

TESTOVANIE DVOJROZMERNÝCH ARCHIMEDOVSKÝCH KOPULA FUNKCIÍ PRI KOINCIDENCII PRIETOKOV HLAVNÉHO TOKU A PRÍTOKU

Veronika Bačová Mitková, Dana Halmová

Príspevok prezentuje testovanie a výber vhodnej kopula funkcie pre dvojrozmernú analýzu synchronne sa vyskytujúcich maximálnych prietokov na hlavnom toku a jeho prítoku. Analyzované boli synchronne priemerné denné prietoky na hlavných slovenských tokoch a ich prítokoch v povodí rieky Dunaj. V prvej časti príspevku je prezentovaná príprava vstupných údajov a korelačná analýza údajov. V ďalšej časti je prezentovaný výber marginálnych distribučných funkcií a výber vhodnej funkcie pre výpočet dvojrozmernej združenej distribučnej funkcie výskytu synchronných prietokov. Ako matematický prostriedok na určenie združenej pravdepodobnosti výskytu dvoch závislých premenných bola použitá Archimedovská trieda kopula funkcií. Najmenšie hodnoty chýb odhadu dosiahli pravdepodobnosti vypočítané pomocou Gumbel-Hougaard kopula funkcie a preto bola táto funkcia vybraná ako najvhodnejšia na ilustráciu združenej pravdepodobnosti výskytu synchronných premenných.

KEÚČOVÉ SLOVÁ: maximálny priemerný denný prietok, korelácia, kopula funkcia, štatistický test, združená pravdepodobnosť prekročenia

TESTING THE TWO-DIMENSIONAL ARCHIMEDEAN COPULA FUNCTIONS IN COINCIDENCE OF FLOOD DISCHARGES OF THE MAIN RIVER AND TRIBUTARY. The paper presents testing and selecting an appropriate copula function for two-dimensional analysis of synchronously occurring maximum discharges at the main stream and its tributary. Synchronous mean daily discharges on the main Slovak streams and their tributaries of the Danube river basin were analysed. The first part of the paper presents the preparation of the input data and correlation analysis of the data. The next part, presents the selection of marginal probability distribution functions and the selection of the appropriate copula function for calculation of the two-dimensional joint probability distribution functions of the synchronous discharges. Some Archimedean copula functions were used as a mathematical tool for determining the joint probability distribution. Tested Archimedes copula functions have achieved relatively equal calculated distribution probabilities. Probabilities calculated using the Gumbel-Hougaard copula function reached the smallest error of the estimation. This function has been selected as the most appropriate for illustrating the joint probability of occurrence of the synchronic variables.

KEY WORDS: maximum daily discharge, correlation, copula function, statistical test, joint probability distribution functions

Úvod

Konvenčné jednorozmerné prístupy hodnotenia hydrologických charakteristík nemusia poskytnúť uspokojivé výsledky pre vyhodnotenie povodňových rizík v situáciách, keď povodne nastanú na dvoch, alebo viacerých riekach súčasne a spoja sa. Aj z tohto dôvodu by mal byť ďalej rozvíjaný a definovaný v rámci hydrologických analýz povodňových vln viacrozmerný prístup štatistického analyzovania povodňových udalostí na

vstupných profiloch hlavného toku a jeho prítokov. Tokov s prirodzeným režimom neustále ubúda a narastá počet tokov, ktoré sú ovplyvnené technickými zásahmi. Úpravou toku (ale aj povodia) často dochádza ku koncentrácii odtoku z povodia ako aj k zvýšeniu postupovej rýchlosti povodňovej vlny. Dôsledkom toho, môže dochádzať k stretávaniu kulminácií na hlavnom toku s prítokom, alebo k nárastu prietokov v dôsledku zrýchleného pohybu vody buď na hlavnom toku, alebo prítoku. Napríklad augustová povodeň v roku 2002 v Če-

chách na rieke Vltava a rieke Dyje vykazovala nárast doby opakovania prietokov s rastom plochy povodia, čo bolo spôsobené stretávaním povodňových vln v profiloch riečnej siete, keď dochádzalo k súčasnému zasiahnutiu susedných povodí intenzívnymi zrážkami (Záverečná správa VUV TGM ČR, 2002). Analyzovaním rizika povodní v dôsledku zhody povodňových prietokov na hlavnom toku a jeho prítokov za použitia vybraných marginálnych distribučných funkcií a viacrozmerných (4D) kopula funkcií sa zaoberali napr. Chen a kol, 2012. Autori vyhodnotili ako vhodnú X-Gumbel kopula funkciu a následne odvodili podmienené distribučné funkcie a doby opakovania združených prietokov. Zhodou synchronne sa vyskytujúcich povodňových vln na Dunaji a jeho prítokoch sa napr. zaoberal Prohaska (1999), ktorý pri analýze vychádzal z teórie štatistiky dvojrozsmernej premennej. Modelovaním hydrologických združených udalostí na rieke Moravy s využitím agregáčnych operátorov (napr. kopula funkcie) sa zaoberal vo svojom príspevku Matúš (2009). Modelovaním viacrozmerných časových radov v oblasti prírodných vied sa zaoberali Gaál a kol. (2010). Szolgay et al. (2016) analyzovali závislosť medzi vrcholovým prietokmi a objemom povodňovej udalosti v regionálnom kontexte pomocou kopula funkcií. Jeden z výsledkov analýzy ukázal, že napr. kopula funkcie ako: Galambos, Gumbel and Hüsler-Reiss pomerne dobre modelujú extrémnu závislosť ako iné vybrané kopula funkcie (Clayton, Joe).

Cieľom príspevku je prezentovať výber vhodnej kopule z najčastejšie používaných triedy kopula funkcií v hydrologii – Archimedovské kopula funkcie (Clayton, Gumbel-Hougaard a Frank) pre frekvenčnú analýzu synchronne sa vyskytujúcich prietokov hlavného toku a jeho prítoku. Analyzované budú synchronne priemer-

né denné prietoky na hlavných slovenských tokoch v povodí rieky Dunaj a ich vybraných prítokoch. Analýza zhodných prietokových charakteristík môže významne prispieť k spoľahlivejšiemu hodnoteniu povodňového nebezpečenstva.

Vybrané toky a ich prítoky

V tomto príspevku sme sa zamerali na povodie Dunaja. Na analýzu boli vybrané najväčšie slovenské toky vlievajúce sa priamo do Dunaja a ich najväčšie prítoky v úsekoch s čo najmenším ovplyvnením toku (obr. 1). Ako vstupné údaje boli použité priemerné denné prietoky. Tabuľka 1 uvádza zoznam vybraných tokov a ich prítokov, 12 vodomerných staníc (hlavné toky 8 staníc a prítoky 4 stanice) a analyzované obdobie.

Metodika

Príprava vstupných údajov

Rady ročných maxim obsahujú iba jednu hodnotu za rok a preto sa v tomto rade môže stratiť informácia o vrcholovom prietoku, ktorý je vyšší ako maximum v inom roku. Tieto obmedzenia vedú k využitiu metódy Peak Over Threshold (POT). Rady údajov pri metóde POT obsahujú všetky hodnoty nad zvolenou prahovou hodnotou (Rao a Hamed, 2000). Pre prvý výber dát sa ako prahová hodnota volí hodnota okolo Q_a (dlhodobý priemerný ročný prietok). Táto hodnota je pomerne nízka a súbor vrcholových prietokov vln môže mať príliš veľký rozptyl, môže teda obsahovať vlny, ktoré nemusia byť vrcholmi významných povodní. Postupným zvyšovaním prahovej hodnoty sa dosiahne požadovaný počet udalostí za dané obdobie (rok), ktoré sú nad touto prahovou hodnotou.



Obr. 1. Schematická mapa hlavných povodí Slovenska.

Fig. 1. The Schematic map of the main Slovakia basins.

Tabuľka 1. Zoznam vybraných tokov a ich prítokov, vodomerných staníc a analyzované obdobie
Table 1. List of the selected streams and their tributaries, gauging stations and analyzed periods

Hlavný tok	Vodomerná stanica hl. tok horná (UP)	Vodomerná stanica hl. tok dolná (DWN)	Prítok	Vodomerná stanica prítok	obdobie
Morava	Strážnica	Moravský Svätý Ján	Myjava	Šaštín Stráža	1968 – 2011
Váh	Liptovský Hrádok	Liptovský Mikuláš	Belá	Podbanské	1950 – 2011
Nitra	Chynorany	Nitrianska Streda	Bebrava	Nadlice	1940 – 2011
Hron	Banská Bystrica	Žiar nad Hronom	Slatina	Zvolen	1977 – 2011

Prahová hodnota sa potom zvyčajne volí tak, aby súbor hodnôt zaradených do štatistického spracovania obsahoval v priemere okolo štyroch vrcholových hodnôt za rok. Pri výbere povodňových udalostí musí byť zabezpečená aj ich nezávislosť, a to určitými kritériami, ktoré sú uvedené napr. v práci Bayliss, 1999. V našich výpočtoch bola prahová hodnota zvolená na hladine 40 – 50% dlhodobého maximálneho ročného prietoku tak, aby bola zabezpečená nezávislosť vln a aby boli do analýzy zahrnuté všetky významné udalosti v roku. Na základe denných prietokov boli vybrané prietokové vlny, ktoré by mohli byť zaradené do súboru potrebného k štatistickému spracovaniu. Z týchto vln boli vybrané maximálne priemerné denné prietoky a prislúchajúce prietoky na hlavnom toku a jeho prítoku podľa nižšie uvedenej schémy kombinácie premenných.

Schéma kombinácie premenných

Pri analýze vzájomného vzťahu synchrónnych prietokov povodňových vln na hlavnom toku a jeho prítoku sme skúmali kombinácie nasledovných premenných:

- 1) $Q_{UP(max)} - Q_{cor1(DWN)}$; 2) $Q_{UP(max)} - Q_{cor1(TR)}$; 3) $Q_{DWN(max)} - Q_{cor2(UP)}$; 4) $Q_{DWN(max)} - Q_{cor2(TR)}$; 5) $Q_{TR(max)} - Q_{cor3(UP)}$; 6) $Q_{TR(max)} - Q_{cor3(DWN)}$.

kde:

$Q_{UP(max)}$ – maximálny priemerný denný prietok v hornej stanici na hlavnom toku;

$Q_{cor1(DWN)}$, $Q_{cor1(TR)}$ – korešpondujúce prietoky na hlavnom toku v dolnej stanici a prítoku v čase výskytu maximálneho priemerného denného prietoku na hlavnom toku v hornej stanici;

$Q_{DWN(max)}$ – maximálny priemerný denný prietok v dolnej stanici na hlavnom toku;

$Q_{cor2(UP)}$, $Q_{cor2(TR)}$ – korešpondujúce prietoky na hlavnom toku v hornej stanici a prítoku v čase výskytu maximálneho priemerného denného prietoku na hlavnom toku v dolnej stanici;

$Q_{TR(max)}$ – maximálny priemerný denný prietok na prítoku;

$Q_{cor3(UP)}$, $Q_{cor3(DWN)}$ – korešpondujúce prietoky na hlavnom toku v hornej stanici a dolnej stanici v čase výskytu maximálneho priemerného denného prietoku na prítoku.

Korelačný koeficient vzájomného vzťahu prietokových charakteristík

Korelačný koeficient vzájomného vzťahu prietokových charakteristík pre zvolené kombinácie premenných bol vypočítaný podľa vzťahu (1).

$$R = \frac{1}{n\delta_x\delta_y} \left[\sum_{i=1}^n x_i y_i - \frac{1}{n} \sum_{i=1}^n x_i \sum_{i=1}^n y_i \right] \quad (1)$$

kde

n – počet meraní; δ_x – štandardná odchýlka premennej x ; δ_y – štandardná odchýlka y ; x_i – premenná x ($Q [m^3 s^{-1}]$); y_i – premenná y ($Q [m^3 s^{-1}]$). Na určenie tesnosti vzájomného vzťahu bolo použité nasledovné kritérium (Prohaska, 1999):

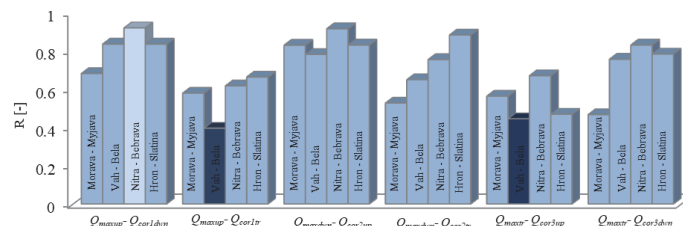
$$|R| \geq 3\sigma_R \quad (2)$$

kde σ_R je odchýlka korelačného koeficienta (rovnica 3), ak $R \gg 0$.

$$\sigma_R = \frac{1-R^2}{\sqrt{n}} \quad (3)$$

Následne boli vypočítané korelačné závislosti medzi zvolenými kombináciami premenných a podľa vzťahu 2 určená tesnosť tejto vzájomnej závislosti (obr.2)

Na základe výsledkov z korelačnej analýzy pre rôzne kombinácie premenných môžeme konštatovať, že všetky kombinácie denných maximálnych prietokov majú štatistickú významnosť a môžu byť použité na ďalšiu viacrozmernú štatistickú (frekvenčnú) analýzu povodňových udalostí na vstupných profiloch hlavného toku a jeho prítoku. Najnižšia korelácia bola zistená medzi maximálnym prietokom na hlavnom toku v hornej stanici a prítokom pre Váh a Belú. Nižší korelačný koeficient pre kombináciu premenných $Q_{UP(max)} - Q_{cor1(TR)}$ a $Q_{TR(max)} - Q_{cor3(UP)}$ na Váhu by mohol byť dôsledkom čiastočne rozdielného povodňového režimu hlavného toku a jeho prítoku. Korelačný koeficient medzi dolnou stanicou na Váhu a prítokom Belou pre kombináciu premenných $Q_{DWN(max)} - Q_{cor2(TR)}$ a $Q_{TR(max)} - Q_{cor3(DWN)}$ dosiahol podstatne vyššie hodnoty.



Obr. 2. Hodnoty korelačného koeficienta R pre volené synchrónne kombinácie premenných.

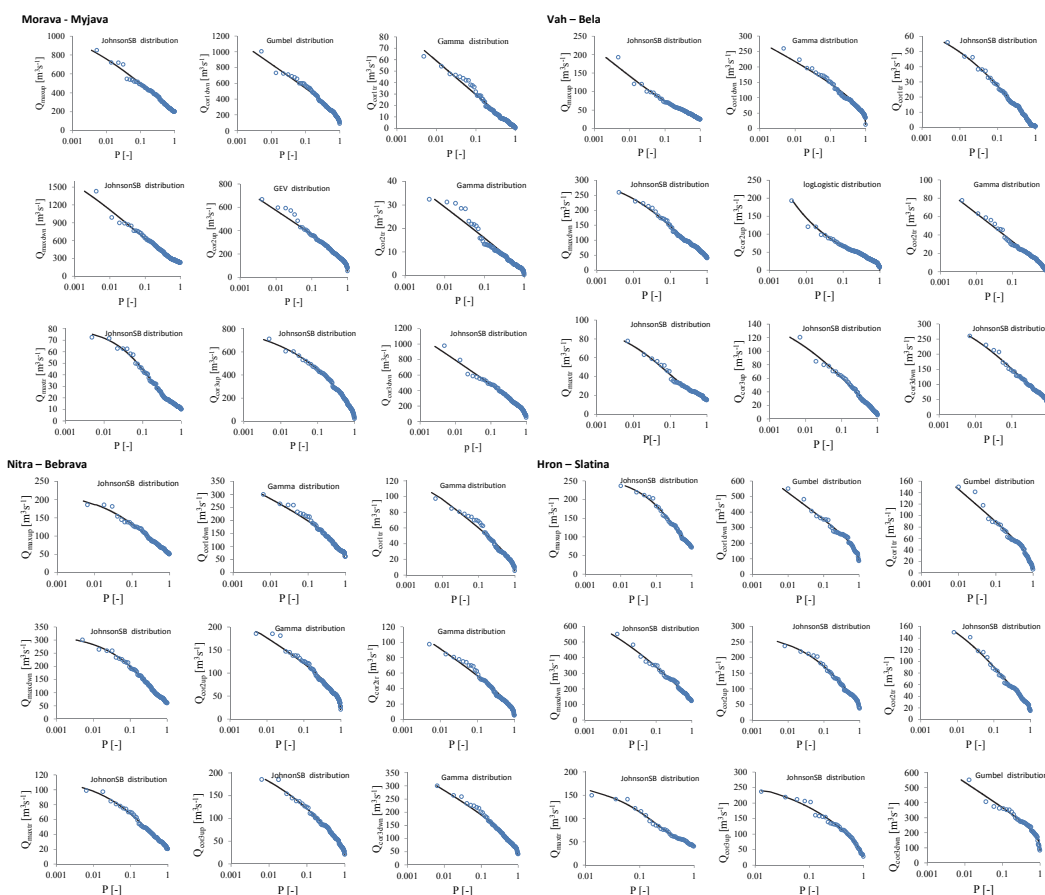
Fig.2. Values of the correlation coefficients R of the various combinations of the synchronous variables.

Atmosférické zrážky patria k časovo a priestorovo najpremenlivejším meteorologickým prvkom. Reliéf, hlavne výškové pomery, členitosť, smer pohoria a poloha zníženín sú v súvislosti s prevládajúcim prúdením vzduchu a zrážkovou expozíciou rozhodujúce pre priestorovú diferenciáciu zrážok. Vplyv týchto charakteristík reliéfu sa v plnej miere uplatňuje aj v povodí horného Váhu (Pekárová a Szolgay, 2005). Najvyššiu korelačnú závislosť ukázala kombinácia premenných $Q_{UP(max)}$ – $Q_{cor1(DWN)}$ medzi Nitrou a Bebravou.

Jednorozmerná analýza synchronných hydrologických premenných

Prvým krokom dvojrozmerné analýzy je identifikácia jednorozmerných (okrajových) rozdelení pravdepodobností hydrologických premenných. Náhodné premenné môžu mať rôzne vlastnosti a preto je potrebné ich transformovať na náhodne premenné s oborom hodnôt v intervale $[0, 1]$. Keď poznáme okrajové rozdelenie, dokážeme oddeliť individuálne správanie sa náhodných premenných od ich štruktúry. Štruktúra závislosti je úplne opísaná združeným rozdelením náhodných premenných získaných z marginálnej distribúcie. Na určenie

parametrov distribučných funkcií sme použili štandardnú metódu MLM (metóda maximálnej virohodnosti). Podľa testov dobrej zhody (Kolmogorov-Smirnov a χ^2) boli vybrané vhodné parametrické rozdelenia pravdepodobnosti. Hodnoty p -value testov ukázali, že nemôžeme zamietnuť hypotézu, že vybrané parametrické rozdelenia pravdepodobnosti dobre vyhovujú pozorovaným údajom na úrovni 5 % významnosti. Parametrické distribúcie JohnsonSB a Gamma sa ukázali ako najlepšie pre väčšinu vybraných premenných odvodených metódou POT. JohnsonSB predstavuje ohraničenú distribúciu s parametrami tvaru γ , δ , lokalizačným parametrom μ a parametrom stupnice σ . Toto rozdelenie súvisí s normálnym rozdelením. Ide o kontinuálnu distribúciu definovanú na ohraničenom rozsahu, ktorá môže byť symetrická alebo asymetrická. Dvoj-parametrické rozdelenie Gamma je veľmi dôležitým modelom v štatistickej hydrológii. Jedná sa o flexibilné rozdelenie schopné mať mnoho rôznych tvarov a je v mnohých krajinách široko používané pre modelovanie povodňových sérií. Empirická čiara prekročenia bola určená pomocou vzorca Cunnane (1978) a porovnaná s vybranými parametrickými rozdeleniami pravdepodobnosti (obr. 3a) -d)).



Obr. 3. a)-d) Porovnanie empirickej čiary rozdelenia pravdepodobnosti so zvolenou parametrickou distribučnou funkciou.

Fig. 3. a)-d) Comparison of the empirical distribution with fitted parametrical distribution functions.

Dvojrozmerná analýza synchrónnych hydrologických premenných

V našej práci sme použili Archimedovskú triedu kopula funkcií pre ich flexibilitu a pomerne ľahkú konštrukciu (tab. 2). Okrem toho táto trieda kopula funkcií zahŕňa celú sadu parametrických kopúl, ktoré pokrývajú širokú škálu štruktúr závislostí. Medzi existujúcimi typmi kopula funkcií je trieda Archimedovských kopúl veľmi populárna a používaná pri hydrologických analýzach (Zhang a Sing, 2006, , De Michele a kol., 2005, Genest a Favre, 2007). Pre každú kopulu možno určiť hodnotu parametra θ pomocou matematického vzťahu (Nelsen, 2006) medzi Kendallovým koeficientom korelácie τ a generujúcou funkciou $\varphi(t)$. Hodnoty odhadnutých parametrov testovaných Archimedovských kopula funkcií sú uvedené v tab. 3.

Testovanie vhodnosti zvolených kopula funkcií

Testovanie vhodnosti kopula funkcie a využívanie testov dobrej zhody je pomerne často diskutovaný problém v odbornej literatúre. Kojadinovic and Yan (2011) porovnávali empirickú kopulu s parametrickým odhadom kopuly odvodeným z nulovej hypotézy. Rozšírili multiplikátorový prístup o multivariačné viacparametrické kopuly a študovali výsledný výsledok testu dobrej zhody (Karmakar and Simonovic (2009); Shiau et al. (2010); Chowdhary et al. (2011)). Chen a Fan (2005) predstavili takzvaný test "pseudo-likelihood ratio", inšpirovaný semi-parametrickou adaptáciou kritéria AIC. Requena a kol. (2012), Szolgay a kol. (2016) využili ako test dobrej zhody na určenie rozdielu medzi empirickou a parametrickou kopula funkciou test založený na Cramér-von Mises kritériu.

Tabuľka 2. Funkcie pravdepodobností, rozsah parametra kopula funkcie, vzťah parametra a Kendallovho τ a funkcia generátora kopule $\varphi(t)$ – Archimedovské kopula funkcie
Table 2. Probability functions, parameter space, generating function and relationship of non-parametric dependence measure with association parameter for the most frequently used Archimedean copulas in hydrology

Kopula	$C(u, v, \theta)$	parameter θ	1 Kendall's τ	2 generátor $\varphi(t)$
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	$-\frac{1}{\theta} \ln[1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{(e^{-\theta} - 1)}]$	$(-\infty, \infty) \setminus \{0\}$	$1 + \frac{4}{\theta}[D_1(\theta^*) - 1]$	$-\ln \frac{e^{-\theta t} - 1}{e^{-\theta} - 1}$

Debye function $D_1 = \frac{1}{\theta} \int_0^\theta \frac{t}{e^t - 1} dt$, $\theta^* = -\log(\theta)$

Tabuľka 3. Hodnoty parametrov testovaných kopula funkcií (C – Clayton, G-H – Gumbel-Hougaard, F – Frank) pre zvolené kombinácie premenných
Table 3. Copula parameters (C – Clayton, G-H – Gumbel-Hougaard, F – Frank), selected combinations of the variables

	$Q_{\max up} - Q_{\text{cor}1dwn}$	$Q_{\max up} - Q_{\text{cor}1tr}$	$Q_{\max dwn} - Q_{\text{cor}2up}$	$Q_{\max dwn} - Q_{\text{cor}2tr}$	$Q_{\max tr} - Q_{\text{cor}3up}$	$Q_{\max tr} - Q_{\text{cor}3dwn}$
Morava – Myjava						
C	1,3	1,22	1,91	0,88	0,83	0,58
G-H	1,65	1,60	1,95	1,44	1,42	1,29
F	4,19	3,98	5,52	3,1	2,88	2,4
Váh – Belá						
C	3,28	0,77	2,42	1,25	0,59	1,66
G-H	2,64	1,39	2,21	1,62	1,29	1,83
F	8,1	2,73	6,72	3,95	2,21	5,21
Nitra – Bebrava						
C	4,06	1,15	5,33	1,89	1,45	2,65
G-H	3,03	1,58	3,60	1,95	1,72	2,32
F	10,41	3,62	13,22	5,61	4,61	7,31
Hron – Slatina						
C	3,75	1,86	2,74	3,15	1,07	2,81
G-H	2,87	1,93	2,37	2,58	1,54	2,41
F	9,65	5,41	7,35	8,05	3,54	7,3

V prvej fáze sme na testovanie vhodnosti kopula funkcie použili porovnanie empirickej pravdepodobnosti výskytu s parametrickým odhadom pravdepodobnosti výskytu podľa kopula funkcií. Pre výpočet empirickej kopula funkcie bola použitá rovnica podľa Gringortena (Rovnica 4) (Gringorten, 1963). Graficky sme porovnali určené empirické pravdepodobnosti so zodpovedajúcimi vypočítanými parametrickými pravdepodobnosťami, ktoré odhalili rozsah, v akom plocha kopuly zodpovedá vybraným radom pozorovaných údajov. Na obr. 4 sú empirické a vypočítané pravdepodobnosti podľa kopula funkcie vynesené proti vzostupným poradiam pozorovania.

$$F(x, y) = P(X \leq x_i, Y \leq y_j) = \frac{\sum_{m=1}^i \sum_{l=1}^j n_{ml} - 0,44}{N + 0,12} \quad (4)$$

kde N je počet premenných, j and i stúpajúca pozícia

hodnoty x_i a y_i , n_{ml} poradie výskytu kombinácií x_i a y_i .

Na posúdenie rozdielov medzi vypočítanou empirickou a parametrickou kopula funkciou boli použité aj nasledovné štatistické indikátory:
Priemerná absolútna chyba odhadu

$$MAE = \frac{1}{N} \sum_{t=1}^N |C_\theta - C_E|. \quad (5)$$

Štandardná chyba odhadu

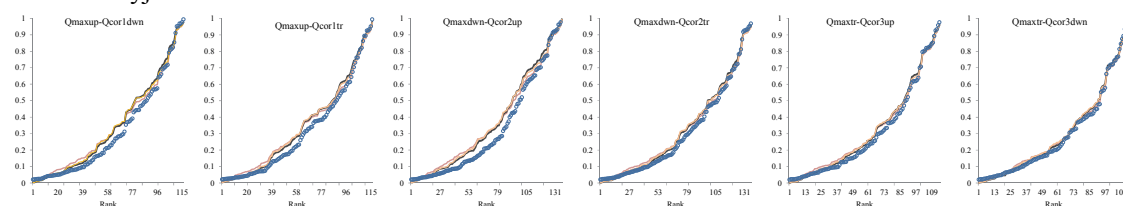
$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (C_\theta - C_E)^2}. \quad (6)$$

Maximálna absolútna chyba odhadu

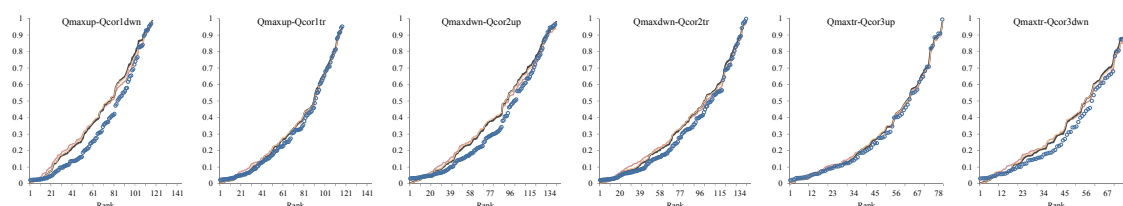
$$ME = \max |C_\theta - C_E|. \quad (7)$$

kde C_θ je parametrická kopula funkcia a C_E je empirická kopula podľa Gringortena.

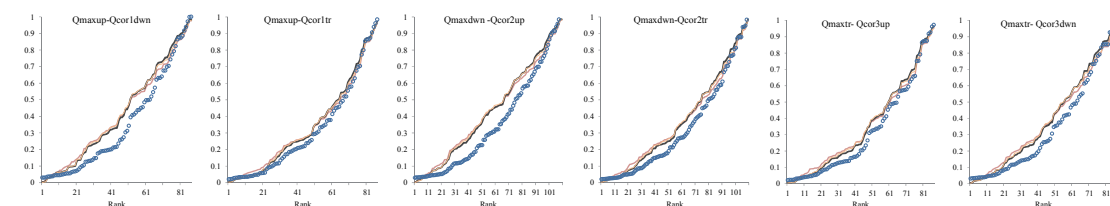
Morava-Myjava



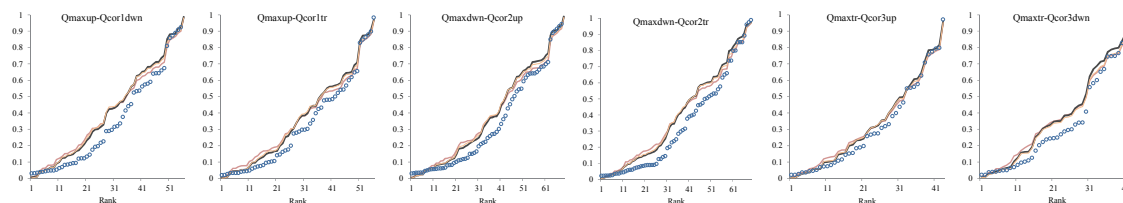
Vah-Bela



Nitra-Bebrava



Hron-Slatina



Obr. 4. Porovnanie empirických a parametrických pravdepodobností podľa kopula funkcií (Clayton – ružová, Gumbel-Hougaard – čierna a Frank – žltá) pre zvolené kombinácie premenných.

Fig. 4. Comparison of the empirical (points) with fitted parametrical copula probabilities (Clayton – pink, Gumbel-Hougaard – black and Frank – yellow).

Z výsledkov porovnania hodnôt chýb odhadu je zrejmé, že všetky testované Archimedovské kopula funkcie dosiahli pomerne rovnaké rozdiely v chybách odhadu. Najmenšie hodnoty chýb odhadu dosiahli pravdepodobnosti vypočítané pomocou Gumbel-Hougaard kopula funkcie. Najlepšia zhoda medzi empirickou a parametrickou kopula funkciou bola dosahovaná pre kombinácie premenných s menšou hodnotou „Rank korelácie“ (korelácia poradia). Následne bol na posúdenie vhodnosti teoretickej kopula funkcie použitý neparametrický Kolmogorov-Smirnov test (KS test).

Na hladine významnosti $\alpha=0,05$ bola pre daný rozsah premenných určená kritická hodnota štatistickej premennej D_k z tabuľky KS testu (napr. Anděl, 2003). Podľa vzťahu (8) (Zhang, 2014) bola vypočítaná hodnota štatistiky D_v , pričom musí platiť, že $D_k \geq D_v$, aby sme nemuseli zamietnuť nulovú hypotézu. Výsledky hodnôt posúdenia vhodnosti kopula funkcie podľa KS testu sú vypísané v Tabuľke 4.

$$D_v = \max_{1 \leq k \leq n} \left\{ \left| C_k - \frac{m_k}{n} \right|, \left| C_k - \frac{m_{k-1}}{n} \right| \right\}. \quad (8)$$

kde C_k je hodnota parametrickej (x_k, y_k) kopula funkcie, m_k je počet párov (x_k, y_k) vyhovujúcich $x \leq x_k$ and $y \leq y_k$.

Podľa KS testu môžeme zamietnuť hypotézu H_0 pre dvojice premenných, ktoré sa prejavujú pomerne vysokou poradovou koreláciou $Q_{maxup} - Q_{cor1dwn}$ (Váh), $Q_{maxup} - Q_{cor1dwn}$ a $Q_{maxdwn} - Q_{cor2up}$ (Nitra) a $Q_{maxdwn} - Q_{cor2tr}$ (Hron – Slatina) pri všetkých troch testovaných kopula funkciách. Keďže hodnota štatistickej premennej D_k narastá s klesajúcim N pre dvojicu premenných $Q_{maxup} - Q_{cor1dwn}$ (Hron) v tomto prípade H_0 nemôžeme zamietnuť. Následne bol pre tieto dvojice použitý ešte doplnujúci χ^2 test na potvrdenie výsledkov KS testu (tabuľka 5). Tento test nepotvrdil zamietnutie H_0 na hladine významnosti $\alpha=0,05$. Najvyššie hodnoty p -value boli dosiahnuté pre Gumbel-Hougaardovú kopula funkciu. Najlepšia zhoda medzi empirickou a parametrickou kopula funkciou ako aj hodnoty p -value boli aj týmito testami dosahované pre kombinácie premenných s menšou hodnotou korelácia poradia. Na základe testov sme ako najvhodnejšiu kopula funkciu na určenie združenej pravdepodobnosti rozdelenia vyhodnotili Gumbel-Hougaardovú funkciu.

Následne bola urobená simulácia 3000 párov podľa Gumbel-Hougaardovej kopula funkcie (obr. 5), a určená združená distribučná funkcia výskytu synchronných prietokov (JPDF-G-H- Joint Probability Distribution Function – Gumbel Hougaard).

Tabuľka 4. Posúdenie vhodnosti zvolených Archimedovských kopula funkcií na výpočet združenej distribučnej funkcie pomocou Kolmogorov-Smirnovho testu
Table 4. Evaluation of Archimedean copula functions for computing the joint distribution function by Kolmogorov-Smirnov test

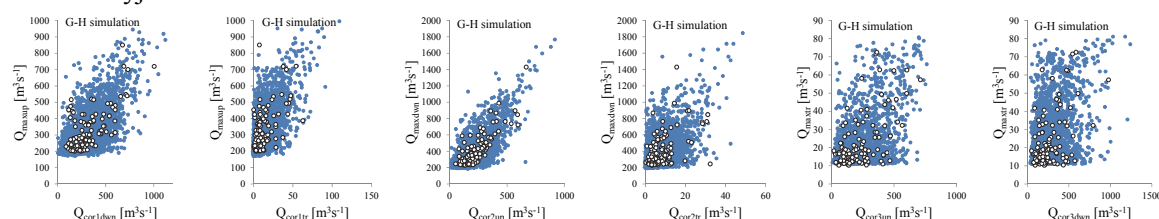
			Clayton			Gumbel-Hougaard		Frank	
	N	τ / ρ^*	D_k	D_v	p -value	D_v	p -value	D_v	p -value
Morava-Myjava									
$Q_{maxup} - Q_{cor1dwn}$	116	0,39 / 0,57	0,126	0,129	0,268	0,095	0,653	0,120	0,364
$Q_{maxup} - Q_{cor1tr}$	116	0,37 / 0,54	0,126	0,121	0,340	0,103	0,541	0,121	0,460
$Q_{maxdwn} - Q_{cor2up}$	137	0,49 / 0,66	0,116	0,146	0,098	0,116	0,290	0,134	0,174
$Q_{maxdwn} - Q_{cor2tr}$	137	0,31 / 0,44	0,116	0,102	0,451	0,083	0,752	0,095	0,547
$Q_{maxtr} - Q_{cor3up}$	115	0,29 / 0,43	0,127	0,122	0,341	0,078	0,859	0,104	0,536
$Q_{maxtr} - Q_{cor3dwn}$	115	0,23 / 0,32	0,127	0,078	0,859	0,078	0,854	0,095	0,647
Váh - Belá									
$Q_{maxup} - Q_{cor1dwn}$	117	0,62 / 0,78	0,126	0,180	0,016	0,174	0,040	0,180	0,028
$Q_{maxup} - Q_{cor1tr}$	117	0,28 / 0,39	0,126	0,111	0,443	0,094	0,658	0,111	0,443
$Q_{maxdwn} - Q_{cor2up}$	140	0,54 / 0,73	0,115	0,143	0,105	0,114	0,302	0,129	0,183
$Q_{maxdwn} - Q_{cor2tr}$	140	0,38 / 0,53	0,115	0,136	0,140	0,107	0,387	0,114	0,302
$Q_{maxtr} - Q_{cor3up}$	80	0,23 / 0,32	0,152	0,088	0,907	0,088	0,907	0,088	0,907
$Q_{maxtr} - Q_{cor3dwn}$	80	0,45 / 0,61	0,152	0,150	0,304	0,113	0,666	0,138	0,409
Nitra - Bebrava									
$Q_{maxup} - Q_{cor1dwn}$	87	0,67 / 0,84	0,146	0,207	0,041	0,201	0,046	0,207	0,041
$Q_{maxup} - Q_{cor1tr}$	87	0,37 / 0,52	0,146	0,150	0,264	0,126	0,464	0,150	0,264
$Q_{maxdwn} - Q_{cor2up}$	110	0,73 / 0,89	0,129	0,182	0,046	0,173	0,670	0,173	0,067
$Q_{maxdwn} - Q_{cor2tr}$	110	0,49 / 0,66	0,129	0,155	0,131	0,127	0,314	0,136	0,240
$Q_{maxtr} - Q_{cor3up}$	87	0,42 / 0,58	0,146	0,184	0,093	0,138	0,335	0,149	0,264
$Q_{maxtr} - Q_{cor3dwn}$	87	0,57 / 0,76	0,146	0,172	0,135	0,143	0,264	0,161	0,191
Hron - Slatina									
$Q_{maxup} - Q_{cor1dwn}$	56	0,65 / 0,84	0,181	0,179	0,303	0,161	0,430	0,163	0,430
$Q_{maxup} - Q_{cor1tr}$	56	0,48 / 0,67	0,181	0,167	0,387	0,132	0,693	0,139	0,619
$Q_{maxdwn} - Q_{cor2up}$	69	0,58 / 0,78	0,163	0,150	0,317	0,130	0,570	0,128	0,568
$Q_{maxdwn} - Q_{cor2tr}$	69	0,61 / 0,82	0,163	0,217	0,066	0,188	0,154	0,203	0,102
$Q_{maxtr} - Q_{cor3up}$	43	0,35 / 0,48	0,203	0,140	0,865	0,116	0,918	0,117	0,917
$Q_{maxtr} - Q_{cor3dwn}$	43	0,58 / 0,78	0,203	0,209	0,269	0,203	0,270	0,209	0,269

* τ / ρ sú hodnoty poradovej korelácie

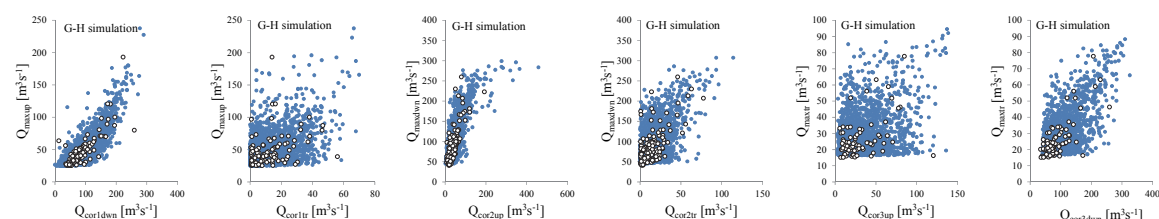
Tabuľka 5. Posúdenie vhodnosti zvolených Archimedovských kopula funkcií na výpočet združenej distribučnej funkcie pomocou χ^2 test
Table 5. Evaluation of Archimedean copula functions for computing the joint distribution function test χ^2 test

	df	Clayton			Gumbel -Hougaard		Frank	
		$\text{crit } \chi^2_{(0,05)}$	χ^2	$p\text{-value}$	χ^2	$p\text{-value}$	χ^2	$p\text{-value}$
$Q_{\maxup} - Q_{\text{cor1dwn}} (Váh)$	6	12,59	18,75	0,005	10,17	0,118	11,92	0,06
$Q_{\maxup} - Q_{\text{cor1dwn}} (Nitra)$	7	14,07	23,89	0,002	13,86	0,054	21,56	0,003
$Q_{\maxdwn} - Q_{\text{cor2up}} (Nitra)$	7	14,07	18,58	0,010	13,47	0,062	16,66	0,020
$Q_{\maxdwn} - Q_{\text{cor2up}} (Hron - Slatina)$	6	12,59	10,86	0,090	4,84	0,560	5,99	0,174

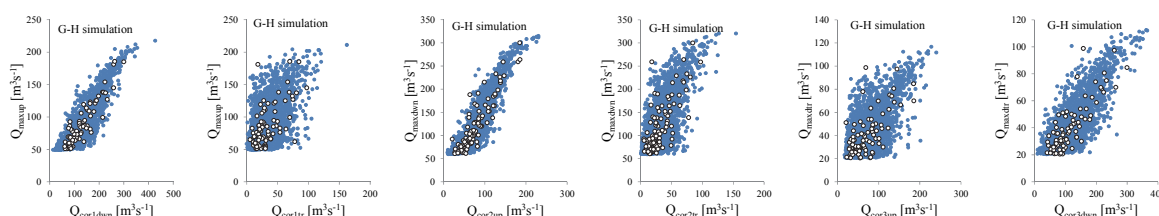
Morava-Myjava



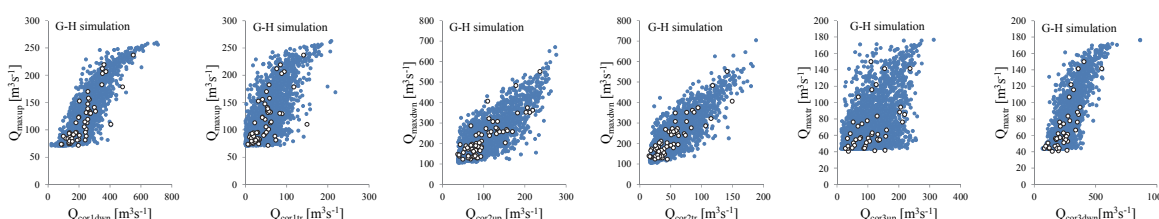
Váh-Bela



Nitra-Bebrava



Hron-Slatina



Obr. 5. Porovnanie simulácie 3000 párov generovaných pomocou Gumbel-Hougaard kopula funkcie pri nezamietnutej hypotéze H_0 .

Fig. 5. Scatter plots of 3000 data pairs generated from Gumbel-Hougaard copula and measured data of the synchronous variables for selected combinations (cannot reject hypothesis H_0).

Tabuľka 6 prezentuje hodnoty JDF-G-H a združenú pravdepodobnosť výskytu maximálnych prietokov nameraných na hlavnom toku a jeho prítoku v daných staniciach $P \cap (x, y)$ rovnica (9).

$$P = 1 - F_{(cor)} - F_{(max)} + C_{(F(max), F(cor))} \quad (9)$$

kde, $F_{(cor)}$ a $F_{(max)}$ sú jednorozmerné okrajové distribučné funkcie a $C_{(F(max), F(cor))}$ je združená distribučná funkcia podľa zvolenej kopula funkcie.

Vo všeobecnosti pravdepodobnosti, že obe premenné prekročia maximálne namerané hodnoty súčasne v tejto štúdii boli pomerne nízke. Takáto najvyššia pravdepodobnosť ($P = 0,89 \%$) bola dosiahnutá pre tok Hron a jeho prítok Slatina pre dvojicu maximálnych premenných ($Q_{maxtr}-Q_{max3corup}$).

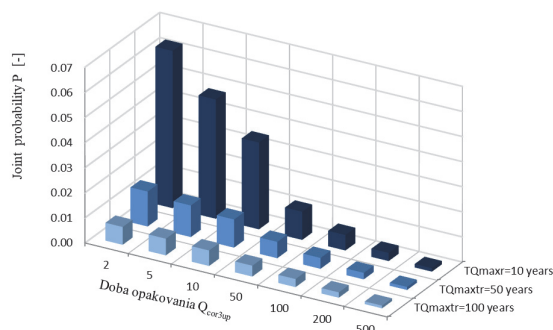
Obr. 6 znázorňuje príklad združenej pravdepodobnosti prekročenia $P \cap (x, y)$, že obe premenné ($Q_{maxtr}-Q_{cor3up}$) na rieke Hron a jeho prítoku Slatina prekročia zvolenú hodnotu súčasne.

Tabuľka 6. Hodnoty JDF-G-H pri maximálnych prietokoch nameraných na hlavnom toku a jeho prítoku v daných staniciach

Table 6. Values of the JDF-G-H for maximal synchronous discharges for selected rivers

	Morava-Myjava	Váh-Bela	Nitra-Bebrava	Hron-Slatina
$Q_{maxup} [m^3 s^{-1}]$	850	193	185,6	236,7
$Q_{max1cortr} [m^3 s^{-1}]$	62,8	56,1	97,4	149,8
F_{maxup}	0,996	0,998	0,99	0,988
$F_{max1cortr}$	0,993	0,996	0,991	0,991
Kopula parameter θ	1,6	1,39	1,58	1,93
JDF-G-H	0,991	0,995	0,985	0,985
P [%]	0,226	0,096	0,430	0,585
$Q_{maxdwn} [m^3 s^{-1}]$	1429	259,9	300	551,1
$Q_{max2cortr} [m^3 s^{-1}]$	32,4	77,7	97,4	149,8
F_{maxdwn}	0,998	0,997	0,996	0,994
$F_{max2cortr}$	0,994	0,997	0,993	0,991
Kopula parameter θ	1,44	1,62	1,95	2,58
JDF-G-H	0,994	0,995	0,992	0,990
P [%]	0,111	0,140	0,289	0,490
$Q_{maxtr} [m^3 s^{-1}]$	72,46	77,7	98,7	149,8
$Q_{max3corup} [m^3 s^{-1}]$	712	120,6	185,6	236,7
F_{maxtr}	0,994	0,995	0,991	0,977
$F_{max3corup}$	0,997	0,996	0,992	0,982
Kopula parameter θ	1,42	1,29	1,72	1,54
JDF-G-H	0,993	0,992	0,987	0,968
P [%]	0,141	0,130	0,429	0,890
$Q_{max3cordwn} [m^3 s^{-1}]$	978,9	259,9	300	551,1
$F_{max3cordwn}$	0,998	0,993	0,994	0,992
Kopula parameter θ	1,29	1,83	2,32	2,41
JDF-G-H	0,993	0,991	0,989	0,976
P [%]	0,091	0,315	0,489	0,729

F – jednorozmerná distribučná funkcia rozdelenia pravdepodobnosti premennej



Obr. 6. Príklad združenej pravdepodobnosti ako funkcie ($Q_{maxtr}-Q_{cor3up}$) na rieke Hron a jeho prítoku Slatina.

Fig. 6. The example of the joint probability as a function of the ($Q_{maxtr}-Q_{cor3up}$) on Hron river and its tributary-Slatina.

Záver

Predložený príspevok prezentuje dvojrozmernú frekvenčnú analýzu synchrónne sa vyskytujúcich maximálnych prietokov na hlavnom toku a jeho prítoku. Analyzované boli synchrónne priemerné denné prietoky na hlavných slovenských tokoch a ich vybraných prítokoch v povodí rieky Dunaj. V prvej časti príspevku bola opísaná príprava vstupných údajov a korelačná analýza vzájomných synchrónne sa vyskytujúcich maximálnych prietokov. V ďalšej časti je prezentovaný výber marginálnych distribučných funkcií rozdelenia pravdepodobnosti a výber vhodnej kopula funkcie pre dvojrozmernú združenú štatistickú analýzu synchrónnych prietokov. Parametrické distribúcie JohnsonSB a Gamma sa ukázali ako najlepšie pre väčšinu vybraných premenných odvodených metódou POT.

V predloženom príspevku sme analyzovali najmä vhodnosť najčastejšie používanej triedy kopula funkcií v hydrológii – Archimedovské kopula funkcie (Clayton, Gumbel-Hougaard a Frank). Najmenšie hodnoty chýb odhadu dosiahli pravdepodobnosti vypočítané pomocou Gumbel-Hougaard kopula funkcie, a preto bola táto funkcia vybraná ako najvhodnejšia na ilustráciu združeného rozdelenia pravdepodobnosti synchrónnych premenných.

Podľa zvoleného neparametrického Kolmogorov-Smirnovho testu bola hypotéza H_0 (dvojice náhodných premenných ukazujú dobrú zhodu medzi empirickou a teoretickou kopula funkciou) zamietnutá pre všetky testované kopula funkcie pri pároch, ktoré dosahovali pomerne vysoké hodnoty poradových závislostí Kendall τ a Spearman ρ . Najlepšia zhoda medzi empirickou a parametrickou kopula funkciou bola dosahovaná pre kombinácie premenných s menšou hodnotou Kendallovho koeficienta korelácie. Podľa KS testu sa taktiež pre ostatné synchrónne páry ukázala ako najvhodnejšia z testovaných Archimedovských kopula funkcií Gumbel-Hougaardová funkcia na určenie združenej distribučnej funkcie. V ďalšej časti bol pre dvojice premenných na potvrdenie zamietnutia hypotézy H_0 použitý ešte doplňujúci neparametrický χ^2 test. Tento test nepotvrdil zamietnutie H_0 na hladine významnosti $\alpha=0,05$. Najvyššie hodnoty p -value boli dosiahnuté pre Gumbel-Hougaardovú kopula funkciu, ktorá bola využitá na určenie združenej pravdepodobnosti výskytu synchrónnych prietokov. Napriek tomu, že pravdepodobnosť súčasného výskytu analyzovaných maximálnych synchrónnych prietokov sú pomerne nízke nemožno vylúčiť, že takto situácia by nemohla nastať. Napr. Tadić a kol. (2016) analyzovali združenú pravdepodobnosť výskytu povodní na Dunaji a Dráve pri meste Osijek. Ich výsledky ukázali, že pravdepodobnosť takejto situácie je nízka (0,79%), ale nie nepravdepodobná. Poukázali na to, že takáto situácia nastala v roku 1966 a bola to jedna z najväčších povodní. Aj preto výsledky získané dvojrozmernou analýzou premenných, ktoré

charakterizujú hydrologický režim, môžu prispieť k spoľahlivejšiemu hodnoteniu povodňových rizík.

Podakovanie

Táto práca bola podporovaná projektom VEGA 2/0009/15 a realizáciou projektu ITMS 26240120004 Centrum excelentnosti integrovanej protipovodňovej ochrany územia, na základe podpory operačného programu Výskum a vývoj financovaného z Európskeho fondu regionálneho rozvoja.

Literatúra

- Anděl, J. (2003): Statistické metody. 2. vyd. Praha: MATFYZPRESS, 2003
- Bayliss, A. C., 1999: Deriving flood peak data. Flood estimation: Handbook. Vol.3, p.273–283.
- Cunnane, C. (1973): A particular comparison of annual maximum and partial duration series methods of flood frequency prediction. Journal of Hydrol., 18, 257–271.
- Cunnane, C. (1978): Methods and merits of regional flood frequency analysis. J. Hydrol., 100, 269–290.
- De Michele, C., Salvadori, G., Canossi, M., Petaccia, A., Rosso, R. (2005): Bivariate statistical approach to check adequacy of dam spillway. J. Hydrol. Eng., 10(1), 50–57.
- Gaál, L., Szolgay, J., Bacigál, T., Kohnová, S. (2010): Copula-based assessment of the relationship between flood peaks and flood volumes using information on historical floods by Bayesian Monte Carlo Markov Chain simulations. Geophysical Research Abstracts, Vol. 12, EGU2010-13534-3.
- Genest, C., Favre, A. C. (2007): Everything You Always Wanted to Know about Copula Modeling but Were Afraid to Ask. J. Hydrol. Eng., 12, 347–368.
- Gringorten, I. I. (1963): A plotting rule for extreme probability paper. J. Geophys. Res., 68(3), 813–814.
- Chen, X., Fan, Y. (2005): Pseudo-likelihood ratio tests for semi-parametric multivariate copula model selection. Can. J. Stat., 33(3), 389–414.
- Chen, L., Singh, V., Shenglian, G., Hao, Z., and Li, T. (2012): Flood Coincidence Risk Analysis Using Multivariate Copula Functions. J. Hydrol. Eng., 10.1061/(ASCE)HE.1943-5584.0000504, 742–755.
- Chowdhary, H., Escobar, L., A., Singh, V. P. (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.
- Matúš, R. (2009): The modeling of hydrological joint events on the Morava river using aggregation operators., Slovak university of Technology, Slovak Journal of Civil Engineering Vol. 3, p. 9–15.
- Nelsen, R. B. (2006): An introduction to Copula. The 2nd edition. Springer, New York.

- Prohaska, S., Isailović, D., Srna, P., Marčetić, I. (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.
- Requena, A., Mediero, L., Garrote, L. (2012): A Monte Carlo procedure based on the use of copulas for risk assessment of dam overtopping. 3rd STAHY international Workshop on Stat. methods for Hydr. and Water Res. Manag., Tunis, 1–11.
- Shiau, J. T., Wang, H. Y., Tsai, Ch. T. (2010): Copula-based depth-duration-frequency analysis of typhoons in Taiwan. *Hydrol. Res.*, 41(5), 414–422.
- Szolgay, J., Gaál, L., Bacigál, T., Kohnová, S., Hlavčová, K., Výleta, R., Parajka, P., Blöschl, G. (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.
- Tadić, L., Bonacci, O. and Dadić, T. (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>
- Výzkumný ústav vodohospodářský T. G. Masaryka (2002): Vyhodnocení katastrofální povodně v srpnu 2002 a návrh úpravy systému prevence před povodněmi., Výsledná zpráva o projektu 2002, Praha, 174 s.
- Zhang, L., Singh, V.P. (2006): Bivariate flood frequency analysis using the copula method. *J. Hydrol. Eng. ASCE*, 11(2), 150–164.
- Zhang, J., Ding, Z., You J. (2014): The joint probability distribution of runoff and sediment and its change characteristics with multi-time scales. *J. Hydrol. Hydromech.*, Vol. 62 (3), 218 – 225, doi: 10.2478/johh-2014-0024.

TESTING THE TWO-DIMENSIONAL ARCHIMEDEAN COPULA FUNCTIONS IN COINCIDENCE OF FLOOD DISCHARGES OF THE MAIN RIVER AND TRIBUTARY

The paper presents the two-dimensional frequency analysis of the synchronous maximum discharges at the main river and its tributary. The daily discharges on the main Slovak rivers and their selected tributaries of the Danube river basin were analysed. The first part of the paper described the preparation of the input data and correlation analysis of synchronic maximum flows. In the next part, the selection of marginal probability distribution functions and selection of the appropriate copula for the two-dimensional joint statistical analysis of synchronous discharges were done. The parametrical JohnsonSB and Gamma distributions showed the best agreement for variables obtained from POT method.

In the presented paper, we analysed the suitability of the most commonly used class of copula functions in hydrology - Archimedean copula functions (Clayton, Gumbel-Hougaard and Frank). At the first some errors of the estimation for an evaluation of the difference between the parametrical (Clayton, Gumbel-Hougaard and Frank) and empirical copula function was used. Based on the statistical criteria and visual comparison, the Gumbel-Hougaard copula function was selected as the most suitable to illustrate the joint probability distribution of the synchronous variables. The best match between empirical and parametrical copula function was achieved for combinations of variables with a lower Kendall correlation coefficient. Next, the non-parametric goodness-of-fit test Kolmogorov-

Smirnov for evaluation of the selected copula was used. According to the KS test the null hypothesis H_0 (variables come from same distribution) was rejected for all pairs with relative high values of the Kendall τ a Spearman ρ . The best agreement between empirical and parametric copula function was achieved for combination of variables with a lower Kendall τ coefficient and for Gumbel-Hougaard copula function. For confirmation of the rejection of the null hypothesis H_0 , another nonparametric χ^2 test was used. This test did not confirm rejection of H_0 at the significance level $\alpha = 0.05$ and achieved the highest values of *p-value* for the Gumbel-Hougaard copula. The Gumbel-Hougaard copula was used to determine the joint probability of occurrence of the synchronic variables. Although the occurrence probability of analysed maximum synchronous discharges was relatively low it cannot be excluded that such a situation would not happen. E.g. 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 not unlikely. They reminded that such a situation occurred in 1966 and it was one of the biggest floods. Therefore, the results obtained by the two-dimensional analysis of the variables which characterize the hydrological regime can contribute to a more reliable assessment of the flood risks.

Ing. Veronika Bačová Mitková, PhD.

Ing. Dana Halmová, PhD

Ústav hydrológie SAV

Dúbravská cesta 9

84104 Bratislava

Slovenská republika

E-mail: mitkova@uh.savba.sk