

Primljen / Received: 13.7.2017.

Ispravljen / Corrected: 22.2.2018.

Prihvaćen / Accepted: 10.3.2018.

Dostupno online / Available online: 10.5.2018.

# Joint probability analysis of flood hazard at river confluences using bivariate copulas

## Authors:



Assist.Prof. **Gordon Gilja**, PhD. CE  
University of Zagreb  
Faculty of Civil Engineering  
[ggilja@grad.hr](mailto:ggilja@grad.hr)



Assoc.Prof. **Eva Ocvirk**, PhD. CE  
University of Zagreb  
Faculty of Civil Engineering  
[ocvirk@grad.hr](mailto:ocvirk@grad.hr)



Prof. **Neven Kuspilić**, PhD. CE  
University of Zagreb  
Faculty of Civil Engineering  
[kuspa@grad.hr](mailto:kuspa@grad.hr)

Original scientific paper

**Gordon Gilja, Eva Ocvirk, Neven Kuspilić**

## Joint probability analysis of flood hazard at river confluences using bivariate copulas

Estimation of flood hazard associated with return period of a hydrologic event is the basis for design of flood protection structures. More frequent occurrence of flood events in recent history has imposed the need to reconsider traditional hydrological approaches to estimation of high flow events. This paper focuses on estimation of peak flood discharge at two confluences of the Sava River. The results show that the peak discharge trend estimated using bivariate copulas is comparable to the values measured during an extreme flood event in 2014.

### Key words:

flood hazard, flood protection, bivariate copulas, Sava River, confluence

Izvorni znanstveni rad

**Gordon Gilja, Eva Ocvirk, Neven Kuspilić**

## Analiza zajedničke vjerojatnosti pojave velikih voda na ušćima primjenom bivarijantnih kopula

Procjena opasnosti od poplava za hidrološki događaj zadanog povratnog razdoblja jest temelj projektiranja građevina namjenjenih sustavu obrane od poplava. Učestalo pojavljivanje velikih voda u kratkom razdoblju upozorava kako treba preispitati tradicionalne pristupe procjenjivanju vjerojatnosti njihovog pojavljivanja. U ovom radu procijenjena je opasnost od istovremene pojave velike vode na rijeci Savi i njezinom pritoku za dva ušća. Rezultati pokazuju da se trend protoka procijenjenih bivarijantnom kopulom poklapa s mjerenjima protoka velike vode 2014. godine.

### Ključne riječi:

opasnost od poplava, obrana od poplava, bivarijantne kopule, rijeka Sava, ušće

Wissenschaftlicher Originalbeitrag

**Gordon Gilja, Eva Ocvirk, Neven Kuspilić**

## Analyse der gemeinsamen Wahrscheinlichkeiten des Auftretens von Hochwasser an Mündungen durch Anwendung bivariater Copula

Die Einschätzung der Hochwassergefahr für hydrologische Ereignisse des festgelegten Rückkehrzeitraums ist die Grundlage für die Planung von Gebäuden, die für das Hochwasserschutzsystem gedacht sind. Das häufige Auftreten von Hochwasser in einem kurzen Zeitraum weist darauf hin, dass die traditionellen Ansätze der Einschätzung der Wahrscheinlichkeit ihres Auftretens überprüft werden müssen. In dieser Abhandlung wird für zwei Mündungen die Gefahr eines gleichzeitigen Auftretens von Hochwasser am Fluss Save und ihrem Nebenfluss eingeschätzt. Die Ergebnisse zeigen, dass sich der Trend des Durchflusses, eingeschätzt anhand bivariater Copula, mit den Messungen des Durchflusses des Hochwassers von 2014 deckt.

### Schlüsselwörter:

Hochwassergefahr, Hochwasserschutz, bivariate Copula, Fluss Save, Mündungen

## 1. Introduction

Flood wave is the result of interaction of numerous variables such as rainfall intensity, duration and spatial distribution of rainfall over catchment area, catchment characteristics and land use, reservoir operation, etc. Although flood peak is just one of flood wave characteristics, it is considered to be the most important from the flood protection aspect and, therefore, it is often associated with flood hazard assessment in standard hydrological design based on univariate frequency analysis methods [1]. Hydrological analyses are conducted to obtain design variables that are used in structural analysis and in the assessment of functionality in the design process. Multivariate nature of flood event implies that simultaneous exceedance of critical thresholds for the correlated variables needs to be taken into account for the description of its characteristics, so as to provide a complete assessment of the probability of exceedance for design flow [2]. Bivariate and multivariate probability analyses for calculating variables relevant for flood hazard under conditional distribution have been used since 1980s in many studies. The results obtained by such methods differ greatly from those obtained using univariate methods [3]. It was established that, during occurrence of a flood wave, a correlation can be made between its maximum flow rate (flood peak discharge), volume and duration, which may affect design requirements in a number of way, depending on the planned use of a particular hydraulic structure [4]. Joint probability distributions of flood peak discharge and volume have been used for the design of flood control reservoirs [2], for estimating parameters for a synthetic design hydrograph [1, 5], for the design of flood control systems [6, 7], and for dam design [8]. The joint probability distribution of flood peak discharge of two rivers is used to model simultaneous occurrence of flood waves in complex river systems [9].

The bivariate flood frequency analysis results in an infinite number of combinations of flood controlling variables along the isoline representing their joint return period. From a practical point of view, bivariate approach provides extensive information on the possible outcome of a hydrologic event, while only specific subset of this information is relevant to the problem that needs to be analysed. Combinations of values closer to the edges of the isoline are less likely to occur because an increase in one variable is compensated by a decrease of another, due to low probability of them coinciding [1]. For practical use, Volpi and Fiori proposed a methodology for the identification and extraction of several combinations within the subset on the isoline, depending on the hydrological loads relevant for a specific structure [2]. The bivariate frequency analysis of flood peak discharge, joined with other hydrologic variables like flood duration or volume, is mostly used for reservoir design [4, 10], while the flood volume does not show a significant correlation with flood peak discharge in river reaches where water regime reflects hydrological cycle duration [6]. It is challenging to define flood hazard for reaches near tributary confluences where

each river contributes to overall flood hazard by introducing backwater effect, thus making determination of discharge rating curve unreliable because the discharge cannot be simply translated into water levels. The flood frequency analysis at the river confluence is of interest because it is necessary for the design of the levee system, the height of which has to be increased to ensure an undisturbed passage of flood wave in case of coincident flood levels at the main river and its tributary. An accurate and practical approach to flood frequency estimation on river confluences is needed, and this approach has to involve calculation of the joint probability distribution of coincident flood peak discharges at the confluent river pair.

The flood frequency analysis of coincident floods at the Sava River and its tributaries is used in this paper to calculate the joint probability distribution at the tributary confluence based on the measured discharge data. The estimation of the joint probability distribution of coincident floods is based on the identified marginal distributions of the annual flood peak discharges of two confluence river pairs using bivariate copulas. According to the proposed hypothesis, an accurate flood peak discharge estimation approach may be developed based on the joint probability distribution of coincident flows of the main river and tributary. Two case studies of confluence into the Sava River are analysed: confluence of the Kupa River and confluence of the Una River. The main aim of this research is to provide a practical approach for the accurate design flood estimation that is applicable to hydraulic problems of the joint exceedance probability of flood peaks at river confluences. For this purpose, discharge recordings from gauging stations situated to the upstream of the confluence are used rather than those from downstream stations. The assumption is that more information on a flood event can be taken into account if each river is represented in the analysis with its own regime, and then their joint influence is quantified through information about joint return period of an exceedance of a data pair. The following methodology was used for evaluation of the proposed approach:

- identify distribution of annual flood peak discharges of the main river and the tributaries
- identify joint probability distribution for the annual extreme discharge pair on the river confluence using the copula function to account for the most severe flood event
- compare joint exceedance probability with traditional univariate flood frequency analysis based on the discharge observation data collected downstream of the confluence
- make comparison with the most recent measurement of the significant flood peak discharge recorded during the event characterized as a 1000-year return period by the Croatian Meteorological and Hydrological Service.

## 2. Theoretical background

Models for multivariate flood frequency analysis require more data than the univariate ones, which is a shortcoming when

analysis of hydrological data is conducted, because the dataset is mostly restricted with the data collection interval and length of operational period, while also being impacted by water regime changes induced either naturally or artificially. Limitation in modelling multivariate distributions is that it requires that all marginal distributions come from the same family of distributions, which is extremely challenging to fulfil because of the different nature of the dependent variables. Taking into account these constraints, the application of multivariate analysis of correlated random variables is in practice mainly reduced to the bivariate case [6]. Copulas can be used for the multivariate hydrologic modelling, including calculation of joint probability distributions for correlated random variables. Copulas are functions that combine several univariate marginal cumulative distribution functions into their joint cumulative distribution function. The biggest advantage of copula approach is its flexibility, i.e. capability of determining the joint multivariate distribution function independent of marginal distributions [2, 3]. More theoretical details can be found in papers proposed by Sklar (1959) and Nelsen (2006). According to Sklar's theorem, any  $n$ -dimensional distribution function  $F$  can be formulated through a copula and its marginal distributions, which can be expressed as:

$$F(x_1, x_2, \dots, x_n) = C(F_{X_1}(x_1), F_{X_2}(x_2), \dots, F_{X_n}(x_n)) \quad (1)$$

where  $F_{X_i}(x_i)$ ,  $i = 1, \dots, n$  are marginal distributions of a random vector  $(X_1, X_2, \dots, X_n)$ .

If these marginal distributions are continuous, then a single copula function  $C$  exists, which can be written as:

$$C(u_1, u_2, \dots, u_n) = F(F_{X_1}^{-1}(u_1), F_{X_2}^{-1}(u_2), \dots, F_{X_n}^{-1}(u_n)) \quad (2)$$

The conditional distribution can be obtained after an appropriate copula function is selected. The conditional distribution function of  $U_1$ , given  $U_2 = u_2$  can be expressed as:

$$C_{(U_1|U_2=u_2)}(u_1) = C(U_1 \leq u_1 | U_2 = u_2) = \lim_{\Delta u \rightarrow 0} \frac{C(u_1, u_2 + \Delta u) - C(u_1, u_2)}{\Delta u} = \left( \frac{\partial}{\partial u_2} C(u_1, u_2) \right) |_{U_2 = u_2} \quad (3)$$

Furthermore, the conditional distribution function of  $U_1$ , given  $U_2 < u_2$  can be expressed as:

$$C_{(U_1|U_2 \leq u_2)}(u_1) = C(U_1 \leq u_1 | U_2 \leq u_2) = \frac{C(u_1, u_2)}{u_2} \quad (4)$$

Furthermore, the probability density function of a copula function can be expressed as:

$$c_{U_1, U_2} = \frac{\partial^2 C(u_1, u_2)}{\partial u_1 \partial u_2} \quad (5)$$

A copula  $C$  is called an extreme value copula if there exists a copula  $C_F$  such that:

$$C_F(u_1^{1/n}, \dots, u_d^{1/n})^n \rightarrow C(u_1, \dots, u_d) \quad n \rightarrow \infty \quad (6)$$

for all  $(u_1, \dots, u_d) \in [0, 1]^d$ .

When the problem consists in modelling extreme values, the extreme value theory tells us that max-stable distributions should be considered and that some restrictions should be put on the copulas to be used [11]. Finally, the joint distribution function can be calculated after selection of an appropriate copula function that fulfils the requirements [12, 13].

Since different marginal and joint probability distribution functions can be chosen in copula functions, the selection of marginal distributions has the greatest impact on the performance of the copula. Flood variables commonly selected for multivariate hydrologic frequency analysis through copula functions are annual maximum peak discharges and the associated hydrograph volumes and durations. The copula function selection has been in the focus of numerous studies and different families of copulas have been proposed for hydraulic models, as described in relevant literature. Volpi and Fiori [2] used Gumbel-Hougaard copulas to model dependency between peak and volume in flood analysis; Gräler *et al.* [1] used synthetic dataset to show difference in performance of 2D and 3D copulas for various combinations of hydrologic variables; Xu *et al.* [5] compared four Archimedean family copulas for derivation of a design flood hydrograph; Bender *et al.* [14] used Gumbel copula for bivariate analysis of concurrent flows at a river confluence, and Bender *et al.* [15] continued to investigate bivariate analysis of discharges on confluences, investigating sites where floods do not occur simultaneously. Szolgay *et al.* [16] evaluated applicability of different copula types to a peak-volume flood relationship over a large region in Austria and its sub-catchments, and also studied the influence of the data series length on bivariate copula performance. Sraj *et al.* [17] compared three copula families for bivariate flood frequency analysis in the Sava River basin in Slovenia, and the results obtained by Ozga-Zielinski *et al.* [18] showed that the Gumbel-Hougaard copula yields better results compared to Gaussian copula for snowmelt floods.

The Archimedean copula family ranks among the most desirable ones for hydrologic analyses because most of its copulas can easily be constructed, and they can also describe the positive and negative correlation between hydrological variables [17]. Based on the literature, Gumbel-Hougaard, Ali-Mikhail-Haq, Frank and Cook-Johnson copulas are the most frequently used one-parameter Archimedean copulas. For the Gumbel-Hougaard copulas, the relationship between the Kendall's coefficient and the generating function shows that only the positive dependence structure of the bivariate data can be analysed, i.e. the flood peak and volume, and the flood duration and volume are positively correlated, and their bivariate distributions can be derived using this copula [5, 19]. On the other hand, the flood duration and flood peak are usually negatively correlated and therefore the Gumbel-

Hougaard copula is not suitable for deriving their joint distribution. The same is true for the Cook-Johnson copulas. The Ali-Mikhail-Haq copula can be applied to both negatively and positively dependent random variables, but this copula is not appropriate if the dependence structure represented by the Kendall's coefficient approaches a very high value of positive correlation, or a very low value of negative correlation. The Frank copula can be applied to deriving joint distributions of both negatively and positively dependent random variables, and there is no restriction with regard to the degree of correlation. Gumbel-Hougaard copula does not allow negative dependence and exhibits strong right-tail dependence and relatively weak left-tail dependence. If outcomes are known to be strongly correlated at high values but are less correlated at low values, then the Gumbel-Hougaard copula is an appropriate choice [3]. Therefore, the Gumbel-Hougaard copula was used to model the dependence between the flood peak discharges on the Sava River and its tributaries in this study. The copula parameter is estimated based on the Kendall's tau coefficient approach.

The identification of marginal distributions of the correlated variables is the initial task in constructing the bivariate copula model. The distribution of population of maximum annual discharges can be compared with a parametric distribution using the so called Chi-square test [5] or the Anderson-Darling test [14]. After estimation of marginal distributions, the joint distribution function can be calculated using one of the copula functions evaluated by, for instance, goodness of fit test based on the Cramér-von-Mises statistics [15]. Using the fitted marginal distribution functions, the discharge data are transformed into uniformly distributed variables for which the joint probability distribution function is calculated using the selected copula model and its respective parameter(s). Once the joint probability distribution function is known, the data can be transformed from uniform variables to the original variables domain using the inverse cumulative distribution function [2]. A detailed description of marginal distribution selection, as well as the copula fitting evaluation using log-likelihood, is given by Gräler et al. [1].

### 3. Study reach

The Sava River is the right-hand tributary of the Danube, and also its largest tributary when considering its total length (990 km) and water volume (average discharge at the confluence is 1564 m<sup>3</sup>/s). It is the second largest Danube tributary as per catchment size (95419 km<sup>2</sup>). The Sava River is formed in Radovljica, Slovenia, by confluence of the rivers Sava Dolinka and Sava Bohinjka. It then flows through Croatia, forms the border with Bosnia and Herzegovina, and finally discharges into the Danube in Serbia. The Sava River exhibits the peri-Pannonian pluvial-nival regime, with high water levels occurring in the autumn and winter period, from August till February.

Two tributaries selected for analyses, the Kupa River and the Una River, are both right-hand side tributaries of the Sava River: the Kupa River with the confluence at rkm 576+000 and the Una River with the confluence at rkm 500+000. The catchment area of the Kupa River is 10605 km<sup>2</sup> overall until its discharge into the Sava River near Sisak, where the Sava River catchment size is 22852 km<sup>2</sup>. The Kupa River regime is showing spring and autumn maxima, which is typical for rivers with the pluvial-nival regime [20]. The Una River has a total length of 214 km and a total catchment of 10816 km<sup>2</sup>. It discharges into the Sava River near Jasenovac, where the Sava River catchment size is 38953 km<sup>2</sup>. The Una exhibits the Posavina variant of the pluvio-nival regime, with the highest water levels in April, and then in March, May, and November [21].

The analysis of flood hazard is based on the continuous record of water levels or discharges on a gauging station that reflects water regime on a particular river section. Discharges, which are the primary generators of flood hazard, can be calculated from water levels recorded using a calibrated discharge rating curve. The analysed Kupa River and Una River confluences are well covered with gauging station records. The Sava River near the Kupa River confluence is monitored by the GS Strelečko upstream of the confluence, and the GS Crnac downstream, while the Kupa River is monitored by the GS Farkašić. All three gauging stations have been recording data from 1955 onwards,



Figure 1. Layout of Sava River basin [22]

Table 1. Characteristic discharges for analysed gauging stations

River	Gauging station	Distance from confluence [rkm]	Characteristic discharges	
			$Q_{SR}$ [m <sup>3</sup> /s]	$Q_{MAX}$ [m <sup>3</sup> /s]
Sava	Strelečko	581+000	427	1544
Kupa	Farkašić	47+150	195	1585
Sava	Crnac	575+000	622	2359
Sava	Jasenovac	500+500	781	2671
Una	Dubica	20+100	227	1808
Sava	Gradiška	453+400	792	3493

with the gap in the period from 1991 to 2000. The discharge is measured at the GS Farkašić and GS Crnac, and so the discharge time series for the GS Strelečko are calculated by subtracting the daily discharge at the GS Farkašić from the one recorded at the GS Crnac. The Sava River near the Una River confluence is monitored by the GS Jasenovac upstream of the confluence, and the GS Gradiška downstream, while the Una River is monitored by the GS Dubica. All three gauging stations have been recording data from 1926 onwards, with the gap in the period from 1991 to 2005. Both tributaries, Una and Kupa, exhibit similar water regimes, with the mean discharge of  $Q_{AVG,KUPA} = 195 \text{ m}^3/\text{s}$  and  $Q_{AVG,UNA} = 227 \text{ m}^3/\text{s}$ ; and maximum discharge of  $Q_{MAX,KUPA} = 1585 \text{ m}^3/\text{s}$  and  $Q_{MAX,UNA} = 1808 \text{ m}^3/\text{s}$ . Records have been kept longer at the Una River confluence gauging station (1926-2015) compared to the records at Kupa River confluence (1955-2015), which might affect the calculated long term regime properties. The discharge is measured at all three stations: GS Jasenovac and GS Gradiška at the Sava River, and GS Dubica at the Una River. Characteristic discharges recorded or calculated for all gauging stations are given in Table 1.

Over the last ten years, an increased number of flood events has been recorded in the Sava River basin: in 2010, the highest Sava water level was recorded in Zagreb - second only to flood event levels registered in 1964; the discharge equalling a 100-year return period was registered in 2014 at the Kupa River in Karlovac area. Finally, the May 2014 flood in the lower Sava River reaches was marked as a 1000-year return period flood. The frequency and intensity of floods registered at the Sava and its tributaries has recently been hard to ignore, especially in May 2014 when the levee collapsed after days of high water flow, resulting in disastrous damage and evacuation of residents from several villages. After a statistical analysis, it was established that the measured flood peak discharge of this event ( $Q = 6000 \text{ m}^3/\text{s}$ , measured on May 17, 2014) downstream from the Bosna River discharge into the Sava River equalled a 1000-years return period [23]. Further upstream, on the river section from Jasenovac to Slavonski Brod, the return period was a bit lower and corresponded to 100-year flood.

It has been shown that the Bosna River (tributary to the Sava) exerted the biggest influence on the 2014 flood in Croatia. Even though the discharge into the Sava River was not extreme, the coincidence of two flood events resulted in maximum discharge

that surpassed all previous records, setting new highest measurement downstream of the Bosna River confluence. Abdulaj *et al.* compared theoretical univariate probability fitted on discharge data from almost 100 years of gauging records and concluded that - even if the most favourable distribution is selected - the recorded discharge would correspond to a 1000-year flood [23]. Because of its erratic nature, this high flow event has prompted a large number of new analyses, including re-examination of traditional approaches to flood event estimates for the Sava River.

#### 4. Results and discussion

Given the flow regime of the Sava River and its tributaries, it can be expected that high flows at the Sava River coincide with high flows in one or more of its tributaries, Figure 2, resulting in flood wave whose severity exceeds the design levee crest elevation, causing in turn flooding of populated areas defended by levees. The coincidental flood occurrence analysed in the present study was conducted on the annual flood peak discharge data pairs, given in following figure (Figure 2) as discharge time series showing annual extremes for the entire monitoring period.

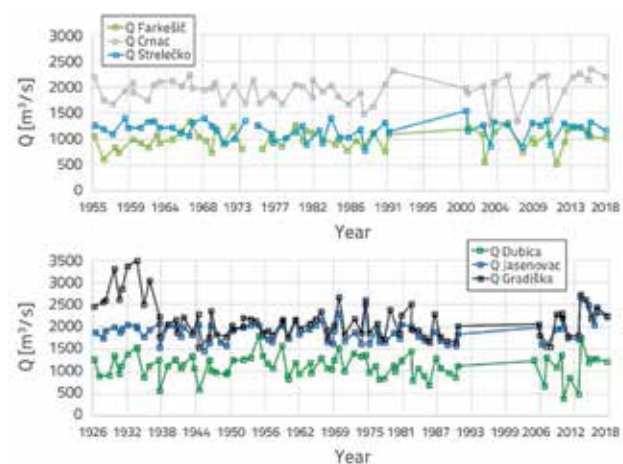


Figure 2. Annual extremes of flood peak discharges: Kupa confluence (up), Una confluence (down)

Pairs of annual extreme discharges into the Sava River and its tributaries show that 22% and 44% of the floods occurred

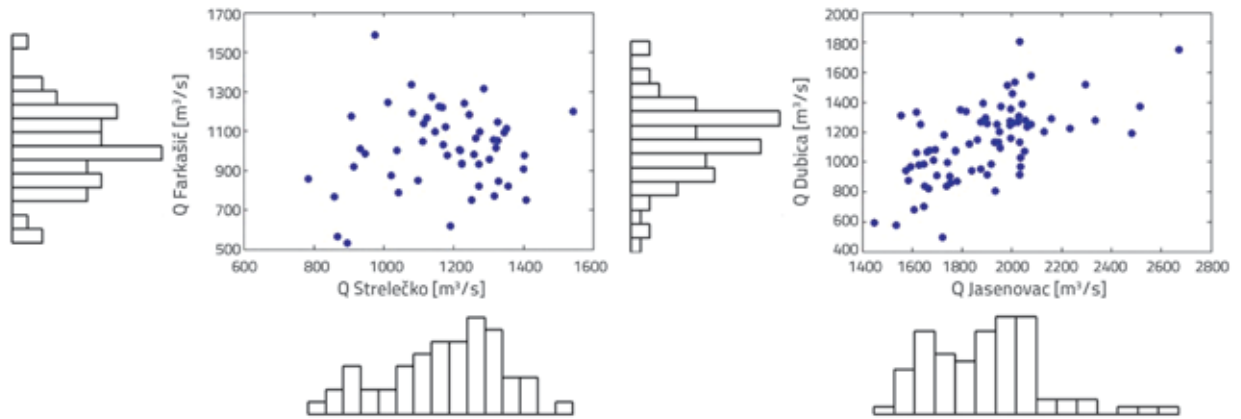


Figure 3. Flood events for annual peak discharge pairs: a) Kupa confluence, b) Una confluence

simultaneously in the Kupa and Una rivers, respectively, taking into account the criterion that simultaneous floods occur when the time difference between the peak at the main river and the tributary is 7 days or less. The fact that the time that elapsed between the flood peaks is relatively low indicates that pairs of annual extremes can be used for the bivariate flood frequency analysis.

Discharge time series of annual flood peak discharges reveal that high annual peaks at the Kupa River do not frequently coincide with high annual peaks at the Sava River (Figure 2 and Figure 3). For the period prior to 1990, the high peak at the Sava River was more frequently paired with the low peak at the Kupa River, and vice versa. In the period after 1990, annual peaks coincide almost every year. This relation becomes more evident when flood events for annual peak discharge pairs at the GS Strelečko and the GS Farkašić are plotted against each other in the scatter plot (Figure 3). This scatter plot shows that any given annual peak discharge at the GS Strelečko can coincide with discharge covering the entire span of annual peaks at the GS Farkašić, and vice versa. At the Kupa River confluence with the Sava River, it can be seen that both rivers have similar discharges, with the Sava having larger annual peaks. However, the extreme discharge for the entire recording period is similar.

The trend of occurrence of annual peaks at the Una River confluence is significantly different: both high and low annual peaks at the Una River frequently coincide with high and low annual peaks at the Sava River, respectively (Figure 2 and Figure 3). The uppermost data point in Figure 3 is the flood event from the year 2013, when the maximum discharge at the GS Jasenovac coincided with 2<sup>nd</sup> highest annual peak discharge at the GS Dubica. Although several similar events occurred between 1970 and 1982, this event was the one with highest combined discharge since 1936. During that flow event, water level at the GS Crnac reached the highest level since the start of the recordings, which prompted installation of additional flood protection measures to ensure that flooding does not reoccur. Similarly, a flood event from 2014 when the levee collapsed was 3<sup>rd</sup> highest annual peak discharge at the GS Jasenovac. In that year, high peaks of the Sava and Una rivers did not coincide, and thus even more disastrous consequences of flooding were prevented. Considering frequent coincidence of maximum annual discharges necessitating temporary levee heightening measures, doubts have been raised regarding suitability of using traditional approach for determining flood hazard at river confluences.

Table 2. Summary of fitted univariate theoretical distributions and resulting characteristic discharges for given return period

Gaugin station	Fitted distribution	p-value	Q <sup>PR</sup> [m <sup>3</sup> /s]					Q <sub>MAX</sub> [m <sup>3</sup> /s]	
			Q <sup>50</sup>	Q <sup>100</sup>	Q <sup>200</sup>	Q <sup>500</sup>	Q <sup>1000</sup>	Measured	Return period [year]
Strelečko	Gamma	0,58	1547	1603	1655	1720	1766	1544	48
Farkašić	Gamma	0,18	1475	1548	1617	1703	1765	1585	145
Crnac	Gamma	0,11	2507	2591	2669	2765	2834	2359	17
Jasenovac	Log-Normal	0,18	2405	2485	2561	2655	2724	2671	585
Dubica	Gamma	0,10	1767	1871	1970	2094	2184	1808	65
Gradiška	Log-Normal	0,56	3093	3250	3402	3595	3737	3493	306

#### 4.1. Univariate distributions of flood peaks using traditional annual extremes method

For all gauging stations, theoretical distributions were fitted to annual flood peak discharge data to obtain characteristic discharges using traditional approach for comparison with the proposed copula method. For each dataset, the following distributions were fitted to data: Log-Normal, Gamma, Pearson, and Weibull. In order to evaluate the goodness of fit for four parametric distributions when quantifying probability distributions of flood variables, the Kolmogorov–Smirnov (K–S) test was performed to test the assumption that the flood magnitudes follow the theoretical distribution. The criterion for goodness of fit was  $p$ -value, i.e. the threshold  $p$ -value was used as a criterion for rejection of the proposed distribution hypothesis. A  $p$ -value greater than the threshold ( $p \geq 0.05$ ) is interpreted as a good fit by a fitted theoretical distribution. The K-S test shows that the assumption could not be rejected at the 5% significance level for all fitted distributions, and the one with the highest  $p$ -value was selected as the best fit to the data. The results of K-S test are given in Table 2 for the selected distribution only, together with the resulting characteristic discharges for a span of return periods. Resulting discharges for the return periods were determined as simple reciprocal of the probability of exceedance.

The distribution of annual peaks follows Gamma or Log-Normal distribution at all stations, but with different parameters. Annual peaks at the Kupa and Sava rivers to the upstream of the Kupa confluence follow similar theoretical distributions, i.e. the best fit for all of them is Gamma distribution, which is quite expected given high similarity of their flow regimes. At the GS Crnac, several low annual peaks ( $<1900 \text{ m}^3/\text{s}$ ) shift the distribution to the left, and measured discharges are grouped closely together, making their distribution very narrow. This results in a relatively low estimate of discharges for low probability of exceedance. When the maximum measured discharge  $Q_{MAX}$  is evaluated from the fitted theoretical distribution, its return period is 17 years only (Table 2). The Sava River to the upstream of the confluence, at GS Strelečko, reflects a similar pattern, where the return period of  $Q_{MAX}$  is 48 years. On the other hand, the  $Q_{MAX}$  measured for the Kupa River has the return period of 145 years. For the Una River confluence, annual peaks at the Sava River follow Log-Normal distributions, while tributary data follow Gamma distribution. Discharges measured at the tributary are grouped closely together, with few low annual peaks. Discharges measured at the Sava River follow theoretical distribution very closely, and discharges with very high annual peaks are scarce towards the distribution tail. When maximum measured discharge  $Q_{MAX}$  is evaluated from fitted theoretical distribution, the Sava River (upstream) and the Una River show different return periods of 585 years and 65 years, respectively (Table 2), which reflects the difference in the annual flood peak distribution. For the Sava River (downstream) at the GS Gradiška, the return period

for  $Q_{MAX}$  is closer to the one at the GS Jasenovac and amounts to 306 years. This confirms conclusions from the analysis of coincidence for annual flood peaks, which reveal that they coincide frequently, shifting measured data histogram towards the distribution tail.

#### 4.2. Joint bivariate distribution of flood peaks using Gumbel-Hougaard copula

The joint probability distribution of two flood peaks at the confluences was evaluated by copulas to calculate the joint exceedance probability for the events resulting from the inflow of the Sava River and tributaries in terms of probabilities of coincidence of a particular flood event at both rivers. Therefore, the corresponding annual flood peaks had to be selected for analysis from available recordings. As preliminary analyses have shown, annual flood peaks at the Kupa River confluence do not coincide frequently, while they do for the Una River confluence (Figure 3). This analysis was conducted on pairs of annual peaks, without determining if they happened simultaneously in the same day or few days apart as a part of a single flood event. The statistical dependencies between the Sava River and tributary discharge annual peaks were evaluated through the non-parametric dependence measure, Kendall's tau. Kendall's tau is not dependent on distributional assumptions and is resistant to outliers. Results show that there is a strong positive dependence between the discharges at both confluences, with values of Kendall's tau at 0.07 for the Kupa River (Strelečko-Farkašič) dependency and 0.40 for the Una River (Jasenovac-Dubica) dependency.

The joint return period of the corresponding flood peaks at two confluences between the Sava River and its tributaries is calculated by copula function. The Gumbel-Hougaard copula was selected for modelling dependence between peak flow discharges based on literature review that shows that it is the only Archimedean copula that is max-stable, meaning that it is an extreme value copula [11, 12, 24]. The isolines of the joint distribution function for flood peak discharges are shown in Figure 4 and Figure 5 for confluences of the Kupa River and the Una River, respectively. Joint return periods calculated using copulas can be used to determine severity of flood for desired relations between the Sava River and its tributary for the exceedance of both variables.

The following figure (Figure 4) shows the scatter plot of measured data pairs and simulated values generated from the copula model (sample size: 10 000) plotted with isolines from the joint distribution function for the discharge data pairs at the Kupa confluence. The following trend can be observed: the value of the conditional distribution function for one flood peak decreases with an increase of the other flood peak value. This correlation between flood peaks indicates that the probabilities of concurrence of large flood at both the Sava River and its tributary are generally lower than the exceedance probability of high peak flows at each of the rivers.

Table 3. Comparison of characteristic discharges based on univariate and copula methods

$Q^{PR}$ [m <sup>3</sup> /s]	Kupa river confluence GS Crnac			Una river confluence GS Gradiška		
	Univariate	Copula	Difference [%]	Univariate	Copula	Difference [%]
$Q^{50}$	2507	3338	+33	3093	4627	+50
$Q^{100}$	2591	3385	+30	3250	4717	+45
$Q^{1000}$	2834	3502	+24	3737	4878	+31

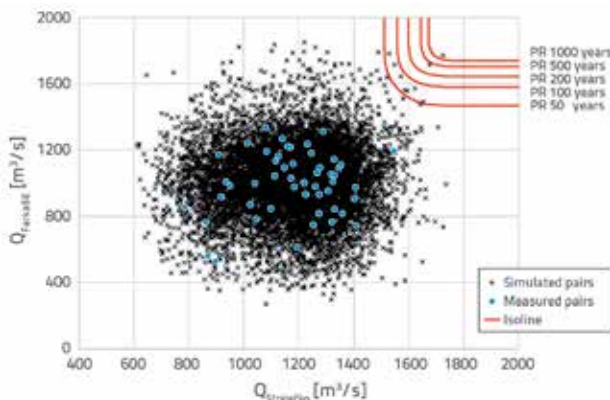


Figure 4. Scatter plot of measured data pairs and simulated values generated from the copula model for Kupa confluence

Figure 5 shows scatter plot of measured data pairs and simulated values generated from the copula model plotted with isolines from joint distribution function for discharge data pairs at the Una confluence. These results differ significantly from the ones presented for the Kupa River confluence, which is due to significant difference in marginal distributions of the peak flows at the Sava and Una rivers, whereas the Sava and Kupa rivers had same type of marginal distributions and a similar flow regime.

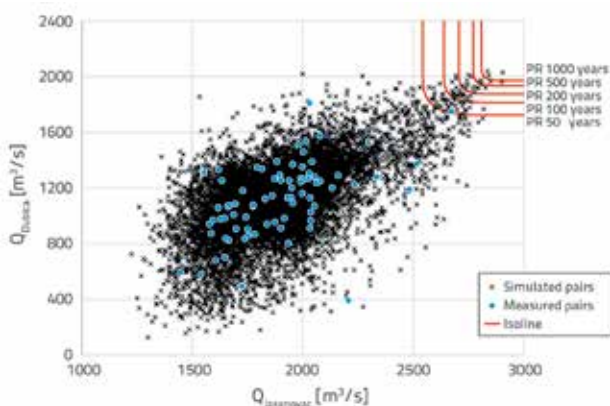


Figure 5. Scatter plot of measured data pairs and simulated values generated from copula model for the Una confluence

The comparison of flood peak discharges calculated using traditional univariate distribution and bivariate copulas at annual flood peak discharges is given in the following table

(Table 3). For traditional method, the resulting discharge for the selected return period is calculated as reciprocal of the probability of exceedance. On the other hand, the discharge calculated using the copula method can lie anywhere on the isoline representing return period and can have infinite combinations. Since the worst-case scenario is representative for the hydraulic problem 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 Sava River and its respective tributary. Discharges estimated from copulas are significantly higher than the ones calculated using the univariate method. Considering that the coincidence of flood peaks is not common at analysed confluences, and that it is therefore rarely reflected in gauging station recordings downstream of the confluence, this outcome is quite expected. For the Kupa River confluence, an increase in discharge calculated by copulas in comparison to traditional approach, ranges from 33% for a 50-year return period to 24% for a 1000-year return period. The configuration of the joint distribution function for the Kupa River confluence resembles closely to an inclined plane, thus giving a uniform increase in discharges compared to univariate distribution. For the Una River confluence, an increase in calculated discharge ranges from 50% for a 50-year return period, to 31% for a 1000-year return period. The curvature of the joint distribution function for the Una River confluence is higher than the one for the Kupa River, and gives therefore a non-uniform difference in comparison to univariate distribution. When the discharge measured during the extreme flood in 2014 is compared to fitted theoretical distributions [23], the measured discharge increases between 25 % and 66 %. Prohaska and Ilić have analysed the aforementioned event from 2014 on the downstream Sava River reaches at the Drina River confluence using multiple probability distribution functions. They have concluded that the probability of coincidence for measured discharges is a 2000-year flood [25]. If we adopt the common conclusion that this was a 1000-year flood, it can be stipulated that this event could be anticipated if the proposed copula method is used for flood hazard assessment at river confluences.

### 5. Conclusion

The research conducted in this paper shows that flood hazard at the Sava River could be underestimated by traditional univariate flood frequency analysis. A recent catastrophic event that



occurred in 2014 at the Sava River highlighted the extreme flow conditions that can occur when flood peaks at the River Sava and its tributary coincide, even when discharge to the Sava River is not extreme at the time. In this paper, advanced possibilities of copula model for modelling hydraulic variables are utilized for bivariate frequency analysis. Gumbel-Hougaard copula is used for bivariate flood frequency analysis of the flood peak discharge data pairs at the confluences of the Sava River and its tributaries to calculate the joint return period of the coincidence floods. The results indicate that discharges estimated using the copula method are significantly higher than the ones calculated using the univariate method. This outcome is expected considering that the coincidence of flood peaks is not common on analysed

confluences and that, therefore, it is rarely reflected in the gauging station recordings downstream of the confluence. Results show that proposed copula approach estimates recent flood events more accurately than the univariate flood frequency analysis based on the observation data. It can be stipulated that the 2014 flood event could have been anticipated if the copula method had been used for flood hazard assessment. The 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 confluences. At river confluences, marginal distributions of inflow discharges rarely follow similar distribution, making copula model especially suitable for flood hazard assessment.

## REFERENCES

- [1] Gräler, B., van den Berg, M. J., Vandenbergh, S., et al.: Multivariate return periods in hydrology: a critical and practical review focusing on synthetic design hydrograph estimation, *Hydrol. Earth Syst. Sci.*, 17 (2013), pp. 1281-1296, <https://doi.org/10.5194/hess-17-1281-2013>
- [2] Volpi, E., Fiori, A.: Design event selection in bivariate hydrological frequency analysis, *Hydrological Sciences Journal - Journal des Sciences Hydrologiques*, 57 (2012) 8, pp. 1506-1515.
- [3] Wang, C.: A joint probability approach for the confluence flood frequency analysis, Iowa State University, 2007.
- [4] Yue, S., Rasmussen P.: Bivariate frequency analysis: discussion of some useful concepts in hydrological application, *Hydrological Processes*, 16 (2002), pp. 2881-2898, <https://doi.org/10.1002/hyp.1185>
- [5] Xu, C., Yin, J., Guo, S., Liu, Z., Hong, X.: Deriving Design Flood Hydrograph Based on Conditional Distribution: A Case Study of Danjiangkou Reservoir in Hanjiang Basin, *Mathematical Problems in Engineering*, (2016), Article ID 4319646, 16 pages.
- [6] Klein, B., Pahlow M., Hundecha, Y., Schumann, A.: Probability Analysis of Hydrological Loads for the Design of Flood Control Systems Using Copulas, *Journal of Hydrologic Engineering*, 15 (2010) 5, pp. 360-369.
- [7] Papaioannou, G., Kohnová, S., Bacigál, T., Szolgay, J., Hlavčová, K., Loukas, A.: Joint modelling of flood peaks and volumes: A copula application for the Danube River, *J. Hydrol. Hydromech.*, 64 (2016) 4, pp. 382-392.
- [8] Requena, A. I., Mediero, L., Garrote, L.: A bivariate return period based on copulas for hydrologic dam design: accounting for reservoir routing in risk estimation, *Hydrol. Earth Syst. Sci.*, 17 (2013), pp. 3023-3038, <https://doi.org/10.5194/hess-17-3023-2013>
- [9] Ilić, A., Prohaska, S.: Multiple Coincidence of Flood Waves in Complex River Systems, *Water & Environmental Dynamics - 6th International Conference on Water Resources and Environment Research*, Koblenz, Germany, pp. 94-113, 2013.
- [10] Dung, N. V., Merz, B., Bárdossy, A., Apel H.: Handling uncertainty in bivariate quantile estimation - An application to flood hazard analysis in the Mekong Delta, *Journal of Hydrology*, 527 (2015), pp. 704-717, <https://doi.org/10.1016/j.jhydrol.2015.05.033>
- [11] Ribatet, M., Sedki, M.: Extreme value copulas and max-stable processes, *Journal of Société Française de Statistique (special edition on copulas)*, 153 (2012) 3, pp. 138-150.
- [12] Fan, Y. R., Huang, W. W., Huang, G. H., Li Y. P., Huang, K., Li Z.: Hydrologic risk analysis in the Yangtze River basin through coupling Gaussian mixtures into copulas, *Advances in Water Resources*, 88 (2016), pp. 170-185, <https://doi.org/10.1016/j.advwatres.2015.12.017>
- [13] Volpi, E., Fiori, A.: Hydraulic structures subject to bivariate hydrological loads: Return period, design, and risk assessment, *Water Resources Research*, 50 (2014), pp. 885-897, <https://doi.org/10.1002/2013WR014214>
- [14] Bender, J., Wahl, T., Mudersbach, C., Jensen, J.: Flood Frequency Analysis at River Confluences - Univariate vs. Multivariate Extreme Value Statistics, *Water & Environmental Dynamics - 6th International Conference on Water Resources and Environment Research*, Koblenz, Germany, pp. 316-328, 2013.
- [15] Bender, J., Wahl, T., Müller, A., Jensen, J.: A multivariate design framework for river confluences, *Hydrological Sciences Journal - Journal des Sciences Hydrologiques*, 61 (2016) 3, pp. 471-482.
- [16] Szolgay, J., Gaál, L., Bacigál, T., et al.: A regional comparative analysis of empirical and theoretical flood peak-volume relationships, *Journal of hydrology and hydromechanics*, 64 (2016) 4, pp. 367-381.
- [17] Sraj, M., Bezak, N., Brilly, M.: Bivariate flood frequency analysis using the copula function: a case study of the Litija station on the Sava River, *Hydrol. Process.*, 29, pp. 225-238, 2015, <https://doi.org/10.1002/hyp.10145>
- [18] Ozga-Zielinski, B., Ciupak, M., Adamowski, J., Khalil, B., Malard, J.: Snow-melt flood frequency analysis by means of copula based 2D probability distributions for the Narew River in Poland, *Journal of Hydrology: Regional Studies*, 6 (2016), pp. 26-51, <https://doi.org/10.1016/j.ejrh.2016.02.001>
- [19] Li, T., Guo, S., Liu, Z., Xiong, L., Yin, J.: Bivariate design flood quantile selection using copulas, *Hydrology Research*, (2016), Article ID nh2016049, 17 pages.
- [20] Čanjevac, I.: Tipologija protočnih režima rijeka u Hrvatskoj, *Hrvatski geografski glasnik*, 75 (2013) 1, pp. 23-42.
- [21] Korjenić, A.: Basic characteristics water regime and water balance of the river Una, *Acta geographica Bosniae et Herzegovinae*, 2 (2014), pp. 65-75.
- [22] ISRBC: Sava River Basin Analysis Summary, International Sava River Basin Commission, Zagreb, Croatia, 2010.
- [23] Abdulaj, R., Miković, N., Oskoruš, D., Vujnović, T.: Velike vode donjeg toka rijeke Save tijekom svibnja 2014., *Hrvatska vodoprivreda*, 207 (2014), pp. 14-17.
- [24] Klein, B., Pahlow, M., Gattke, C., Hundecha, Y., Schumann, A.: Probabilistic analysis of hydrological loads to optimize the design of flood control systems, *4<sup>th</sup> International Symposium on Flood Defence: Managing Flood Risk, Reliability and Vulnerability*, Toronto, Canada, 2008.
- [25] Prohaska, S., Ilić, A.: Koincidencija velikih voda Save i Drine, *Hrvatske vode*, 24 (2016) 95, pp. 1-18.