pairs of  $Q_{maxup} - Q_{cor1tr}$ .

Where:

$Q_{amaxup}$ ,  $Q_{amaxdwn}$ ,  $Q_{amaxtr}$  – are annual peak discharges on upstream and downstream on main river and the annual maximum discharges on tributary;

$Q_{maxdwn}$  – is maximum daily discharge on the main river downstream from the confluence (same event like in upstream station);

$Q_{maxup}$  – is maximum daily discharge on the main river upstream from the confluence;

$Q_{cor1tr}$  – is corresponding discharge on the tributary in the moment of occurrence of the maximum daily discharge on the main river upstream from the confluence (may be moved one or two days).

Discharges time series of the annual peak discharges Morava and Myjava reveal that high annual peaks at the Myjava River do not frequently coincide with high annual peaks at Morava River (Fig. 2).

The analysis of flood hazard is based on the continuous record of discharges on a gauging station that reflects water regime in particular section. As the first we analysed pairs of annual peak discharges  $Q_{amaxup} - Q_{amaxtr}$ . In this study only the combination of the values of annual peak discharge were analysed not they occurrence within the year. Selected and analysed pairs of annual peak discharges Morava – Myjava confluence are presented in Fig. 3a. The linear Pearson correlation coefficient of this pairs is  $R=0.63$  and Kendall rank correlation coefficient is  $\tau=0.42$ .

The annual maximum (AM) series approach is the most frequently used in probabilistic hydrology. But this data series are limited by two factors: 1) the length of the series of annual maxima can be very short and 2) the annual maxima time-series may be interrupted and thus they may not allow us to infer the antecedent conditions in the basin preceding a given peak. The first limiting factor produces uncertainties in interpreting statistical analyses, while the latter constrain implies that statistical models built on a phenomenological basis must rely on ancillary data in order to validate the underlying hypotheses (Claps and Lio, 2003). This situation is avoided in the Peaks Over Threshold method (POT). Data series of the POT method consider all values exceeding a certain predefi-

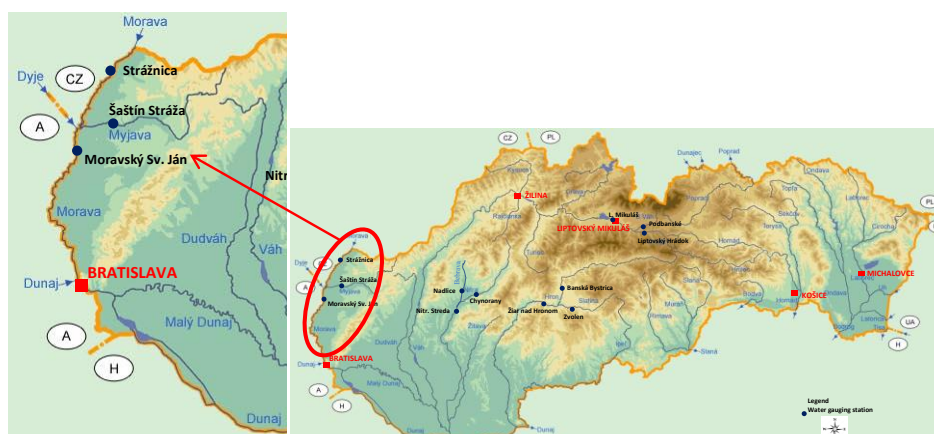


Fig. 1. The map of the Morava and Myjava confluence.

Table 2. Selected main river and tributary, gauging stations and analysed periods

River	Gauge station	River kilometer [km]	$Q_{amax}$ [m <sup>3</sup> s <sup>-1</sup> ]	period
Morava	Strážnica (CZ)	133.5	901	1968–2011
	Moravský Svätý Ján (SK)	67.15	1502	1968–2011
Myjava	Šaštín Stráža (SK)	15.18	82	1968–2011

ned threshold (Bayliss, 1999; Rao and Hamed, 2000). The POT method has been proposed as an alternative analytical tool to the AM approach in analysis of extreme hydrological events. This method was discussed in a number of papers (Langbein, 1949; Todorovic, 1970; Cunneane, 1973; Rosbjerg, 1977; Madsen et al., 1997a; Madsen et al., 1997b; Lang et al., 1999; Bača and Bačová Mitkova, 2007; Bačová Mitková and Onderka, 2010). Therefore, in practice it seems to be more meaningful to consider not only the annual discharge maxima but also flood events that exceed safety limits. The first threshold can be chosen near the long-term mean discharge. This value is rather low; POT series can have high diffusion and can include some insignificant maxims. Therefore, a threshold value is usually chosen so that POT data series includes in average 4 maximum values per year

and flood events must be independent (Bayliss, 1999). The threshold value in our calculations was appointed at the level of 40–50% of the long-term maximum annual discharge in order to ensure the independence of the waves and to include all significant events in the analysed year. On the basis of daily discharges and POT method the flood waves were selected for bivariate statistical analysis. The maximum mean daily discharges and synchronous discharges according to the above mentioned scheme were selected.

The pairs of maximum daily discharges and corresponding daily tributary discharges of the Morava – Myjava confluence based on POT method are presented in Fig. 3b. The linear the Pearson correlation coefficient of this pairs are  $R=0.57$  and Kendall rank correlation coefficient is  $\tau=0.37$ .

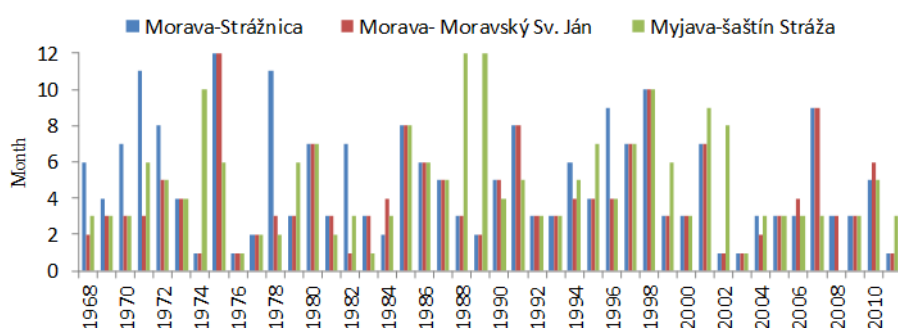


Fig. 2. Occurrence of the annual peak discharges during within the year (1968–2011).

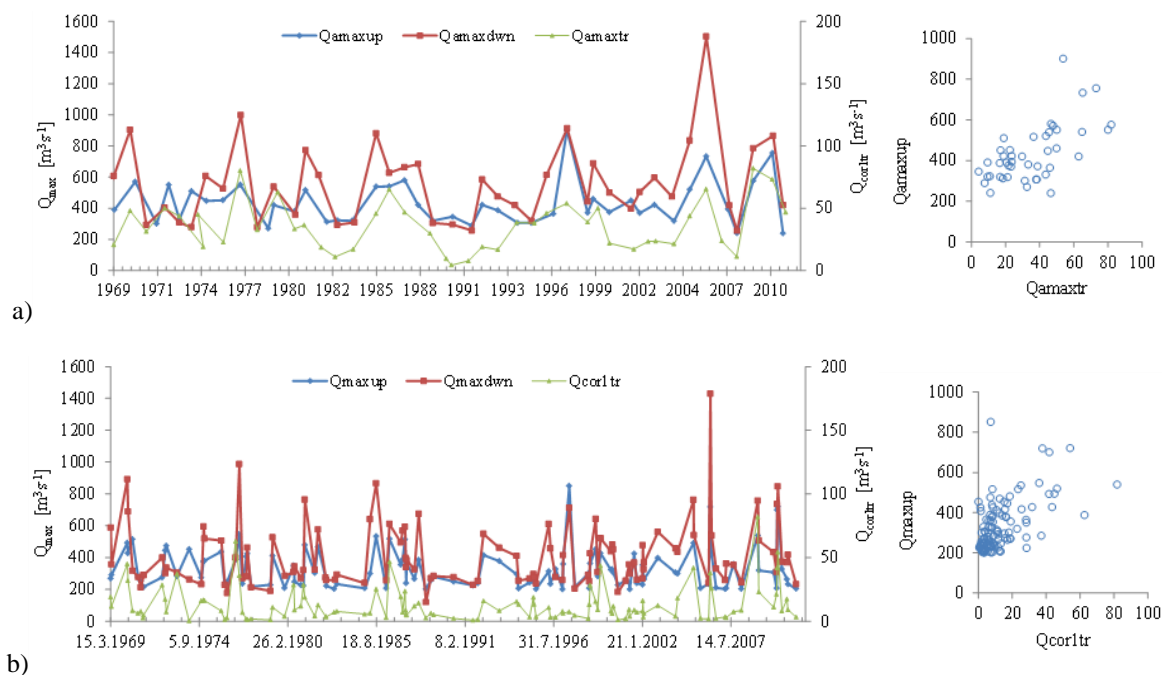


Fig. 3. Selected and analysed pairs of a) annual peak discharges Morava – Myjava confluence and b) maximum daily discharges and corresponding daily discharge Morava – Myjava confluence (based on POT method).

The results of the correlation analysis of different combinations of the variables (discharge) show a statistically significant correlation. All of the combinations of the variables can be used on the following bivariate joint frequency analysis to investigate, how the relationship of the hydrological characteristics may affect the size and course of extreme hydrological situations.

### Univariate Analysis

In order to determine univariate parametric distribution functions (marginal distributions), standard MLM (maximum likelihood method) method was used. According to the goodness-of-fit test (Kolmogorov-Smirnov) the marginal distributions were selected.

The variables of the annual maximum (AM) approach were preferable fitted with Gumbel, Gamma and LogPearson III. The Gumbel distribution is asymmetric, extreme value distribution (EVD) and is used to model the distribution of the maximum (or the minimum). The Gamma two-parametric distribution is a very important model in statistical hydrology. It is a flexible function capable taking many different shapes and has been widely used in many countries for flood series modelling. The Pearson type III distribution is sometimes called three-parameter Gamma distribution, since it can be obtained from the two-parameter Gamma distribution by introducing location parameter  $\tau$ . It is very flexible since it has three parameters which can produce a wide variety of shapes of density function.

The variables derived by POT method was preferably fitted with distributions JohnsonSB, Gamma and LogPearson III. The JohnsonSB distribution is a continuous four-parametric distribution defined on bounded range, and the distribution can be symmetric or asymmetric. The Kolmogorov-Smirnov test was performed to test the assumption that the discharge magnitudes follow the theoretical distributions. The  $p$ -value ( $p \geq 0.05$ ) was used as a criterion for rejection of the proposed distribution hypothesis. The fitted distributions,  $p$ -value of the goodness-of-fit test and calculated designed discharges are presented in Table 3. We cannot reject the hypothesis that selected distributions fit well to the observed data at a 5% significance level. Comparison of the AM and POT approaches shows significant differences (in average 13%) for discharges with return period

of 50 and 100 years on tributary. The lower estimated values of designed discharges may be result of the creation of the analysed POT data series, especially for tributary where are corresponding discharges with  $Q_{maxup}$  and not only  $Q_{maxtr}$ . It effects the estimation of  $Q_N$ .

### Joint bivariate analysis of the discharges at the rivers confluence using copulas

The joint probability distribution of two hydrologic variables at the confluence was evaluated by copulas to calculate the joint exceedance probability for the analysis of the flood hazard occurred at the confluence of the main river and its tributary. Results of the correlation show that there is a strong positive dependence between the discharges at river confluence with values of Kendall's coefficient of rank correlation ranked over 0.3. The values of the estimated parameters of selected Archimedean copula functions are listed in Table 4.

Results of the comparison of the joint empirical probabilities with fitted parametric copula showed that computed errors of the estimation reached relatively small differences between all three tested Archimedean copula functions (Fig. 4). The Kolmogorov-Smirnov test was performed to test the assumption that the joint pair magnitudes follow the theoretical joint distributions (copula). The  $p$ -value ( $p \geq 0.05$ ) was used as a criterion for rejection of the proposed distribution hypothesis. The fitted joint distributions (copula) and  $p$ -value of the goodness-of-fit test are presented in Table 4. We cannot reject the hypothesis that selected joint distributions fit well to the observed data at a 5% significance level. Based on the results of non-parametric test the all three Archimedean copula functions were used to determine the joint probability distribution of the pair variables.

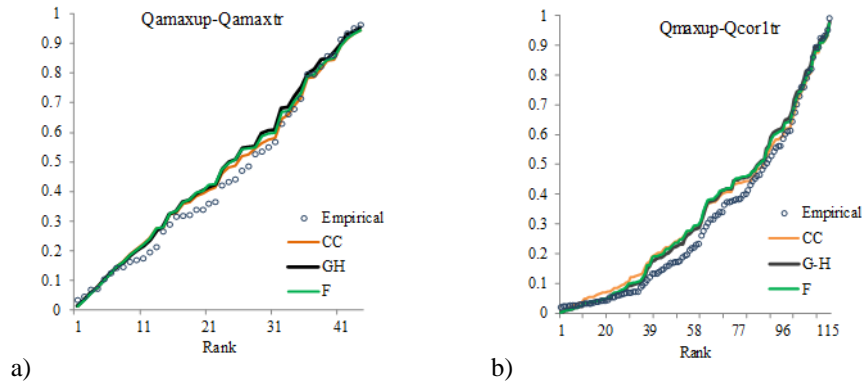
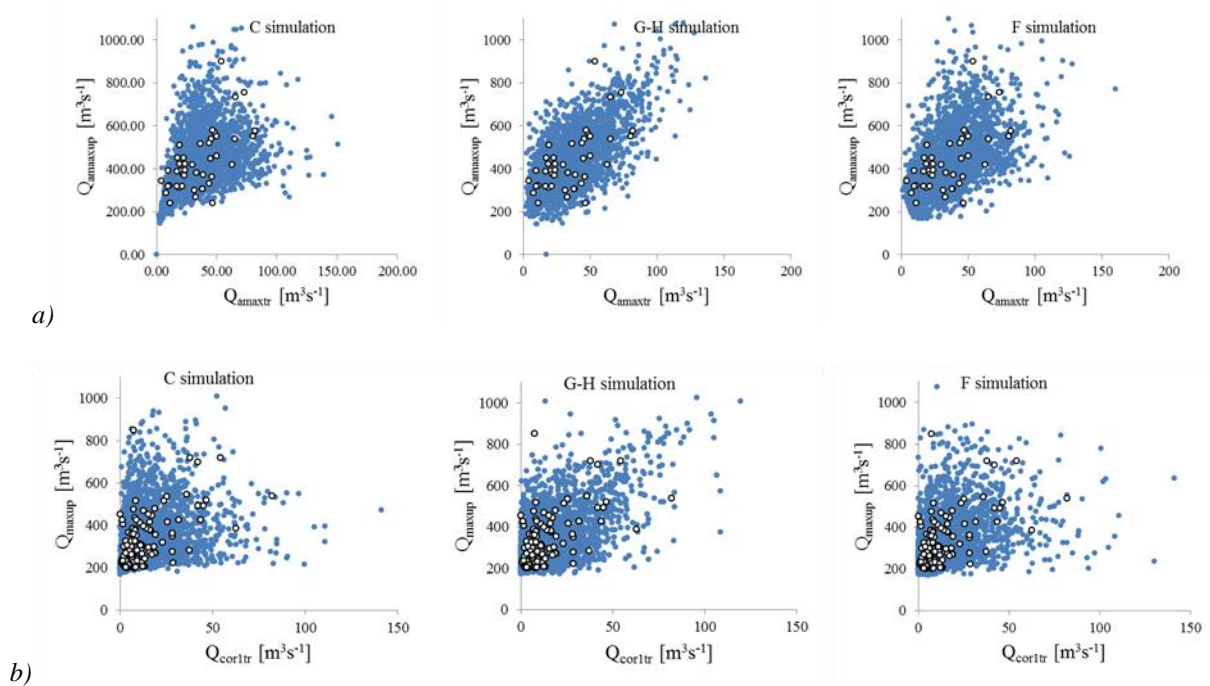
Subsequently, the all selected copula functions were used for simulation of 3000 pairs for all combination of variables on selected river (Fig. 5). The figures show the scatter plot of measured data pairs and simulated values generated from copula models for the discharge pairs at the confluence. Simulated pairs were performed to determine the joint probability distribution using copulas and consequently to determine the joint occurrence probability of the variables. For traditional method, the resulting discharge for the selected return period is

**Table 3. Best fitted parametric univariate distributions of the variables and  $p$ -values of the Kolmogorov-Smirnov test ( $\alpha=0.05$ )**

Confluence	$Q$ [m <sup>3</sup> s <sup>-1</sup> ]	$\lambda$	Distr.	$p$ -value	Estimated $Q_N$ [m <sup>3</sup> s <sup>-1</sup> ]					Measured $Q_{max}$ [m <sup>3</sup> s <sup>-1</sup> ]	
					$Q_{50}$	$Q_{100}$	$Q_{200}$	$Q_{500}$	$Q_{1000}$	$Q_{amax}$ $Q_{max}$	$T$ [year]
Morava Myjava	$Q_{amaxup}$	1	Gumbel	0.96	791	867	942	1041	1116	901	137
	$Q_{amaxtr}$	1	Gamma	0.8	87	97	107	119	129	82	37
	$Q_{amaxdwn}$	1	LogPear III.	0.51	1293	1489	1699	1999	2246	1506	106
	$Q_{maxup}$	2.6	Jon SB	0.48	798	863	920	983	1025	850	96
	$Q_{cor1tr}$	2.6	Gamma	0.51	77	86	98	112	119	82	74
	$Q_{maxdwn}$	2.6	LogPear III.	0.55	1223	1416	1624	1942	2190	1430	105

**Table 4.** Copula parameters (C – Clayton, G-H – Gumbel-Hougaard, F – Frank), selected combinations of the variables

Confluence	pair	$R_{meas}$	$\tau_{meas}$	parameter		
				Clayton	Gumbel-Hougaard	Frank
Morava – Myjava	$Q_{amaxup} - Q_{amaxtr}$	0.63		1.45	1.7	4.4
			$R_{sim}$	0.46	0.68	0.54
			$\tau_{sim}$	0.42	0.43	0.42
	$Q_{maxup} - Q_{corltr}$	0.57		1.2	1.6	3.9
			$R_{sim}$	0.35	0.62	0.42
			$\tau_{sim}$	0.36	0.38	0.36

**Fig. 4.** Comparison of the joint empirical with fitted parametric copula probabilities Morava – Myjava confluence a) annual peak pairs and b) daily maximum and corresponded pairs.**Fig. 5.** Scatter plots of 3000 data pairs generated from copulas (C – Clayton, G-H – Gumbel-Hougaard and F – Frank) and measured data of for selected combinations a)  $Q_{amaxup} - Q_{amaxtr}$  Morava – Myjava, b)  $Q_{maxup} - Q_{corltr}$  Morava – Myjava.



calculated as reciprocal of the probability of exceedance. The discharge calculated using the copulas can lie anywhere on the isoline representing return period and can have infinite combinations. Since the worst-case scenario is regarding flood hazard at confluences, the extreme value of combined discharges was calculated, i.e. the maximum discharge resulting from infinite combinations of discharges at the Morava River and its tributary Myjava River. The comparison of the characteristic designed discharges ( $Q_N$ ) calculated using traditional univariate distribution and bivariate copula distribution at annual peak discharges and maximum daily discharges are presented in Table 6.

## Conclusions

The paper presents the bivariate joint frequency analysis of the discharges at the main river and its tributary. The annual peak discharges and mean daily discharges of the Morava and Myjava confluence were analysed (period of 1968–2011). For analysis the two statistical approaches were used. The first was traditional univariate approach and the second one was the bivariate joint distribution approach using the copulas.

The results in this paper show that:

- The type of theoretical probability distribution as well as the type of used input data series significantly affects the estimated  $Q_N$ .
- Comparison of the AM and POT approaches shows significant differences (in average 13%) for discharges with return period of 50 and 100 years on tributary.
- Comparison of the AM and POT univariate approaches with selected marginal distributions didn't show

significant differences for estimated maximum discharges with return period up to 100 years.

- All tested Archimedes copula function achieved relatively small differences between calculated errors of estimation.
- According to the KS test we cannot reject any of tested Archimedean copula function that the joint pair magnitudes follow the theoretical joint distributions.
- Discharges estimated from Gumbel-Hougaard copula are higher than the ones calculated using the univariate method about 12%. Gumbel-Hougaard copula can well describe the multivariate relationship and improve the abnormal crossing phenomena, so it can give more reasonable results and be further applied to hydrological extreme analysis.

In conclusion, we can say, that the selection of the distribution function to estimate  $T$ -year discharges and also the processing of the statistical data series affects the results of the estimation. Determining the specific value of a 500- or 1000-year flood for engineering practice is extremely complex and in interpreting the results, it should be kept in mind that estimated values with very high return periods are extrapolated values. Each statistical method includes some uncertainty that may be caused by the method but also the data may be affected by certain measurement error therefore, is also necessary to specify confidence intervals in which the flow of a given 100-, 500-, or 1000-year flood may occur with probability, for example, 90%. Results of the bivariate analysis showed that bivariate copula model can be successfully applied at locations where significant change in flow regime is present, or flood intensity is governed by several variables, such as at river confluence.

**Table 5. Evaluation of Archimedean copula functions for computing the joint distribution function by Kolmogorov-Smirnov test ( $\alpha = 0.05$ )**

Confluence	pair	Clayton	Gumbel-Hougaard	Frank
		$p$ -value	$p$ -value	$p$ -value
Morava – Myjava	$Q_{amaxup} - Q_{amaxtr}$	0.93	0.81	0.8
	$Q_{maxup} - Q_{cor1tr}$	0.37	0.56	0.46

**Table 6. Comparison of the characteristic discharges based on univariate and copula method for gauging station Moravský Svätý Ján of the Morava River**

Confluence	$Q$ [m <sup>3</sup> s <sup>-1</sup> ]	Univariate $Q_{Ndown}$	Copula Estimated $Q_N$ [m <sup>3</sup> s <sup>-1</sup> ]			Differences [%]			
			C	G-H	F	C	G-H	F	
Morava – Myjava	$Q_{amaxdown}$	$Q_{50}$	<b>1293</b>	1310	<b>1496</b>	1315	1.3	<b>+13.6</b>	1.7
		$Q_{100}$	<b>1489</b>	1496	<b>1704</b>	1502	0.5	<b>+12.6</b>	0.9
		$Q_{500}$	<b>1999</b>	2003	<b>2251</b>	2004	0.2	<b>+11.2</b>	0.2
		$Q_{1000}$	<b>2246</b>	2248	<b>2513</b>	2248	0.1	<b>+10.6</b>	0.1
	$Q_{maxdown}$	$Q_{50}$	<b>1223</b>	1229	<b>1386</b>	1229	0.5	<b>+11.8</b>	0.5
		$Q_{100}$	<b>1416</b>	1419	<b>1600</b>	1419	0.2	<b>+11.5</b>	0.2
		$Q_{500}$	<b>1924</b>	1942	<b>2170</b>	1942	0.9	<b>+11.3</b>	0.9
		$Q_{1000}$	<b>2190</b>	2192	<b>2462</b>	2210	0.1	<b>+11.0</b>	0.9



ces. At river confluences, marginal distributions of inflow discharges rarely follow similar distribution, making copula model especially suitable for flood hazard assessment. When analysing the complex stochastic character of a river system on streams where the extreme event does not occur nearly simultaneous and where the tributary contributes significantly to the main river is necessary to analyse also pairs  $Q_{maxtr} - Q_{corup}$  (Bender et al., 2016).

### Acknowledgements:

*This work was supported by the project VEGA No. 2/0004/19 and it results from the project implementation of the ITMS 26210120009 Infrastructure completion of hydrological research stations, supported by the Research & Development Operational Programme funded by the ERDF.*

### References

- Bača, P., Bačová Mitková, V. (2007): Analysis of seasonal extreme flows using Peaks Over Threshold method. *Journal of Hydrol. and Hydromech.* Vol.55, No.1, 16–22.
- Bačová Mitková, V., Onderka, M. (2010): Analysis of extreme hydrological events on the Danube using the Peak Over Threshold method. In *Journal of Hydrology and Hydromechanics*, Vol. 58, no. 2, 88–101.
- Bayliss, A. C. (1999): Deriving flood peak data. *Flood estimation: Handbook*. Vol.3, 273–283.
- Bender, J., Wahl, T., Müller, A., Jensen, J. (2016): A multivariate design framework for river confluences. *Hydrologica. Science Journal-Journal des Sciences Hydrologiques.*, 61(2016) 4, 367–381
- Claps, P., Lio, F. (2003): Can continuous stream flow data support flood frequency analysis? An alternative to the partial duration series approach. *Water Resour. Res.*, 39 (8), p. 1216.
- Cunnane, C. (1973): A particular comparison of annual maximum and partial duration series methods of flood frequency prediction. *Journal of Hydrology*, 18, 257–271.
- Cunnane, C. (1978): Methods and merits of regional flood frequency analysis. *J. Hydrol.*, 100, 269–290.
- Dehghani, M.; Saghaian, B.; Zargar, M. (2019): Probabilistic hydrological drought index forecasting based on meteorological drought index using Archimedean copulas *Hydrology Research* (2019) 50 (5): 1230–1250. <https://doi.org/10.2166/nh.2019.051>
- De Michele, C. et al. (2005): Bivariate statistical approach to check adequacy of dam spillway. *J. Hydrol. Eng.*, 10 (1), 50–57.
- Espinoza, J. C. et al. (2013): The Major Floods in the Amazonas River and Tributaries (Western Amazon Basin) during the 1970–2012 Period: A Focus on the 2012 Flood. *Journal of Meteorology*, Vol. 14, 1000–1008. doi: 10.1175/JHM-D-12-0100.1.
- Favre, A. C. et al. (2004): Multivariate hydrological frequency analysis using copulas. *Water Resour. Res.*, 40, W01101, 1–12.
- Gargouri-Ellouze, E., Eslamian, S. (2014): Application of Copulas in Hydrology: Geomorphological Instantaneous Unit Hydrograph and Intensity Index of Infiltration Frequency, in *Handbook of Engineering Hydrology*, Ch. 1, Vol. 2: Modelling, Climate Changes and Variability, Ed. By Eslamian, S., Francis and Taylor, CRC Group, USA, 1–18.
- Genest, C. et al. (2009): Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics*, 44, 199–213. doi:10.1016/j.insmatheco.2007.10.005.
- Gilja, G., Ocvirk, E., Kuspilić, N. (2018): Joint probability analysis of flood hazard at river confluences using bivariate copulas. *Journal of the Croatian Association of Civil Engineers* (0350-2465). 70. 267–275. Doi:10.14256/jce.2173.2017.
- Gringorten, I. I. (1963): A plotting rule for extreme probability paper. *J. Geophys. Res.*, 68(3), 813–814.
- Gupta, V. K. et al. (1976): On the joint distribution of the largest flood and its time of occurrence, *Water Resour. Res.*, 12(2), 295–304, doi:10.1029/WR012i002p00295.
- Hawkes P. J. et al. (2002): The joint probability of waves and water levels in coastal engineering design. *Journal of Hydraulic Research*, 40(3), 241–251.
- Chen, L. et al. (2012): Flood Coincidence Risk Analysis Using Multivariate Copula Functions. *J. Hydrol. Eng.*, 10.1061/(ASCE)HE.1943-5584.0000504, 742–755.
- Chowdhary, H. et al. (2011): Identification of suitable copulas for bivariate frequency analysis of flood peak and flood volume data. *Hydrol. Res.*, 42(2-3), 193–215.
- Karmakar, S., Simonovic, S. P. (2009): Bivariate flood frequency analysis. Part 1: Determination of marginals by parametric and nonparametric techniques. *J. Flood Risk Manag.*, 1, 190–200.
- Kojadinovic, I., Yan, J. (2011): A Goodness-of-Fit Test for Multivariate Multiparameter Copulas Based on Multiplier Central Limit Theorems. *Statistics and Computing*, 21(1), 1–26.
- Lang, M., et al. (1999): Towards operational guidelines for over threshold modelling. *J. Hydrol.*, 225, 103–117.
- Langbein, W. B. (1949): Annual floods and the partial-duration flood series. *Transactions, AGU* 30 (6), 879–881.
- Li, C. H., Li1, N. et al. (2013): A relative vulnerability estimation of flood disaster using data envelopment analysis in the Dongting Lake region of Hunan. *Nat. Hazards Earth Syst. Sci.*, 13, 1723–1734. doi:10.5194/nhess-13-1723-2013
- Madsen, H. et al. (1997a): Comparison of annual maximum series and partial duration series methods for modelling extreme hydrologic events: 1. At-site modelling, *Water Resour. Res.*, 33(4), 747–757.
- Madsen, H. et al. (1997b): Comparison of annual maximum series and partial duration series methods for modelling extreme hydrologic events: 2. Regional modelling, *Water Resour. Res.*, 33(4), 759–769.
- Meylan, P. et al. (2012): Predictive hydrology: a Frequency Analysis Approach. Boca Raton, FL: CRC Press; 212 p.
- Nelsen, R. B. (2006): An introduction to Copula. The 2<sup>nd</sup> edition. Springer, New York.
- Prohaska, S. et al. (1999): Coincidence of flood flow of the Danube River and its tributaries. A Hydrological monograph of the Danube River Basin. Belgrade University.
- Rao, A. R., Hamed, K. H. (2000): Flood frequency analysis. CRC Press LLC, N. W. Corporate Blvd., Boca Raton, Florida.
- Report of T. G. MASARYK WATER RESEARCH INSTITUTE (2002): Assessment of the catastrophic floods in August 2002 and proposal for adaptation of the flood prevention system, Final report on the 2002 project, Prague, pp. 174. (In Czech)
- Rosbjerg, D. (1977): Returns periods of hydrological events. *Nordic Hydrol.*, 8, 57–61.

- Salvadori, G., De Michele, C. (2006): Frequency analysis via copulas: theoretical aspects and applications to hydrological events. *Water Resour Res.* 40, W12511.
- Shiau, J. T. (2003): Return period of bivariate distributed extreme hydrological events. *Stoch. Envir. Res. Risk Assess.*, 17, 42–57.
- Shiau, J. T. et al. (2010): Copula-based depth-duration-frequency analysis of typhoons in Taiwan. *Hydrol. Res.*, 41(5), 414–422.
- Sklar, A. (1959): Fonction de répartition à  $n$  dimensions et leurs marges. *Publications de L'Institut de Statistique, Université de Paris*, 229–231.
- Szolgay, J. et al. (2016): A regional comparative analysis of empirical and theoretical flood peak-volume relationships. *J. Hydrol. Hydromech.* Vol. 64(4), 367–381. doi: 10.1515/johh-2016-0042.
- Todorovic, P. (1970): On some problems involving random number of random variables. *Ann. Math. Statistics*, Vol. 41(3), 1059–1063.
- Tadić, L. et al. (2016): Analysis of the Drava and Danube rivers floods in Osijek (Croatia) and possibility of their coincidence. *Environ Earth Sci*, 75:1238, p. 12. <https://doi.org/10.1007/s12665-016-6052-0>
- Yue, S. et al. (1999): The Gumbel mixed model for flood frequency analysis. *J. Hydrol.* 226(1–2), 88–100.
- Yue, S. (2000): Joint probability distribution of annual maximum storm peaks and amounts as represented by daily rainfalls. *Hydrological Science Journal*, Vol. 45 (2), 315–326.
- Zhang, L., Singh, V. P. (2006): Bivariate flood frequency analysis using the copula method. *J. Hydrol. Eng. ASCE*, 11(2), 150–164.

Ing. Veronika Bačová Mitková, PhD. (\*corresponding author, e-mail: mitkova@uh.savba.sk)  
Institute of Hydrology SAS  
Dúbravská cesta 9  
84104 Bratislava  
Slovak Republic