

UDK/UDC: 519.21:532.433:556

Prejeto/Received: 04. 02. 2014

Izvirni znanstveni članek – Original scientific paper

Sprejeto/Accepted: 14. 05. 2014

PRIMERJAVA MED KLASIČNIMI UNIVARIATNIMI VERJETNOSTNIMI ANALIZAMI IN BIVARIATNIMI Z UPORABO FUNKCIJE KOPULA COMPARISON BETWEEN CLASSICAL UNIVARIATE FREQUENCY ANALYSIS AND BIVARIATE ANALYSIS WITH COPULA

Mojca Šraj^{1,*}, Nejc Bezak¹, Mitja Brilly¹

¹ Fakulteta za gradbeništvo in geodezijo, Univerza v Ljubljani, Hajdrihova 28, 1000 Ljubljana

Izvelek

Verjetnostne analize so osnova za določanje projektnih pretokov. Običajno se pri analizah upošteva le ena spremenljivka, večinoma letna konica pretoka. Ker pa so hidrološki pojavi določeni z več medsebojno odvisnimi spremenljivkami, je pri analizah smiselno multivariatno pristop. Primer takega postopka je funkcija kopula. Klasične, univariatne verjetnostne analize so še vedno predpogoj za izvedbo analiz z uporabo funkcije kopula. Bivariatno verjetnostno analizo z uporabo funkcije kopula smo naredili za letne konice pretokov in pripadajoče volumne vodomerne postaje Litija na reki Savi. Uporabili smo tri funkcije kopula iz Arhimedove družine in parametre ocenili s pomočjo Kendallovega koeficienta korelacije (metoda momentov). Izračunali smo skupne povratne dobe in jih primerjali z univariatnimi povratnimi dobami. Ugotovili smo, da razlike med povratnimi dobami niso zanemarljivo majhne. Pri dogodku iz leta 1990, ki je bil največji v opazovanem obdobju, povratna doba T^{AND} znaša 92 let, T^{OR} pa 17 let. Univariatni povratni dobi konic in volumnov pa ležita v območju med omenjenima povratnima dobama. S pomočjo statističnih in grafičnih kriterijev ustreznosti smo ugotovili, da kopula Gumbel-Hougaard za obravnavani primer izkazuje boljše rezultate kot kopula Clayton ali kopula Frank.

Ključne besede: funkcija kopula, bivariatne analize, Kendallov koeficient korelacije, Arhimedova družina kopol, skupna povratna doba OR, skupna povratna doba AND.

Abstract

Frequency analyses are a basis for designing discharge estimations. Univariate flood frequency analyses are usually applied in hydrological practice. Hydrological processes are multivariate, however multivariate analyses are needed. Copula function can be used for multivariate modelling. Classical univariate flood frequency analyses are a precondition for the copula analyses. Flood frequency analyses were made on the annual maximum series data from gauging station Litija on the Sava River. Three copulas from the Archimedean family were used; parameters were estimated with method of moments (based on the Kendall correlation coefficient). Some joint return periods were calculated and compared with the univariate return periods. Differences between return periods were not negligible. In the case of a flood event in 1990, which was the largest in the observed period, T^{AND} was 92 years and T^{OR} was 17 years. Univariate return periods lay

* Stik / Correspondence: mojca.sraj@fgg.uni-lj.si

© Šraj M. et al.; Vsebina tega članka se sme uporabljati v skladu s pogoji [licence Creative Commons Priznanje avtorstva – Nekomercialno – Deljenje pod enakimi pogoji 4.0.](#)

© Šraj M. et al.; This is an open-access article distributed under the terms of the [Creative Commons Attribution – Non Commercial – Share Alike 4.0 Licence.](#)

between these two values. Statistical and graphical performance measures were used to choose the best fit copula function. Gumbel-Hougaard copula gave better results than Clayton and Frank copulas.

Keywords: copula function, bivariate analyses, Kendall's correlation coefficient, Archimedean copula family, joint return period OR, joint return period AND.

1. Uvod

Verjetnostne analize določajo razmerje med velikostjo in verjetnostjo nastanka opazovanega pojava. So osnova za določanje projektnih pretokov pri dimenzioniranju hidrotehničnih objektov, določanju poplavnih območij, pri upravljanju z vodami, obvladovanju tveganja zaradi poplav ipd. Ker gre večinoma za ukrepe, ki neposredno vplivajo na ogroženost določenega območja in prebivalstva (Đurović & Mikoš, 2004; Brilly & Polič, 2005; Kobold et al., 2005; Mikoš et al., 2004;), je potrebno verjetnostne analize izvesti s čim večjo zanesljivostjo in s hkratnim upoštevanjem več spremenljivk, ki določajo visokovodni val. V svetovni praksi se največkrat uporabljajo univariatne verjetnostne analize, kjer se večinoma upoštevajo le konice pretokov. To velja tudi za Slovenijo. Če želimo verjetnostne analize izvesti zanesljivo, je smiselno poleg konic pretokov statistično obravnavati tudi volumne in čase trajanja visokovodnih valov (Karmakar & Simonovic, 2008). Z upoštevanjem le ene spremenljivke lahko podcenimo dejansko povratno dobo opazovanega ekstremnega dogodka (Favre et al., 2004; Renard & Lang, 2007). Za izvedbo verjetnostne analize lahko uporabimo klasične bivariatne porazdelitvene funkcije, kot so: normalna, logaritemsko normalna, gama in generalizirana porazdelitev ekstremnih vrednosti (Genest & Favre, 2007). Pogoji za uporabo klasične bivariatne analize je, da sta obe spremenljivki definirani z isto porazdelitveno funkcijo (Genest & Favre, 2007). Kot alternativa klasičnemu bivariatnemu pristopu se v zadnjih letih v svetovni hidrološki praksi uveljavlja funkcija kopula, ki so jo pred tem že uspešno uporabljali zlasti v ekonomiji. V zadnjem desetletju so jo že uporabili tudi v nekaterih hidroloških študijah, predvsem pri analizah visokovodnih valov (Genest & Favre, 2007; Karmakar & Simonovic, 2008; Karmakar & Simonovic, 2009; Zhang & Singh, 2006; Zhang & Singh, 2007), ter tudi pri analizah suše (Ma et al.,

2011; Wong et al., 2010), analizah padavin (Balistocchi & Bacchi, 2011; Grimaldi & Serinaldi, 2006), preverjanju ustreznosti prelivnega objekta (De Michele et al., 2005), analizah tveganja (angl. *risk*) (Chen et al., 2012; Ganguli & Reddy, 2013). Glavna prednost funkcije kopula pred klasičnimi bivariatnimi analizami je, da lahko spremenljivke opišemo z različnimi robnimi porazdelitvenimi funkcijami (angl. *marginal distribution functions*) (Genest & Favre, 2007). Poleg tega sta izbira ustreznih robnih porazdelitev in uporaba funkcije kopula dva ločena koraka. Je pa res, da take kompleksne, multivariatne verjetnostne analize z uporabo funkcije kopula zahtevajo predhodno natančno analizo lastnosti in odvisnosti posameznih spremenljivk ter verjetnostne analize vsake posamezne spremenljivke (Bezak et al., 2013; Šraj et al., 2012), kar vsekakor pomeni mnogo več vloženega časa in dela. Zastavlja se vprašanje, kolikšna je razlika razmerij med velikostjo in verjetnostjo pojava pri uporabi univariatnih verjetnostnih analiz in bivariatnih analiz z uporabo funkcije kopula. V prispevku so prikazani rezultati bivariatnih analiz z uporabo funkcije kopula za konice in volumne, možno pa je tudi hkratno upoštevanje vseh treh spremenljivk (konic, volumni in časi trajanja), ki določajo visokovodne valove (trivariatne analize).

Cilji raziskave so bili: (i) narediti klasične, univariatne verjetnostne analize konic in volumnov visokovodnih valov, (ii) izvesti bivariatne verjetnostne analize konic in volumnov z uporabo funkcije kopula, (iii) primerjati univariatne povratne dobe z različnimi skupnimi povratnimi dobami (angl. *joint return periods*), določenimi z bivariatno kopulo.

2. Podatki

Za potrebe raziskave smo uporabili podatke z vodomerne postaje Litija I na reki Savi (slika 1). Hidrološka postaja Litija je najstarejša postaja

Agencije RS za okolje, kjer vodostaje opazujejo od leta 1893, podatki pa so na razpolago od leta 1895 (Ulaga, 2011). Leta 1953 je bila postaja zaradi zasipavanja dna prestavljena približno 500 m gorvodno in opremljena z limnigrafom, kar pomeni, da so od takrat dalje na voljo podatki o celotnih hidrogramih. Za analize smo torej uporabili 58 let podatkov (1953-2010) dnevni vrednosti pretokov z vključenimi maksimalnimi konicami, ki so bili izmerjeni z limnigrafom Seba Omega.



Slika 1: Hidrološka postaja Litija I na reki Savi (ARSO, 2013).

Figure 1: Hydrologic station Litija I on the Sava River (ARSO, 2013).

Za uporabo dnevni vrednosti pretokov z vključenimi absolutnimi konicami smo se odločili po pregledu podatkov in analizi ter primerjavi vseh treh spremenljivk pri uporabi urnih in dnevni vrednosti (Šraj & Bezak, 2013). Ker smo za analize potrebovali tudi volumne visokovodni valov, smo najprej izločili bazni odtok z uporabo grafične tri-točkovne metode. Nato smo določili vzorec letni maksimumov (Bezak, 2012), ki so ga sestavljale maksimalne vrednosti pretokov in pripadajočih volumnov posameznih let.

3. Metode

3.1 Univariatne verjetnostne analize

Vzorec smo določili po metodi letni maksimumov (Bezak et al., 2013). Za vsako spremenljivko

posebej (konice pretokov in pripadajoči volumni) smo najprej naredili univariatne verjetnostne analize (Šraj et al., 2012).

Uporabili smo različne porazdelitvene funkcije, ki se v svetovni hidrološki praksi najpogosteje uporabljajo (logaritemsko normalno porazdelitev, Pearsonovo III porazdelitev, logaritemsko Pearsonovo III porazdelitev, Gumbelovo porazdelitev, normalno porazdelitev, generalizirano porazdelitev ekstremnih vrednosti (GEV) in generalizirano logistično porazdelitev (GL)) in tri metode ocenjevanja parametrov (metodo momentov, metodo momentov L in metodo največjega verjetja) (Bezak, 2012; Turk, 2012; Šraj et al., 2012). Z uporabo statistični testov, kriterijev ujemanja in grafični prikazov ujemanja smo določili najprimernejšo porazdelitev za izvedbo verjetnostni analiz konic in volumnov (Šraj et al., 2012).

3.2 Funkcija kopula

Funkcija kopula predstavlja alternativo klasičnemu, univariatnemu pristopu k verjetnostni analizam. Postopek analize je sestavljen iz naslednjih korakov, ki jih lahko združimo v dve skupini:

- Oblikovanje vzorca letni maksimumov ali metode vrednosti nad izbranim pragom, izločanje baznega odtoka, izvedba univariatni verjetnostni analiz in izbira najustrežnejše porazdelitvene funkcije (Bezak, 2012; Šraj et al., 2012).
- Določitev robni porazdelitveni funkcij, ocena odvisnosti obravnavanih spremenljivk, pregled nabora funkcij in izbira ustreznih kopul, določitev parametrov izbranih funkcij, uporaba statistični testov, kriterijev ujemanja in grafični prikazov ujemanja za določitev najprimernejše kopule ter izračun različni skupni ter pogojni (angl. *conditional*) povratni dob.

Prva skupina korakov je identična tistim pri univariatni verjetnostni analizah. Zapišemo lahko, da je izvedba klasični verjetnostni analiz, kjer upoštevamo le eno spremenljivko, predpogoj za izvedbo verjetnostni analiz z uporabo funkcije kopula.

Robne porazdelitvene funkcije so tiste, ki najbolj opišejo spremenljivke pri univariatnih verjetnostnih analizah (Karmakar & Simonovic, 2009). Za prikaz odvisnosti spremenljivk, ki določajo visokovodne valove, lahko uporabimo grafične prikaze kot so: diagram K (Kendallov diagram), diagram Chi (Genest & Favre, 2007) ali razsevni diagram podatkov (Brilly & Šraj, 2005). Fisher in Swatzer (2001) sta podala več informacij o diagramu Chi, Genest in Boies (2003) pa o diagramu K.

Za oceno parametrov kopul lahko podobno kot pri univariatnih verjetnostnih analizah uporabimo različne metode: metodo momentov (angl. *method of moments*) (Genest & Favre, 2007), metodo največjega verjetja (angl. *maximum likelihood method*) (Dupuis, 2007), pseudo metodo največjega verjetja (angl. *maximum pseudo-likelihood method*) (Genest & Favre, 2007). Pri metodi momentov parametre določimo s pomočjo Kendallovega ali Spearmanovega koeficienta korelacije (Genest & Favre, 2007). Ta način ocenjevanja parametrov je v praksi pogost (Genest & Favre, 2007; Karmakar & Simonovic, 2009; Zhang & Singh, 2006). Kendallov koeficient korelacije temelji na vrstnem redu podatkov in ga lahko izračunamo z naslednjim izrazom (Genest & Favre, 2007):

$$\tau = \frac{4}{n(n-1)} P_n - 1, \quad (1)$$

kjer P_n označuje število parov spremenljivk (X, Y) za katere velja: $(x_i - x_j)(y_i - y_j) > 0$, n pa je velikost vzorca.

Na podlagi Kendallovega koeficienta korelacije lahko določimo kopule, ki so primerne za modeliranje odvisnosti med obravnavanimi spremenljivkami. Tako lahko z nekaterimi funkcijami opišemo le pozitivne vrednosti Kendallovega koeficienta korelacije, nekatere pa so primerne le za vrednosti koeficienta blizu vrednosti 0 (neodvisni spremenljivki). Razvitih je že veliko različnih tipov kopul, ki pripadajo posameznim družinam kopul, kot so: Arhimedova družina (angl. *Archimedean copulas*), eliptična družina (angl. *Elliptical copulas*), družina kopul ekstremnih vrednosti (angl. *Extreme value copulas*), Farlie-Gumbel-Morgensternova družina kopul (Genest & Favre, 2007) in še nekatere druge (Nelsen, 1999). V prispevku smo uporabili le kopule iz Arhimedove družine kopul (preglednica 1). Kopula Gumbel-Hougaard je primerna le za pozitivne vrednosti Kendallovega koeficienta korelacije, medtem ko lahko kopuli Frank in Clayton uporabimo tudi pri negativnih vrednostih korelacijskega koeficienta. Do negativnih vrednosti lahko pride predvsem pri hkratnem obravnavanju konic pretokov in časov trajanja visokovodnih valov.

Preglednica 1: Osnovne lastnosti nekaterih kopul iz Arhimedove družine.

Table 1: Some properties of copulas from Archimedean family.

Ime / Name	$C_\theta(u, v)$	τ	$\theta \in$
Gumbel-Hougaard	$\exp\{-((-\ln u)^\theta + (-\ln v)^\theta)^{1/\theta}\}$	$1 - \theta^{-1}$	$[1, \infty)$
Clayton (Cook-Johnson)	$\{u^{-\theta} + v^{-\theta} - 1\}^{-1/\theta}$	$\frac{\theta}{\theta + 2}$	$[-1, \infty) \setminus \{0\}$
Frank	$-\frac{1}{\theta} \ln \left\{ 1 + \frac{(\exp(-\theta u) - 1)(\exp(-\theta v) - 1)}{\exp(-\theta) - 1} \right\}$	$1 - \frac{4}{\theta} \left(1 - \frac{1}{\theta} \int_0^\theta \frac{t}{e^t - 1} dt \right)$	$(-\infty, \infty) \setminus \{0\}$

Multivariatno porazdelitveno funkcijo H , ki jo določa n spremenljivk, lahko zapišemo z naslednjim izrazom (Zhang & Singh, 2006):

$$H(x_1, \dots, x_n) = C_\theta\{F_{x_1}(x_1), \dots, F_{x_n}(x_n)\}. \quad (2)$$

Porazdelitvena funkcija H za bivariatni primer pa ima naslednjo obliko:

$$H(x, y) = C_\theta\{F(x), G(y)\}, \quad (3)$$

kjer sta $F(x)$ in $G(y)$ robni porazdelitveni funkciji, C_θ pa označuje teoretično funkcijo kopula (preglednica 1). V preglednici 1 so zapisani tudi izrazi, s katerimi lahko določimo parameter kopule θ na podlagi Kendallovega koeficienta korelacije τ (metoda momentov) in območje vrednosti parametra izbrane funkcije kopula.

Z uporabo statističnih testov, kriterijev ujemanja in grafičnih prikazov ujemanja lahko določimo kopulo, ki se najbolj prilega vzorcu letnih maksimumov. Večina testov in kriterijev ujemanja temelji ali na Kendallovi porazdelitveni funkciji (Genest & Boies, 2003; Genest & Favre, 2007) ali na empirični kopuli (Genest & Favre, 2007). Skupna empirična verjetnost (angl. *empirical joint probability*) je določena z naslednjim izrazom (Karmakar & Simonovic, 2009; Zhang & Singh, 2006):

$$H(x_i, y_i) = P(X \leq x_i, Y \leq y_i) = \frac{P_n}{n+1}. \quad (4)$$

Najprej je potrebno vzorec urediti v naraščajočem redu po spremenljivki X , nato pa s štetjem parov, za katere velja $x_j \leq x_i$ in $y_j \leq y_i$, določimo skupno empirično verjetnost posameznega para (x_i, y_i) . V enačbi (4) n označuje velikost vzorca oziroma število let podatkov, ki jih analiziramo. Ane in Kharoubi (2003) sta za izbiro kopule predlagala dva kriterija ujemanja, ki temeljita na primerjavi empirične in teoretične funkcije kopula (Anderson-Darling in Integrirani Anderson-Darling):

$$AD = \max_{1 \leq u \leq n, 1 \leq v \leq n} \frac{|C_n - C_\theta|}{\sqrt{C_\theta(1 - C_\theta)}}, \quad (5)$$

$$IAD = \sum_{u=1}^n \sum_{v=1}^n \frac{(C_n - C_\theta)^2}{C_\theta(1 - C_\theta)}, \quad (6)$$

kjer je C_θ teoretična funkcija kopula, C_n pa je empirična kopula (Genest & Favre, 2007). Kriterija ujemanja, ki sta zapisana z enačbama (5) in (6), se lahko uporabita le za primerjavo posameznih kopul, ne moremo pa jih uporabiti za testiranje hipotez (Genest et al., 2009). Gre za podoben tip kriterijev ujemanja kot RMSE ali MAE (Šraj et al., 2012) testa.

Zadnji korak pri uporabi funkcij kopula je določitev povezave med ocenjenimi vrednostmi spremenljivk in različnimi skupnimi (ali pogojnimi) povratnimi dobami. Salvadori in sodelavci (Salvadori et al., 2007) so v teorijo kopul vpeljali dve skupni primarni (angl. primary) povratni dobi:

- OR, ko se zgodi vsaj eden od obravnavanih dogodkov,
- AND, ko se zgodita oba obravnavana dogodka hkrati.

Izraza za izračun skupnih povratnih dob imata pri metodi letnih maksimumov naslednjo obliko (Salvadori et al., 2007):

$$T_{u,v}^{OR} = \frac{1}{1 - C_\theta(u,v)}, \quad (7)$$

$$T_{u,v}^{AND} = \frac{1}{1 - u - v + C_\theta(u,v)}, \quad (8)$$

kjer so uporabljene spremenljivke razložene že pri enačbi 3.

4. Rezultati

4.1 Rezultati univariatnih verjetnostnih analiz

Rezultat verjetnostne analize je povezava med ocenjeno vrednostjo spremenljivke in povratno dobo oziroma verjetnostjo. Šraj s sodelavci (Šraj et

al., 2012) je prikazala rezultate verjetnostnih analiz ter postopek izbire najustreznejše porazdelitvene funkcije v primeru podatkov z vodomerne postaje Litija I na reki Savi. Pri analizi konic smo kot najprimernejšo določili logaritemsko Pearsonovo III porazdelitev (LP3), pri analizi volumnov visokovodnih valov pa smo izbrali Pearsonovo III porazdelitev (P3) (Šraj et al., 2012). Metoda momentov L je pokazala boljše rezultate kot metoda momentov in metoda največjega verjetja (Šraj et al., 2012).

Preglednica 2: Ocenjene vrednosti pretokov in volumnov.

Table 2: Estimated discharge and wave volume values.

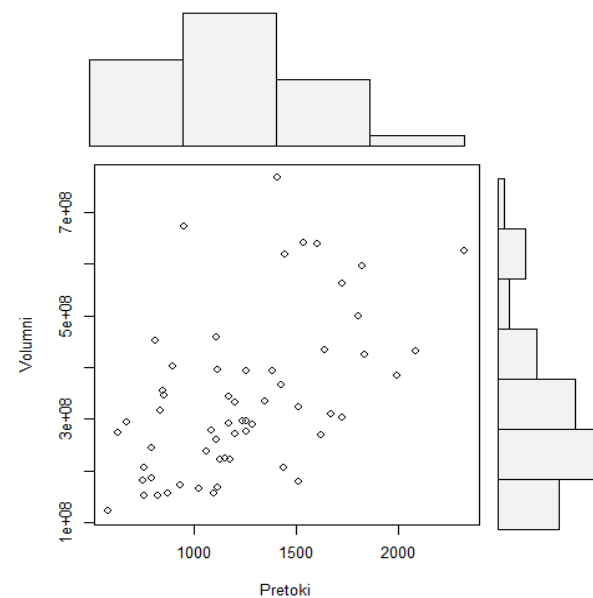
Povratna doba/Return period	Q-LP3 [m ³ /s]	V-P3 [10 ⁶ m ³]
2	1184,2	305,6
10	1775,1	546,9
50	2237,4	751,5
100	2422,6	835,2
500	2835,6	1023,3

V preglednici 2 so prikazane ocenjene vrednosti pretokov (LP3) in volumnov (P3) za nekatere značilne povratne dobe (univariatne) pri ocenjevanju parametrov z metodo momentov L. Izbrane porazdelitve so izkazale najboljše rezultate pri večini uporabljenih grafičnih in statističnih testov (Šraj et al., 2012).

4.2 Rezultati bivariatnih verjetnostnih analiz s funkcijo kopula

Pri izvedbi verjetnostne analize z uporabo funkcije kopula smo sledili postopku, ki je opisan v poglavju 3.2. Rezultate prve skupine korakov je prikazala Šraj s sodelavci (Šraj et al., 2012). Za robno porazdelitveno funkcijo pretokov smo določili logaritemsko Pearsonovo III porazdelitev (parametri: $\alpha = 182,88$; $\beta = -0,01$; $c = 4,99$), za robno porazdelitev volumnov visokovodnih valov pa smo izbrali Pearsonovo III porazdelitev (parametri: $\alpha = 2,53$; $\beta = -98,2 \cdot 10^6$; $c = 88,7 \cdot 10^6$). Na sliki 2 je prikazan razsevni diagram letnih

maksimumov. Vidimo lahko, da je za obravnavani spremenljivki značilna pozitivna odvisnost. Vrednosti Pearsonovega, Kendallovega in Spearmanovega koeficienta korelacije so: 0,525; 0,385; 0,541 in potrjujejo ugotovitve s slike 2.



Slika 2: Razsevni graf pretokov [m³/s] in volumnov [m³] visokovodnih valov Save v Litiji (1953-2010).

Figure 2: Scatter plot of discharges [m³/s] and wave volumes [m³] for Sava in Litija (1953-2010).

Z enačbami iz preglednice 1 in s Kendallovim koeficientom korelacije (metoda momentov) lahko določimo parametre izbranih teoretičnih kopul iz Arhimedove družine. Vrednosti parametrov za kopule Gumbel-Hougaard, Clayton in Frank so: 1,624; 1,248 in 3,947. Glede na izračunano vrednost Kendallovega koeficienta korelacije lahko uporabimo vse tri funkcije kopula iz preglednice 1.

Ko smo določili parametre enoparametrskih kopul, lahko s pomočjo statističnih testov, kriterijev ujemanja in grafičnih prikazov ujemanja izberemo funkcijo, ki se najboljše prilega vzorcu letnih maksimumov.

Prvi grafični prikaz ujemanja je sestavljen iz generiranja velikega števila podatkov. Najprej smo generirali podatke, ki pripadajo testirani teoretični kopuli. Tako smo določili 10.000 parov podatkov, ki pripadajo testirani kopuli s parametrom, ki je bil določen iz vzorca letnih maksimumov s hidrološke

postaje Litija I na reki Savi. Generirane pare podatkov smo nato primerjali z merjenimi. Ujemanje med generiranimi (črne oznake) in izmerjenimi podatki (rdeče oznake) je prikazano na slikah 3 in 4. Za transformacijo generiranih podatkov smo uporabili izbrani robni porazdelitveni funkciji. Iz slik 3 in 4 lahko vidimo, da prikazane kopule iz Arhimedove družine izkazujejo različno obnašanje. Razlike so očitne predvsem v primeru ekstremnih dogodkov (desni zgornji kot grafov; angl. *upper tail dependence*) in v primeru dogodkov z nizkimi povratnimi dobami (levi spodnji kot grafov; angl. *lower tail dependence*). Vidimo lahko, da ima kopula Clayton večjo gostoto dogodkov v levem spodnjem kotu (slika 3), zato ta funkcija ni najprimernejša za modeliranje podatkov z vodomerne postaje Litija I. V vzorcu letnih maksimumov ni veliko dogodkov, ki imajo relativno majhne vrednosti pretokov in volumnov visokovodnih valov. Kopula Gumbel-Hougaard ima večji poudarek na dogodkih, za katere so značilne velike vrednosti pretokov in volumnov. Kopula Frank nima poudarka niti na levem spodnjem niti na desnem zgornjem robu. Zaradi teh lastnosti lahko z uporabo kopule Frank podcenimo opazovani pojav (Poulin et al., 2007). Ujemanje med merjenimi podatki (točke) in teoretičnimi funkcijami kopula (črte) je prikazano tudi na sliki 5. Vidimo lahko, da se s podatki s hidrološke postaje Litija I nekoliko slabše ujema kopula Clayton. Kopuli Frank in Gumbel-Hougaard pa sta izkazali dokaj podobne rezultate. Za natančnejšo določitev razlik med uporabljenimi funkcijami kopula je potrebno uporabiti statistične teste ali kriterije ujemanja. Za izbrane funkcije smo uporabili Anderson-Darlingov in Integrirani Anderson-Darlingov kriterij ujemanja (enačbe (5) in (6)). Rezultati kriterijev ujemanja so prikazani v preglednici 3. Najboljše ujemanje med empirično

in teoretično funkcijo kopula smo dobili pri uporabi kopule Gumbel-Hougaard iz Arhimedove družine kopul.

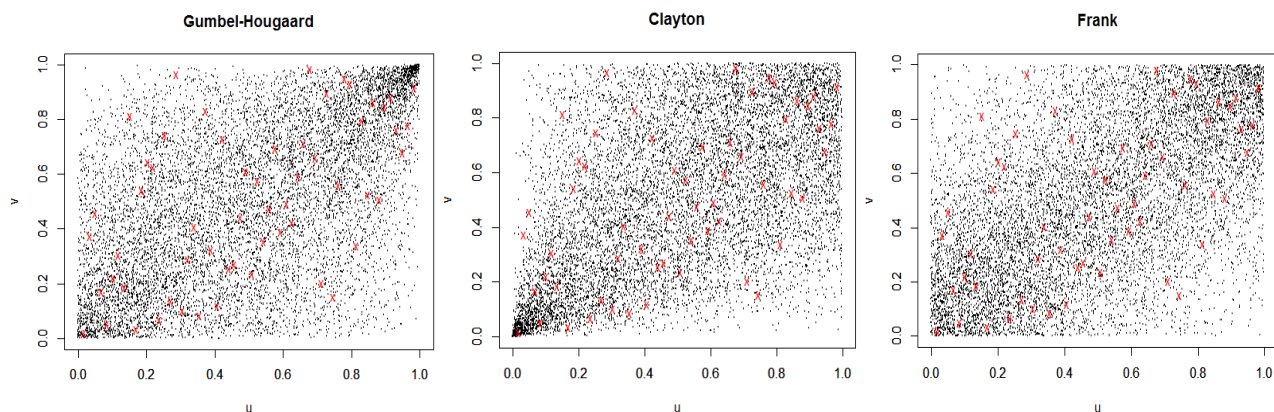
Preglednica 3: Rezultati Anderson-Darlingovega in Integriranega Anderson-Darlingovega testa .

Table 3: Test results for Anderson-Darling and Integrated Anderson-Darling tests.

Funkcija kopula/ Copula function	Anderson- Darling	Integrirani Anderson- Darling
Gumbel-Hougaard	0,16	7,47
Clayton	0,22	17,83
Frank	0,21	9,79

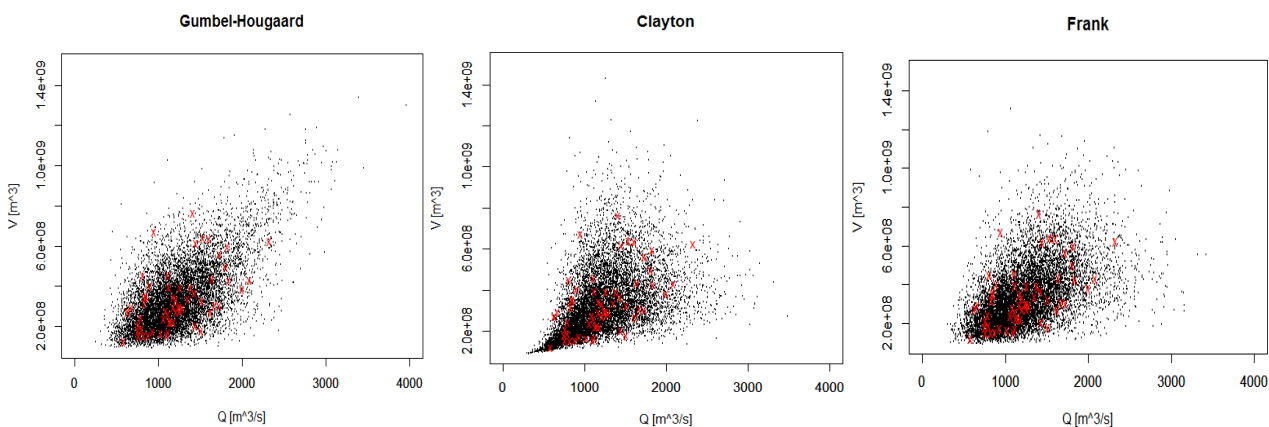
Kopula Gumbel-Hougaard je tudi pri analizah nekaterih drugih avtorjev izkazala najboljše rezultate. Tako sta Zhang in Singh (2006) ugotovila, da je omenjena funkcija dala najboljše rezultate pri analizah podatkov z dveh vodomernih postaj v Kanadi, Karmakar in Simonovic (2009) pa sta zapisala, da je kopula Gumbel-Hougaard dala boljše rezultate kot kopuli Clayton in Ali-Mikhail-Haq pri analizah podatkov s hidrološke postaje v ZDA. Vsi avtorji so ugotovitve podali na podlagi bivariatnih analiz konic in volumnov visokovodnih valov.

V preglednici 4 so zapisane skupne povratne dobe za primer OR, ki označuje situacijo, ko se zgodi vsaj eden izmed obravnavanih pojavov. V preglednici 5 so prikazane skupne povratne dobe za primer AND, ko se zgodita oba dogodka hkrati, tako konica kot volumen visokovodnega vala. Rezultati so izračunani z uporabo kopule Gumbel-Hougaard iz Arhimedove družine kopul, kjer smo parameter funkcije določili na podlagi Kendallovega koeficienta korelacije (metoda momentov).



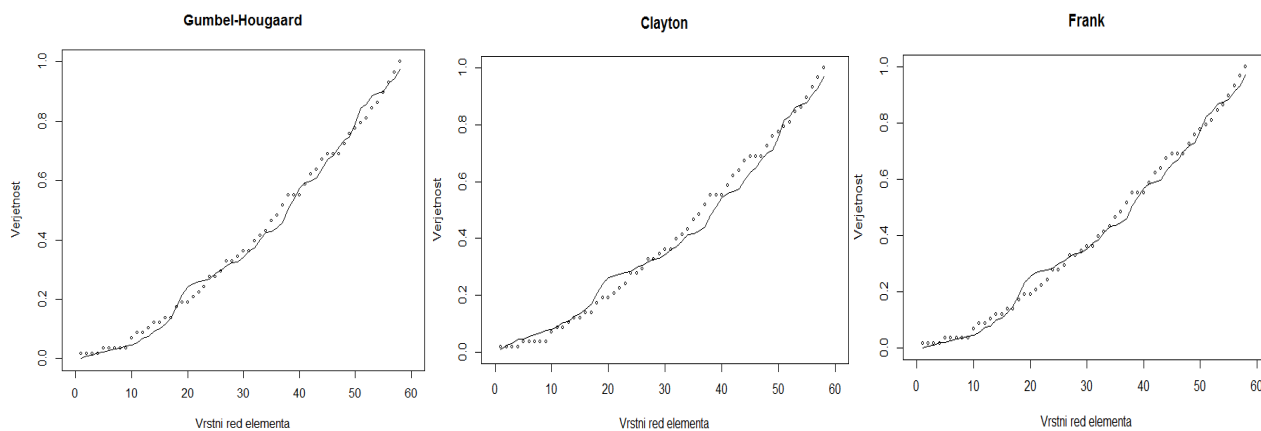
Slika 3: Prikaz ujemanja med generiranimi in izmerjenimi podatki za tri teoretične funkcije kopula.

Figure 3: Presentation of graphical fit between generated and measured data for three copula functions.



Slika 4: Prikaz ujemanja med generiranimi in izmerjenimi podatki pri uporabi robnih porazdelitvenih funkcij.

Figure 4: Presentation of graphical fit between generated and measured data in case of use of marginal distribution functions.



Slika 5: Primerjava med skupno empirično verjetnostjo in teoretičnimi funkcijami kopula.

Figure 5: Comparison between joint empirical probabilities and theoretical copula functions.

Preglednica 4: Vrednosti povratnih dob za primer OR (zgodí se vsaj eden izmed dogodkov).

Table 4: Return period values for the OR case (at least one event occurs).

V [10^6 m^3]/Q [m^3/s]	1000	1500	2000	2500
100	1,0	1,0	1,0	1,0
250	1,2	1,5	1,5	1,5
500	1,4	3,6	6,5	7,1
750	1,4	4,3	18,7	44,4

Preglednica 5: Vrednosti povratnih dob za primer AND (zgodita se oba dogodka).

Table 5: Return period values for the AND case (both events occur).

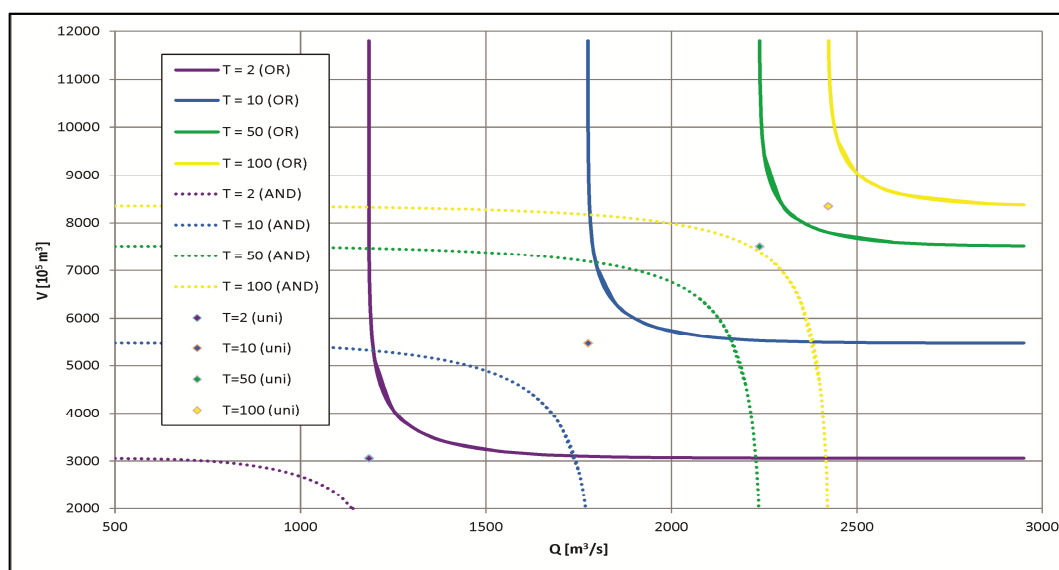
V [10^6 m^3]/Q [m^3/s]	1000	1500	2000	2500
100	1,4	4,3	21,3	134,4
250	1,9	4,8	22,0	135,7
500	7,5	10,6	28,6	146,3
750	50,3	54,9	74,1	195,4

4.3 Primerjava rezultatov

Preglednica 6: Primerjava povratnih dob za nekatere ekstremne dogodke obravnavanega obdobja.

Table 6: Comparison of return periods for some extreme events in the discussed period.

Leto dogodka/Year of the event	Q [m^3/s]	V [10^6 m^3]	$T^{\text{UNI-Q}}$ [let/years]	$T^{\text{UNI-V}}$ [let/years]	T^{OR} [let/years]	T^{AND} [let/years]
1965	1837	425	12,2	4,2	3,9	15,9
1968	1411	766	3,4	56,3	3,4	60,7
1982	1994	383	17,6	3,2	3,1	23,7
1990	2326	624	69,5	18,0	17,0	91,9
2007	1510	179	4,5	1,1	1,1	4,6
2010	2087	431	29,0	4,4	4,3	33,6



Slika 6: Primerjava med skupnimi povratnimi dobami (bivariatni primer) in univariatnimi povratnimi dobami.

Figure 6: Comparison between joint bivariate return periods (bivariate case) and the classical univariate return periods.

Na sliki 6 je prikazana primerjava med univariatnimi in skupnimi bivariatnimi povratnimi dobami. Prikazani so rezultati tako za primer OR kot za primer AND. Opazimo lahko, da velja pravilo: $T^{OR} < T^{UNI} < T^{AND}$ (slika 6). Do podobnih ugotovitev so prišli tudi nekateri drugi raziskovalci (Salvadori & De Michele, 2007; Salvadori et al., 2007). Te ugotovitve so deloma pričakovane, saj je verjetnost, da se zgodita oba dogodka, konica in volumen hkrati, manjša kot verjetnost, da se zgodi vsaj eden izmed njiju. Univariatne povratne dobe smo izračunali z uporabo logaritemsko Pearsonove III porazdelitve za konice in Pearsonove III porazdelitve za volumne, kjer smo parametre določili z uporabo metode momentov L (Šraj et al., 2012).

Primerjava med različnimi povratnimi dobami nekaterih ekstremnih dogodkov, ki so se zgodili v opazovanem obdobju, je predstavljena v preglednici 6. Maksimalna konica v opazovanem obdobju se je zgodila leta 1990. Izmerjena vrednost pretoka je takrat znašala $2326 \text{ m}^3/\text{s}$, pripadajoči volumen visokovodnega vala pa je znašal $624 \cdot 10^6 \text{ m}^3$ (preglednica 6). ARSO (Kobold, 2011) je temu dogodku pripisal več kot stoletno povratno dobo za konico pretoka, z upoštevanjem podatkov do leta 2008. Iz preglednice 6 lahko vidimo, da smo z uporabo univariatnih verjetnostnih analiz (metoda momentov L) in nekoliko daljšega niza podatkov ocenili vrednost tega dogodka na 70 let, pripadajočega volumna pa na 18 let. Poleg tega lahko iz rezultatov vidimo, da skupne bivariatne povratne dobe precej odstopajo od univariatnih. Tako povratna doba za dogodek iz leta 1990 za primer OR znaša 17 let, za primer AND pa 92 let. Za val iz leta 2010 je ARSO (Kobold, 2011) določil povratno dobo konice 50 let, iz preglednice 6 lahko vidimo, da smo z upoštevanjem nekoliko daljšega niza izračunali manjšo povratno dobo, manjša pa je tudi skupna povratna doba za primer AND. Ker lahko v nekaterih primerih dogodek, ki ima le nadpovprečno veliko konico, povzroči manj škode kot visokovodni val, za katerega sta značilna povprečna konica in volumen (Renard & Lang, 2007), je hkratno upoštevanje obeh spremenljivk v analizah smiselno. Z upoštevanjem le ene izmed

spremenljivk lahko namreč podcenimo dejansko skupno povratno dobo opazovanega pojava, kot jo da primer AND. (Salvadori & De Michele, 2007; Salvadori et al., 2007).

5. Zaključki

S podatki visokovodnih valov s hidrološke postaje Litija I na reki Savi smo najprej naredili klasične univariatne verjetnostne analize (Šraj et al., 2012). Te so predpogoj za izvedbo verjetnostnih analiz z uporabo funkcij kopula (določitev robnih porazdelitvenih funkcij). Za oceno parametrov izbranih teoretičnih kopul iz Arhimedove družine smo uporabili metodo momentov, ki temelji na povezavi med Kendallovim koeficientom korelacije ter parametrom kopule in se v praksi pogosto uporablja (Genest & Favre, 2007; Karmakar & Simonovic, 2009; Zhang & Singh, 2006). Z uporabo nekaterih statističnih in grafičnih kriterijev ujemanja smo ugotovili, da se s podatki najbolje ujema kopula Gumbel-Hougaard. Ta funkcija je tudi pri nekaterih drugih študijah visokovodnih valov dala dobre rezultate (Karmakar & Simonovic, 2009; Zhang & Singh, 2006). Zhang in Singh (2006) sta ugotovila, da se je funkcija kopula boljše ujemala s podatki kot klasična bivariatna normalna porazdelitev. V študiji smo izračunali tudi nekatere skupne povratne dobe (primera OR in AND) ter ugotovili, da lahko z upoštevanjem le ene izmed spremenljivk, ki določa visokovodni val, s precejšno razliko ocenimo povratno dobo opazovanega dogodka (preglednica 6). Izračunali smo tudi skupne povratne dobe za nekatere ekstremne dogodke, ki so se zgodili na območju srednjega toka reke Save v obravnavanem obdobju. Poleg tega smo ugotovili, da je povratna doba v primeru OR manjša kot pri primeru AND, kar je v skladu z rezultati drugih raziskovalcev (Salvadori & De Michele, 2007; Salvadori et al., 2007).

V hidrološki praksi so analize z uporabo funkcije kopula šele v razvoju in še ne povsem uveljavljene. V primerjavi s klasičnimi, univariatnimi verjetnostnimi analizami, je postopek uporabe kopule precej bolj zapleten, potrebno je izvesti

številne korake, ki pri univariatnih analizah niso potrebni. Končni rezultati pa so večinoma drugačni od tistih, ki jih dobimo pri klasičnih analizah. Pri pomembnejših hidrotehničnih posegih, kot je dimenzioniranje hidrotehničnih objektov, je torej v verjetnostnih analizah smiselno in priporočljivo upoštevati več spremenljivk visokovodnih valov hkrati, kar pa lahko naredimo z uporabo funkcije kopula, kot je prikazano v prispevku.

Zahvala

Zahvaljujemo se Agenciji RS za okolje za posredovane podatke z vodomerne postaje Litija I na reki Savi. Rezultati raziskave so nastali v okviru temeljnega raziskovalnega projekta J2-4096, ki ga financira Javna agencija za raziskovalno dejavnost Republike Slovenije.

Viri

Agencija RS za okolje (2013). Atlas okolja. http://gis.arso.gov.si/atlasokolja/profile.aspx?id=AtlasOkolja_AXL@Arso (Pridobljeno 6. 2. 2013).

Ane, T., Kharoubi, C. (2003). Dependence structure and risk measure, *Journal of Business*, **76**(3), 411–438.

Balistrocchi, M., Bacchi, B. (2011). Modelling the statistical dependence of rainfall event variables through copula functions, *Hydrology and Earth System Sciences*, **15**(6), 1959–1977.

Bezjak, N. (2012). Verjetnostna analiza visokovodnih konic z metodo vrednosti nad izbranim pragom in z metodo letnih maksimumov = Flood frequency analysis with peaks over threshold method and annual maximum series method. Thesis, Univerza v Ljubljani, FG, 106 p. (in Slovenian).

Bezjak, N., Brilly, M., Šraj, M. (2013). Comparison between the peaks over threshold method and the annual maximum method for flood frequency analyses, *Hydrological Sciences Journal*. DOI:10.1080/02626667.2013.831174

Brilly, M., Polič, M. (2005). Public perception of flood risks, flood forecasting and mitigation, *Natural Hazards and Earth System Sciences*, **5**(3), 345–355.

Brilly, M., Šraj, M. (2005). *Osnove hidrologije*. Ljubljana, Univerza v Ljubljani, Fakulteta za gradbeništvo in geodezijo, 309 p. (in Slovenian).

Chen, L., Singh, V. P., Shenglian, G., Hao, Z., Li, T. (2012). Flood Coincidence Risk Analysis Using Multivariate Copula Functions, *Journal of Hydrologic Engineering*, **17**(6), 742–755.

De Michele, C., Salvadori, G., Canossi, M., Petaccia, A., Rosso, R. (2005). Bivariate Statistical Approach to Check Adequacy of Dam Spillway, *Journal of Hydrologic Engineering*, **10**(1), 50–57.

Dupuis, D. J. (2007). Using copulas in hydrology: Benefits, cautions, and issues, *Journal of Hydrologic Engineering*, **12**(4), 381–393.

Đurović, B., Mikoš, M. (2004). Preventivno obvladovanje tveganj zaradi naravnih nevarnosti – postopki v alpskih državah in Sloveniji = Preventive management of risks due to natural hazards – Procedures in the alpine countries and in Slovenia, *Acta hydrotechnica*, **22**(36), 17–35.

Favre, A.C., El Adlouni, S., Perreault, L., Thiemonge, N., Bobee, B. (2004). Multivariate hydrological frequency analysis using copulas, *Water Resources Research*, **40**(1), 1–12.

Fisher, N. I., Switzer, P. (2001). Graphical assessment of dependence: Is a picture worth 100 tests?, *American Statistician*, **55**(3), 233–239.

Ganguli, P., Reddy, M. J. (2013). Probabilistic assessment of flood risks using trivariate copulas, *Theoretical and Applied Climatology*, **111**(1-2), 341–360.

Genest, C., Boies, J. C. (2003). Detecting dependence with Kendall plots, *American Statistician*, **57**(4), 275–284.

Genest, C., Favre, A. C. (2007). Everything You Always Wanted to Know about Copula Modeling but Were Afraid to Ask, *Journal of Hydrologic Engineering*, **12**(4), 347–368.

Genest, C., Remillard, B., Beaudoin, D. (2009). Goodness-of-fit tests for copulas: A review and a power study, *Insurance: Mathematics and Economics*, **44**(2), 199–213.

Grimaldi, S., Serinaldi, F. (2006). Design hyetograph analysis with 3-copula function, *Hydrological Sciences-Journal-des Sciences Hydrologiques*, **51**(2), 223–238.

Karmakar, S., Simonovic, S. P. (2008). Bivariate flood frequency analysis: Part 1. Determination of marginals by parametric and nonparametric techniques, *Journal of Flood Risk Management*, **1**(4), 190–200.

- Karmakar, S., Simonovic, S. P. (2009). Bivariate flood frequency analysis: Part 2: a copula-based approach with mixed marginal distributions, *Journal of Flood Risk Management*, **2(1)**, 32–44.
- Kobold, M. (2011). Primerljivost poplave septembra 2010 z zabeleženimi zgodovinskimi poplavnimi dogodki = Comparison of Floods in September 2010 with Registered Historic Flood Events, *Ujma*, **25**, 48–56 (In Slovenian).
- Kobold, M., Zgonc, A., Sušnik, M. (2005). Nezanesljivost padavinskih meritev in napovedi pri modeliranju hudourniških poplav = Uncertainty of precipitation measurements and predictions in flash flood modelling, *Acta hydrotechnica*, **23(39)**, 79–98 (in Slovenian).
- Ma, M., Song, S., Ren, L., Jiang, S., Song, J. (2011). Multivariate drought characteristics using trivariate Gaussian and Student-t copulas, *Hydrological Processes*, **27(8)**, 1175–1190.
- Mikoš, M., Brilly, M., Ribičič, M. (2004). Poplave in zemeljski plazovi v Sloveniji = Floods and landslides in Slovenia, *Acta hydrotechnica*, **22(37)**, 113–133 (in Slovenian).
- Nelsen, R. B. (1999). *An introduction to copulas*. Springer, New York, 269 p.
- Poulin, A., Huard, D., Favre, A.C., Pugin, S. (2007). Importance of tail dependence in bivariate frequency analysis. *Journal of Hydrologic Engineering*, **12(4)**, 394–403.
- Renard, B., Lang, M. (2007). Use of a Gaussian copula for multivariate extreme value analysis: Some case studies in hydrology, *Advances in Water Resources*, **30**, 897–912.
- Salvadori, G., De Michele, C., Kottegoda, N. T., Rosso, R. (2007). *Extremes in nature an approach using Copulas*. Springer, Dordrecht, 292 p.
- Salvadori, G., De Michele, C. (2007). On the use of copulas in hydrology: Theory and practice. *Journal of Hydrologic Engineering*, **12(4)**, 369–380.
- Šraj, M., Bezak, N. (2013). Analiza visokovodnih valov Save v Litiji = The analysis of flood waves on the Sava River in Litija, *Ujma*, **27**, 228–235. (In Slovenian).
- Šraj, M., Bezak, N., Brilly, M. (2012). Vpliv izbire metode na rezultate verjetnostnih analiz konic, volumnov in trajanj visokovodnih valov Save v Litiji = The influence of the choice of method on the results of frequency analysis of peaks, volumes and durations of flood waves of the Sava River in Litija, *Acta hydrotechnica*, **25(42)**, 41–58. (In Slovenian).
- Turk, G. (2012). *Verjetnostni račun in statistika = Probability and Statistics*. Ljubljana, Univerza v Ljubljani, Fakulteta za gradbeništvo in geodezijo, 264 p. (in Slovenian).
- Uлага, F. (2011). Hidrološka postaja Litija na Savi = Hydrological station Litija on the Sava River, *Naše okolje*, **18(8)**, 81–85 (in Slovenian).
- Wong, G., Lambert, M. F., Leonard, M., Metcalfe, A. V. (2010). Drought Analysis Using Trivariate Copulas Conditional on Climatic States, *Journal of Hydrologic Engineering*, **15(2)**, 129–141.
- Zhang, L., Singh, V. P. (2006). Bivariate Flood Frequency Analysis Using the Copula Method, *Journal of Hydrologic Engineering*, **11(2)**, 150–164.
- Zhang, L., Singh, V. P. (2007). Trivariate Flood Frequency Analysis Using the Gumbel-Hougaard Copula, *Journal of Hydrologic Engineering*, **12(4)**, 431–439.