

UDK/UDC: 519.21:556.53:627.53(497.4)

Prejeto/Received: 6. 11. 2012

Izvirni znanstveni članek – *Original scientific paper*

Sprejeto/Accepted: 21. 4. 2013

VPLIV IZBIRE METODE NA REZULTATE VERJETNOSTNIH ANALIZ KONIC, VOLUMNOV IN TRAJANJ VISOKOVODNIH VALOV SAVE V LITIJ

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 LITIJ

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

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

Izveleček

Kompleksni hidrološki pojavi, kot so poplave, se vedno pojavljajo kot posledica več med seboj odvisnih slučajnih spremenljivk (npr. konica, volumen, trajanje poplavnega vala). Z verjetnostno analizo le ene spremenljivke lahko zato le približno ocenimo verjetnost pojava določenega dogodka. Za popolno razumevanje vseh treh spremenljivk in njihovo medsebojno povezanost je nujen multivariatni statistični pristop. Predpogoj za to je obsežna analiza posameznih spremenljivk. V raziskavi smo obravnavali 58 visokovodnih valov z vodomerne postaje Litija na reki Savi. Verjetnostno analizo smo izvedli za vse tri spremenljivke, ki določajo visokovodni val. S statističnimi in grafičnimi testi smo primerjali v praksi najpogosteje uporabljene porazdelitvene funkcije in načine ocenjevanja parametrov porazdelitev. Ugotovili smo, da je konice in trajanja valov najbolje opisala logaritemsko Pearsonova III porazdelitev, volumne pa Pearsonova III porazdelitev. Pri večini porazdelitev je najboljše ocene parametrov izkazala metoda momentov L, ki se v slovenski hidrološki praksi do sedaj še ne uporablja.

Ključne besede: verjetnostna analiza, metoda momentov, metoda momentov L, metoda največjega verjetja, robne porazdelitve, statistični testi.

Abstract

Complex hydrological events such as floods always appear as a consequence of several correlated random variables (peak, volume, duration). Therefore single-variable frequency analysis can only provide limited assessment of these events. To fully understand all three variables and their relationship, a multivariate statistical approach is necessary. A precondition for such an approach is a complete analysis of all individual variables. 58 years of data from gauging station Litija on the Sava River were analyzed. Peaks, volumes and durations of flood waves were considered in the flood frequency analysis. Statistical and graphical tests were used to assess performance of the most commonly used distribution functions and parameter estimation techniques. Log-Pearson III distribution gave the best results in case of peaks and durations of flood waves and Pearson III distribution in case of volumes. Method of L-moments, which has not yet been used in Slovenian hydrological practice, gave the best estimation of parameters for most distributions.

* 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](#).

Keywords: flood frequency analysis, method of moments, method of L-moments, maximum likelihood method, marginal distributions, statistical tests.

1. Uvod

Upravljanje z vodami, načrtovanje, preventivno obvladovanje tveganja zaradi poplav, določanje poplavnih območij in projektnih pretokov, dimenzioniranje objektov na vodi ipd. – vse to zahteva temeljito poznavanje značilnosti visokovodnih dogodkov (Brilly in Polič, 2005; Đurović in Mikoš, 2004; Kobold et al., 2005; Mikoš et al., 2004; Šraj et al., 2010). V hidrološki praksi se za opis verjetnosti pojava visoke vode oziroma poplave uporablja pojem povratna doba. Povratna doba nekega dogodka je navadno definirana kot povprečen čas, ki preteče med pojavom dveh enakovrednih in zaporednih dogodkov (Salvadori in De Michele, 2007). Povratna doba je enostaven način določanja tveganja oziroma verjetnosti pojava nekega dogodka. Trenutno se povratna doba ekstremnih dogodkov v Sloveniji računa na podlagi verjetnostne analize maksimalnih visokovodnih konic, torej na podlagi le ene spremenljivke. To lahko pripelje do podcenjevanja tveganja pojava določenega dogodka (Favre et al., 2004). Večina hidroloških pojavov je v resnici odvisna od hkratnega vpliva več naključnih spremenljivk (npr. konica, volumen, trajanje itd.), ki so v večini primerov medsebojno odvisne. Dogodek s 100-letno povratno dobo konice vala lahko na primer povzroči manj škode kot dogodek z 10-letno povratno dobo konice in volumna vala (Renard in Lang, 2007). Kompleksni hidrološki pojavi kot so poplave, se vedno pojavljajo kot posledica več med seboj odvisnih slučajnih spremenljivk. Torej lahko z verjetnostno analizo le ene spremenljivke le približno ocenimo možnost pojava določenega dogodka. Za popolno razumevanje vseh treh spremenljivk in njihove medsebojne povezanosti je nujen multivariatni statistični pristop. V zadnjih nekaj letih se je pojavilo kar nekaj multivariatnih pristopov v hidroloških in okoljskih študijah po svetu. Večina teh študij temelji na vsaj eni od naslednjih predpostavk: spremenljivke so medsebojno neodvisne, vse spremenljivke sledijo

normalni porazdelitvi ali (in) vse spremenljivke pripadajo istemu tipu porazdelitvene funkcije (Zhang in Singh, 2007). V resnici so v naravi spremenljivke, ki opisujejo poplavni val, med seboj odvisne, v splošnem ne sledijo normalni porazdelitvi in vseh spremenljivk ne moremo opisati z istim tipom robne porazdelitvene funkcije (Zhang in Singh, 2007). Glede na vsa ta spoznanja so znanstveniki po svetu v zadnjem času za verjetnostno analizo poplav poskušali uporabiti funkcijo kopula (Favre et al., 2004; Grimaldi in Serinaldi, 2006; Renard in Lang, 2007; Salvadori in De Michele, 2004; Salvadori in De Michele, 2007). Funkcija kopula je bila do sedaj uspešno uporabljena predvsem v ekonomiji na področju financ. S funkcijo kopula lahko modeliramo odvisnost več spremenljivk, neodvisno od tipa porazdelitve, ki mu sledijo posamezne spremenljivke. Poleg tega je določitev robnih porazdelitvenih funkcij ločena od izbire najustreznejše funkcije kopula. Razvoj in uporaba novega pristopa je v polnem zamahu. Modeli z uporabo funkcije kopula se pravzaprav v svetovni hidrološki praksi šele uveljavljajo. Cela vrsta vprašanj pri izbiri vzorcev, metodologiji in analizi rezultatov je še odprta. Seveda pa tako kompleksna, multivariatna verjetnostna analiza zahteva predhodno natančno analizo lastnosti in odvisnosti posameznih spremenljivk ter verjetnostno analizo vsake posamezne spremenljivke, kar je prikazano v tem članku.

Cilji raziskave so bili: (i) določiti najprimernejšo natančnost podatkov za multivariatne verjetnostne analize, (ii) narediti pregled v svetovni hidrološki praksi najpogosteje uporabljenih porazdelitev in enačb za oceno parametrov porazdelitev, (iii) pripraviti pregled testov, ki se največkrat uporabljajo pri izbiri najprimernejše porazdelitve, (iv) posameznim spremenljivkam visokovodnih valov (konica, volumen, trajanje) izbrane vodomerne postaje poiskati najboljše prilegajočo se robno porazdelitveno funkcijo, ne nujno iz iste družine porazdelitev, (v) ugotoviti vpliv izbire

porazdelitve in metode ocenjevanja parametrov posamezne porazdelitve na rezultate, (vi) rezultate analiz in ugotovitve primerjati z rezultati drugih avtorjev.

2. Podatki

V raziskovalni nalogi smo se odločili za obdelavo podatkov s hidrološke postaje Litija na reki Savi (slika 1), saj je le-ta najstarejša hidrološka postaja v Sloveniji. Na postaji spremljajo vodostaje že od leta 1893. Postaja je bila leta 1953 zaradi zasipavanja dna prestavljena 500 m gorvodno in tam stoji še danes (Ulaga, 2011). Ob prestavitvi so jo opremili z limnigrafom (Seba Omega), kar pomeni, da imamo od takrat dalje podatke o celotnih hidrogramih in ne le o konicah. Zato so vse analize predstavljene raziskave narejene za obdobje 1953-2010.

Pri analizah smo uporabili dnevne vrednosti pretokov z vključenimi izmerjenimi maksimalnimi konicami. Za tako izbiro podatkov smo se odločili po natančni analizi in primerjavi spremenljivk naključno izbranih valov. V ta namen smo primerjali 13 naključno izbranih visokovodnih valov, ki smo jim bazni odtok izločili z uporabo preproste tro-točkovne grafične metode. Primerjavo smo izvedli med urnimi vrednosti merjenih pretokov in dnevnimi vrednostmi z upoštevanje največje konice. V obeh primerih tako dobimo enake konice in trajanja valov, razlika v volumnih valov pa je v povprečju znašala 4,3 %. Do podobnih zaključkov je prišla tudi Pugelj (2012), ki je analizirala podatke z vodomerne postaje Šentjakob na Savi. Tudi pregled tuje literature je pokazal, da je uporaba dnevnih vrednostih v podobnih tujih študijah zelo pogosta (Karmakar in Simonovic, 2008; Zhang in Singh, 2006; Zhang in Singh, 2007). Zato smo nadaljnje analize izvedli na dnevnih vrednostih z upoštevanjem absolutnih konic.

Za Savo pri Litiji je značilen alpski dežno-snežni režim (Frantar in Hrvatina, 2008). Značilnost takega režima sta dva dokaj izenačena viška, eden spomladi in eden jeseni. V primeru Litije je jesenski višek nekoliko bolj izrazit.

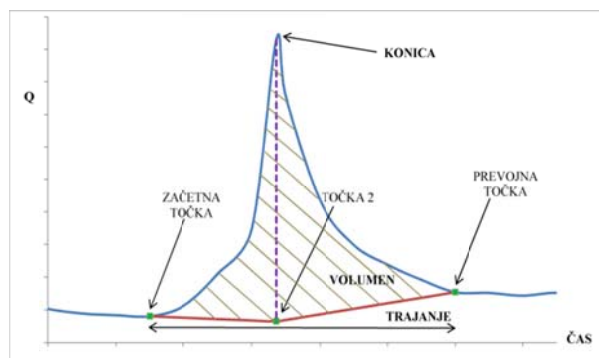


Slika 1: Hidrološka postaja Litija na reki Savi (foto: Nejc Bezak).

Figure 1: Hydrologic station Litija on the Sava River (photo: Nejc Bezak).

3. Metode

Vzorec smo določili po metodi letnih maksimumov (Bezak, 2012). Za vsako leto obravnavanega obdobja (1953-2010) smo določili visokovodni val z največjo konico. Vsem 58 tako izbranim visokovodnim valom smo nato s tro-točkovno metodo izločili bazni del odtoka (slika 2) in nato določili vrednosti vseh treh spremenljivk vala: volumna, konice in trajanja.



Slika 2: Metoda izločanja baznega odtoka.

Figure 2: Base flow separation method.

3.1 Verjetnostne porazdelitve

Pri verjetnostni analizi letnih maksimumov vseh treh spremenljivk smo uporabili teoretične porazdelitve, ki se najpogosteje uporabljajo v svetovni hidrološki praksi: logaritemsko normalno porazdelitev, Pearsonovo III porazdelitev, logaritemsko Pearsonovo III porazdelitev,

Gumbelovo porazdelitev, normalno porazdelitev, generalizirano porazdelitev ekstremnih vrednosti (GEV) in generalizirano logistično porazdelitev (GL). Merjene podatke smo umestili v verjetnostne mreže s pomočjo empirične Weibullove porazdelitve (Haan, 2002), ki se najpogosteje

uporablja v te namene. Za vsako porazdelitev smo izračunali tudi intervale zaupanja. Poleg tega smo oceno parametrov posameznih porazdelitev izvedli po treh različnih metodah: metodi momentov, metodi momentov L in metodi največjega verjetja.

Preglednica 1: Porazdelitvene funkcije in enačbe za oceno parametrov porazdelitev.

Table 1: Distribution functions and equations for distribution parameters estimation.

Porazdelitev/Distribution	Porazdelitvena funkcija; enačbe za oceno parametrov z metodo momentov; metodo momentov L in logaritmirana funkcija verjetja/ Distribution function; equations for method of moments; L-moments and logarithm of likelihood function
Normalna	$F_X(x) = \frac{1}{\sigma_X \sqrt{2\pi}} \int_{-\infty}^x e^{-(x-\mu_X)^2/2\sigma_X^2} dx; \mu_X = \bar{x}, \sigma_X = S_X; \mu_X = l_1, \sigma_X = \sqrt{\pi} l_2;$ $\ln L(\mu_X, \sigma_X) = -n \ln(2\pi) - n \ln \sigma_X - \frac{1}{2\sigma_X^2} \sum_{i=1}^n (x_i - \mu_X)^2.$
Logaritmsko normalna	$F_X(x) = \int_0^x \frac{1}{x\sigma_Y \sqrt{2\pi}} e^{-(\ln x - \mu_Y)^2/2\sigma_Y^2} dx, \sigma_Y = \left[\ln \left(1 + \frac{\sigma_X^2}{\mu_X^2} \right) \right]^{1/2}, \mu_Y = \ln \mu_X - \frac{\sigma_Y^2}{2}, \mu_Y = \bar{y},$ $\sigma_Y = S_Y; \mu_Y = l_1, \sigma_Y = \sqrt{\pi} l_2;$ $\ln L(\mu_Y, \sigma_Y) = -n \ln(2\pi) - n \ln \sigma_Y - \sum_{i=1}^n \ln x_i - \frac{1}{2\sigma_Y^2} \sum_{i=1}^n (\ln x_i - \mu_Y)^2.$
Pearsonova III	$F_X(x) = \int_c^x \frac{1}{\beta \Gamma(\alpha)} \left(\frac{x-c}{\beta} \right)^{\alpha-1} e^{-(x-c)/\beta} dx; \alpha = \frac{4}{c_{sx}^2}, \beta = \frac{\sigma_X c_{sx}}{2}, c = \mu_X - \alpha\beta;$ <p>Za: $0 < t_3 < 1/3$: $z = 3\pi t_3^2, \alpha = \frac{1+0,2906z}{z+0,1882z^2+0,0442z^3},$</p> <p>Za: $1/3 < t_3 < 1$: $z = 1 - t_3, \alpha = \frac{0,36067z-0,59567z^2+0,25361z^3}{1-2,78861z+2,56096z^2-0,77045z^3},$</p> <p>Za vse t_3: $\beta = \text{sign}(t_3) \sqrt{\pi} l_2 \frac{\Gamma(\alpha)}{\Gamma(\alpha+0,5)}, c = l_1 - \alpha\beta;$</p> $\ln L(\alpha, \beta, c) = -n \ln(\beta) - n \ln(\Gamma(\alpha)) - n(\alpha - 1) \ln(\beta) + (\alpha - 1) \sum_{i=1}^n \ln(x_i - c) - \frac{1}{\beta} \sum_{i=1}^n (x_i - c).$
Logaritmsko Pearsonova III	$F_Y(y) = \int_0^y \frac{1}{\Gamma(\alpha)} \left(\frac{y-c}{\beta} \right)^{\alpha-1} e^{-(y-c)/\beta} dy, y = \log x; \alpha = \frac{4}{c_{sy}^2}, \beta = \frac{\sigma_Y c_{sy}}{2}, c = \mu_Y - \alpha\beta; \text{Enake}$ <p>enačbe kot pri Pearsonovi III porazdelitvi;</p> $\ln L(\alpha, \beta, c) = -n\alpha \ln(\beta) - n \ln(\Gamma(\alpha)) - \sum_{i=1}^n y_i + (\alpha - 1) \sum_{i=1}^n \ln(y_i - c) - \frac{1}{\beta} \sum_{i=1}^n (y_i - c).$
Gumbelova	$F_X(x) = e^{-e^{-(x-u)/\alpha}}, \alpha = \frac{\sqrt{6}\sigma_X}{\pi}, u = \mu_X - 0,5772\alpha; \alpha = \frac{l_2}{\ln 2}, u = l_1 - 0,5772157\alpha;$ $\ln L(u, \alpha) = -n \ln(\alpha) - \frac{1}{\alpha} \sum_{i=1}^n (x_i - u) - \sum_{i=1}^n e^{-(x_i-u)/\alpha}.$
GEV	$F_X(x) = \exp \left(- \left\{ 1 - k \left[\frac{x-\xi}{\alpha} \right]^{1/k} \right\} \right); c_s = \frac{k}{ k } \frac{-\Gamma(1+3k)+3\Gamma(1+k)\Gamma(1+2k)-2\Gamma^3(1+k)}{[\Gamma(1+2k)-\Gamma^2(1+k)]^{3/2}}, \alpha =$ $\sqrt{\frac{k^2 \sigma_X^2}{\Gamma(1+2k)-\Gamma^2(1+k)}}, \xi = \mu_X - \frac{\alpha}{k} [1 - \Gamma(1+k)]; c = \frac{2}{3+t_3} - \frac{\ln 2}{\ln 3}, k = 7,8590c + 2,9554c^2,$ $\alpha = \frac{k l_2}{\Gamma(1+k)(1-2^{-k})}, \xi = l_1 + \frac{\alpha(\Gamma(1+k)-1)}{k}; \ln L(k, \alpha, \xi) = -n \ln(\alpha) - (1-k) \sum_{i=1}^n y_i - \sum_{i=1}^n e^{-y_i}, y = -\frac{1}{k} \ln \left(1 - k \left(\frac{x_i-\xi}{\alpha} \right) \right).$
GL	$F_X(x) = \left(1 + \left\{ 1 - \frac{k}{\alpha} [x - \xi]^{1/k} \right\} \right)^{-1}; k = \frac{2}{3\pi} \tan^{-1}(-0,59484c_s),$ $\alpha = \text{sign}(k) \frac{k\sigma_X}{[\Gamma(1+2k)\Gamma(1-2k)-\Gamma^2(1+k)\Gamma^2(1-k)]^{1/2}},$ $\xi = \mu_X - \frac{\alpha}{k} [1 - \Gamma(1+k)\Gamma(1-k)]; k = -t_3, \alpha = \frac{l_2}{\Gamma(1+k)\Gamma(1-k)}, \xi = l_1 + \frac{l_2-\alpha}{k};$ $\ln L(k, \alpha, \xi) = -n \ln(\alpha) + \left(\frac{1}{k} - 1 \right) \sum_{i=1}^n \ln \left(1 - k \frac{x_i-\xi}{\alpha} \right) - 2 \sum_{i=1}^n \ln \left(1 + \left(1 - k \frac{x_i-\xi}{\alpha} \right)^{1/k} \right).$

Vse uporabljene porazdelitvene funkcije in enačbe za oceno parametrov porazdelitev so prikazane v preglednici 1 (Bezjak, 2012). Z uporabo enačb iz preglednice 1 lahko določimo parametre normalne porazdelitve (μ_X, σ_X), logaritemsko normalne porazdelitve (μ_Y, σ_Y), Pearsonove III porazdelitve (α, β, c), logaritemsko Pearsonove III porazdelitve (α, β, c), Gumbelove porazdelitve (u, α), generalizirane porazdelitve ekstremnih vrednosti (k, α, ξ) in generalizirane logistične porazdelitve (k, α, ξ). Ostale neznanke v preglednici 1 označujejo momente in momente L, ki jih izračunamo iz vzorca.

Trenutno se v slovenski hidrološki praksi uporablja le prvih pet omenjenih teoretičnih porazdelitev in ocenjevanje parametrov po metodi navadnih momentov, ki je najpreprostejša. Velikokrat pa se verjetnostno analizo izvede celo le z eno porazdelitveno funkcijo. V praksi se pogosto uporablja logaritemsko Pearsonova porazdelitev tipa III. Bolj smiselno je uporabiti več porazdelitvenih funkcij in nato primerjati rezultate analiz.

Različni avtorji so pokazali, da lahko izbira porazdelitve in metode ocenjevanja parametrov bistveno vpliva na rezultate verjetnostne analize. Metoda momentov (MOM) da npr. dobre ocene vrednosti parametrov le v primeru simetričnih podatkov (Hosking in Wallis, 1997). V primeru, ko imamo na razpolago manj let podatkov, so lahko vzorčni momenti precej različni od teoretičnih momentov porazdelitvene funkcije. V tem primeru je bolj primerno uporabiti momente L, ki izhajajo iz verjetnostno obteženih momentov (angl. *probability weighted moments*) in so manj občutljivi na osamelce, ki se lahko nahajajo v vzorcu. Tudi Hosking (1990) je ugotovil, da je metoda momentov L (LM) v primeru majhnih in srednje velikih vzorcev bolj učinkovita kot metoda največjega verjetja (MLE).

3.2 Testi

Za izbiro najustreznejše porazdelitve smo izvedli različne grafične teste: grafična primerjava empirične in teoretične porazdelitvene funkcije, diagrami P-P (črka P označuje verjetnost oziroma angl. *probability*), histogrami z gostoto verjetnosti

porazdelitve ter različne statistične teste ustreznosti porazdelitve, ki jih lahko razdelimo v štiri skupine:

- Testi, ki primerjajo izmerjene in ocenjene vrednosti spremenljivk: RMSE (koren povprečne kvadratne napake) (angl. *root mean square error*), MAE (povprečna absolutna napaka) (angl. *mean absolute error*), RMAE (koren povprečne absolutne napake) (angl. *relative mean absolute error*) (preglednica 2).
- Testi, ki primerjajo empirične in teoretične verjetnosti podatkov: absolutni in relativni testi (preglednica 3).
- Testi, ki služijo za testiranje hipotez: Hi-kvadrat, Kolmogorov-Smirnov, Anderson-Darling in test PPCC (angl. *probability plot correlation coefficient*) (preglednica 4).
- Test, ki pri izračunu testne statistike upošteva tudi število parametrov porazdelitve: AIC test (angl. *Akaike information criteria*) (enačba (1)).

Testi RMSE, MAE in RMAE primerjajo izmerjene in ocenjene vrednosti pretokov. Pri teh testih ne obstajajo slabe ali dobre vrednosti testnih statistik. Izmerjenim podatkom se najbolje prilega porazdelitev, ki ima najnižjo testno statistiko. Testne statistike izračunamo z izrazi, ki so prikazani v preglednici 2 (Maidment, 1993). Velik vpliv na testno statistiko RMSE imajo osamelci, testna statistika MAE pa je na njih manj občutljiva (Maidment, 1993; Swanson et al., 2011). Pri diagramih P-P primerjamo vrednosti teoretične in empirične porazdelitve. Bolj kot se točke prilegajo premici $y=x$, boljše je ujemanje med podatki in testirano teoretično porazdelitveno funkcijo.

Preglednica 2: Statistični testi (I).

Table 2: Statistical tests (I).

Test/Test	Enačba/Equation
RMSE	$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - Q_{max,i})^2}$,
MAE	$MAE = \frac{1}{n} \sum_{i=1}^n x_i - Q_{max,i} $,
RMAE	$RMAE = \frac{MAE}{Q_{max}}$.

V preglednici 2 x_i predstavlja izračunano vrednost spremenljivke, $Q_{max,i}$ izmerjeno vrednost

spremenljivke in \bar{Q}_{max} povprečno vrednost spremenljivke.

Preglednica 3: Statistični testi (II).

Table 3: Statistical tests (II).

Test/Test	Enačba/Equation
ABS ₁	$ABS_1 = \frac{\sum_{i=1}^n F_i - F_x }{n}$,
ABS ₂	$ABS_2 = \sqrt{\frac{\sum_{i=1}^n (F_i - F_x)^2}{n}}$,
REL ₁	$REL_1 = \frac{\sum_{i=1}^n F_i - F_x }{\sum_{i=1}^n F_i}$,
REL ₂	$REL_2 = \frac{\sqrt{\sum_{i=1}^n (F_i - F_x)^2 / n}}{\sum_{i=1}^n F_i / n}$,
REL ₃	$REL_3 = \sqrt{\frac{\sum_{i=1}^n (F_i - F_x)^2}{\sum_{i=1}^n F_i^2}}$.

V drugo skupino uvrščamo teste, ki primerjajo empirične in teoretične verjetnosti podatkov (Ricci, 2005). F_i označuje empirično verjetnost, F_x pa teoretično verjetnost danega podatka (preglednica 3). Teoretična verjetnost je odvisna od testirane porazdelitvene funkcije. Tudi pri tem sklopu testov ne obstajajo dobre ali slabe vrednosti testnih statistik. Ti testi nam služijo predvsem za primerjavo obravnavanih porazdelitvenih funkcij in metod ocenjevanja parametrov.

V preglednici 4 so prikazani testi, s katerimi lahko testiramo hipotezo, da obravnavani podatki sledijo testirani porazdelitveni funkciji. Pri tem nas zanima skladnost porazdelitvene funkcije in podatkov na celotnem intervalu. V primeru, da vrednost statistike pade nad mejno vrednost testa (oziroma pod kritično vrednost pri testu PPCC), lahko testirano porazdelitev zavrnemo z izbrano stopnjo značilnosti. V primeru, da parametre porazdelitev določimo iz vzorca, ki ga testiramo, moramo običajne mejne vrednosti ustrezno korigirati (Maidment, 1993; Kottogoda in Rosso, 2008).

Nekaj neznank iz preglednice 4 je bilo uporabljenih že v preglednicah 2 in 3. O_i in E_i pa označujeta dejansko ter pričakovano število elementov v posameznem razredu testa Hi-kvadrat. Pri testu Hi-kvadrat moramo razrede določiti tako, da so v posameznem razredu vsaj 3 oziroma vsaj 5

elementov (Haan, 2002). Dela, ki obravnavajo tematiko mejnih vrednosti, je v svoji nalogi zbral Bezak (2012).

Preglednica 4: Statistični testi (III).

Table 4: Statistical tests (III).

Test/Test	Enačba/Equation
Hi-kvadrat	$\chi^2_c = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$,
Kolmogorov-Smirnov	$D = \max F_x(Q_{max,i}) - F_i(x) $,
Anderson-Darling	$A^2 = -n - \sum_{i=1}^n \frac{(2i-1) \ln F_x(Q_{max,i}) }{n} + \sum_{i=1}^n \frac{(2i-1) \ln \{1 - F_x(Q_{max,n-i+1})\}}{n}$,
PPCC	$PPCC = \frac{\sum_{i=1}^n (x_i - \bar{x})(Q_{max,i} - \bar{Q}_{max})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (Q_{max,i} - \bar{Q}_{max})^2}}$

Test, ki pri izračunu testne statistike upošteva tudi število parametrov porazdelitve, je test AIC. V enačbi (1) je zapisan izraz, s katerim lahko izračunamo vrednost testne AIC statistike:

$$AIC = n \log(RMSE^2) + 2k \quad (1)$$

Tako obliko testa, ki pri izračunu testne statistike upošteva tudi število parametrov porazdelitve (neznanka k), sta v svoji raziskavi uporabila tudi Karmakar in Simonovic (2008). Tudi test AIC nam lahko pomaga pri primerjavi porazdelitvenih funkcij in metod ocenjevanja parametrov. Pri vseh uporabljenih testih n označuje število letnih maksimumov, ki jih imamo na razpolago za analize.

Različni avtorji v svojih raziskavah uporabljajo različne teste. Karmakar in Simonovic (2008) sta uporabila teste RMSE, AIC, BIC (angl. *Bayesian information criteria*) in Hi-kvadrat. Testa RMSE in AIC sta bila uporabljena tudi v študiji, ki sta jo naredila Zhang in Singh (2006) ter Zhang in Singh (2007). Haddad in Rahman (2011) pa sta uporabila teste AIC, BIC in Anderson-Darling pri določevanju najustreznejše porazdelitve na več kot 40 vodomernih postajah v Avstraliji.

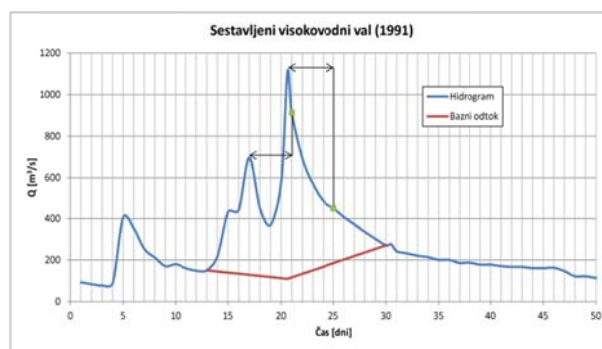
4. Osnovne statistične analize

Za vse tri spremenljivke (konic, volumni, časi trajanja valov) smo najprej izvedli osnovne statistične analize z izračunom ocen karakterističnih števil oz. statistik obravnavanih vzorcev (preglednica 5). S karakterističnimi števili prikažemo lastnosti statističnih populacij (Brilly in Šraj, 2005). Določamo jih predvsem z momenti. Med karakteristična števila spadata tudi srednja vrednost in standardna deviacija. Srednja vrednost ali matematično upanje je moment prvega reda. Centralni moment drugega reda pa predstavlja varianco, katere kvadratni koren je enak standardni deviaciji, ki je mera razpršenosti vzorca (Turk, 2012).

Največji pretok, ki je bil zabeležen v obravnavanem obdobju, se je zgodil leta 1990 in je znašal 2326 m³/s, pripadajoči volumen je drugi največji, kljub temu, da je val trajal 15 dni. Največji volumen v obravnavanem obdobju pripada valu iz leta 1968, ki ima maksimalno konico veliko 1411 m³/s in je trajal kar 30 dni. Tako dolga trajanja visokovodnih valov so posledica kompleksnih hidrogramov (sestavljene valovi), ki jih dobimo, kadar se zgodi več zaporednih in odvisnih konic vala (slika 3). Odvisnost oz. neodvisnost zaporednih konic smo določali s pomočjo empirične enačbe, ki izračuna čas med konico vala in koncem površinskega dela odtoka v odvisnosti od velikosti prispevne površine (Pugelj, 2012).

Povprečna vrednost obravnavanih konic znaša 1234 m³/s, povprečni volumen 337*10⁶ m³ in

povprečno trajanje valov 12,9 dneva. Več kot 25 % konic se je zgodilo v mesecu novembru in več kot polovica letnih maksimumov obravnavanega obdobja (1953 do 2010) se je zgodila v jesenskem času. Prav tako so se v jesenskem obdobju zgodile tri največje konic in sicer leta 1990, 2010 in 1982. Vse to je v skladu z značilnim dežno-snežnim režimom reke Save v spodnjem toku.



Slika 3: Primer kompleksnega hidrograma in izločanja baznega odtoka.

Figure 3: Example of complex hydrograph and base flow separation.

Izračunali smo tudi vrednosti korelacijskih koeficientov za pare obravnavanih spremenljivk (vzorec letnih maksimumov). Za pare konica - volumen znaša Pearsonov koeficient korelacije 0,52, za pare volumen - čas trajanja 0,68 in za pare konica - čas trajanja -0,14. Relativno visok koeficient korelacije med volumni in časi trajanja visokovodnih valov ter nizek koeficient korelacije med konicami in časi trajanja je v največji meri posledica kompleksnih hidrogramov (sestavljenih valov).

Preglednica 5: Osnovne statistike posameznih spremenljivk.

Table 5: Basic statistics of individual variables.

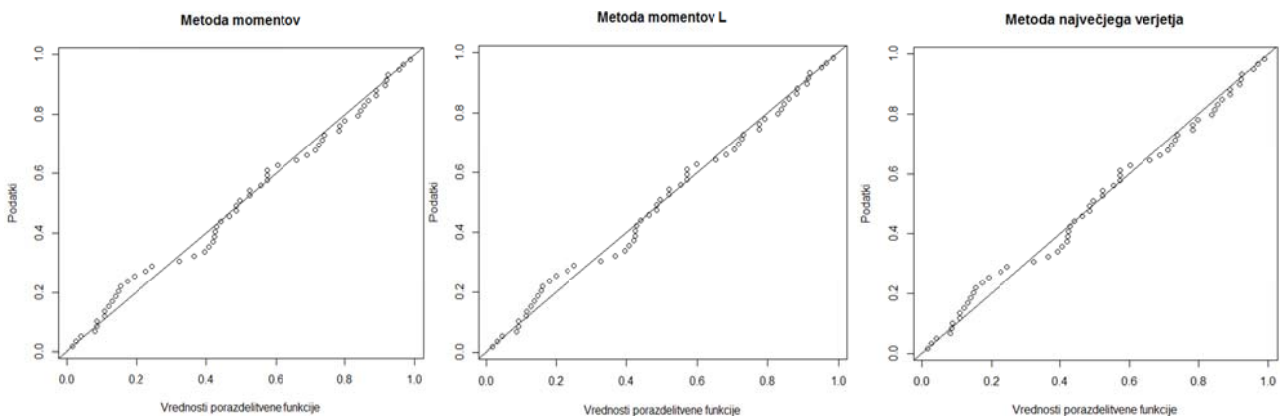
Statistika/Statistic	Pretok/Discharge [m ³ /s]	Volumen/Volume [10 ⁶ m ³]	Trajanje/Duration [dni/days]
Povprečje	1234,4	337,5	12,9
Maksimum	2326	765,8	30
Minimum	579	122,2	6
Standardna deviacija	391,6	152,2	4,8
Mediana	1174	300,1	12
Mod	1254	/	12

5. Verjetnostna analiza visokovodnih konic

5.1 Rezultati analiz z uporabo različnih teoretičnih porazdelitev in enačb za oceno parametrov

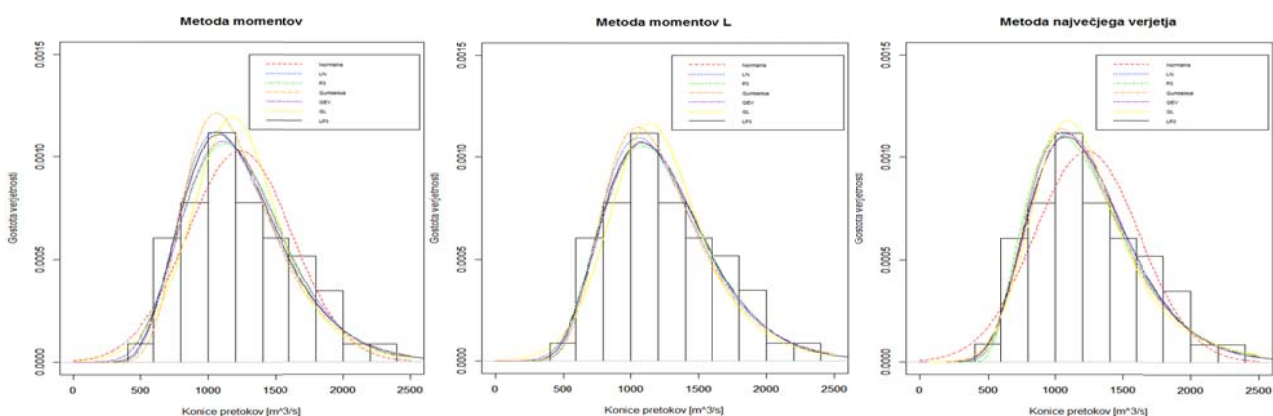
Verjetnostno analizo smo naredili s sedmimi različnimi porazdelitvenimi funkcijami in pri tem uporabili tri načine ocenjevanja parametrov. Za primerjavo med metodami in izbiro najustrežnejše porazdelitve in metode ocenjevanja smo najprej

izvedli različne grafične teste. Na slikah 4 so prikazani diagrami P-P pri uporabi logaritemsko Pearsonove porazdelitve tipa III in vseh treh načinov ocenjevanja parametrov. Merjeni podatki so prikazani z uporabo Weibullove enačbe. Opazimo lahko, da je z diagrami P-P težko določiti razlike med posameznimi načini ocenjevanja parametrov. Vidimo lahko le, da se je metoda momentov L nekoliko boljše ujemala s podatki kot metoda navadnih momentov ali metoda največjega verjetja.



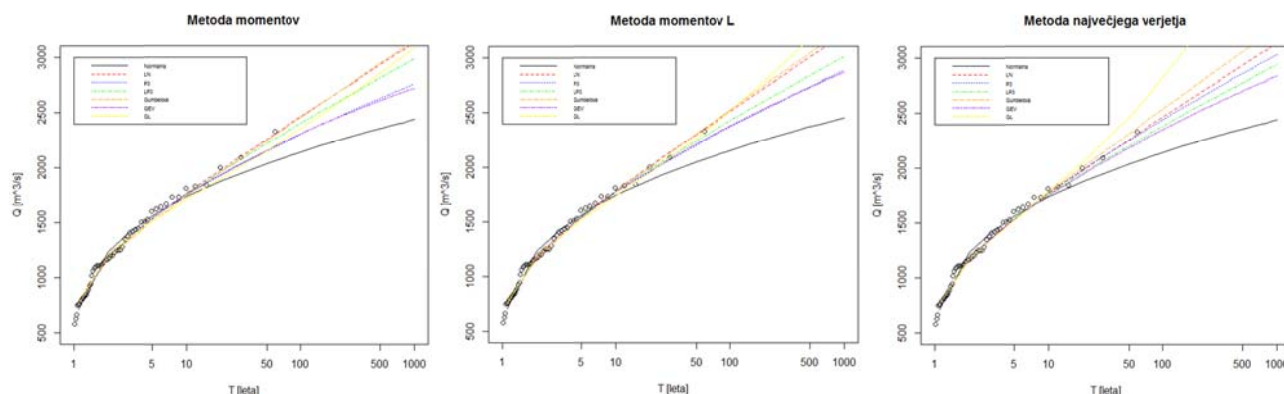
Slika 4: Diagrami P-P za podatke o pretokih in logaritemsko Pearsonovo III porazdelitev pri ocenjevanju parametrov po vseh treh uporabljenih metodah (metoda momentov, metoda momentov L, metoda največjega verjetja).

Figure 4: P-P plots for discharge data and log-Pearson III distribution when parameters are estimated with method of moments, method of L-moments and maximum likelihood method, respectively.



Slika 5: Primerjava med histogrami podatkov o konicah in gostotami verjetnosti za vse uporabljene porazdelitvene funkcije in vse tri načine ocenjevanja parametrov (metoda momentov, metoda momentov L, metoda največjega verjetja).

Figure 5: Comparison between histograms (discharge data) and probability density functions (PDF) for all considered distribution functions for method of moments, method of L-moments and maximum likelihood method, respectively.



Slika 6: Rezultati verjetnostne analize konic visokovodnih valov (vzorec letnih maksimumov) z uporabo različnih porazdelitvenih funkcij in treh načinov ocenjevanja parametrov.

Figure 6: Flood frequency analysis results for discharge values (annual maximum series sample) for different distribution functions and parameter estimation techniques.

Če grafično prikažemo histograme podatkov o konicah in gostote verjetnosti posameznih porazdelitvenih funkcij za vse tri načine ocenjevanja parametrov (slika 5), lahko zaključimo, da se normalna porazdelitev občutno slabše ujema s podatki kot ostale uporabljene porazdelitvene funkcije. Pri ocenjevanju parametrov po metodi momentov in metodi momentov L se s podatki nekoliko slabše ujemata tudi Gumbelova in generalizirana logistična porazdelitev (GL).

Grafični prikaz rezultatov verjetnostne analize (slika 6) pokaže podobno sliko. Tudi tu so merjeni podatki prikazani z Weibullovo enačbo. Opazimo lahko, da je normalna porazdelitev pri vseh treh načinih ocenjevanja parametrov dala najnižje ocenjene vrednosti pretokov. Pri metodi momentov L in metodi največjega verjetja pa je najvišje ocenjene vrednosti pretokov dala generalizirana logistična porazdelitev, medtem ko je pri metodi momentov najvišje vrednosti ocenjenih pretokov izkazala Gumbelova porazdelitev. Med ostalimi porazdelitvami pri vseh treh načinih ocenjevanja parametrov ni bilo zelo izrazitih razlik. Iz slike 6 lahko vidimo, da so krivulje, ki povezujejo ocenjene vrednosti pretokov in povratne dobe pri metodi momentov L najmanj razpršene, medtem ko so razlike največje pri metodi največjega verjetja. Razlike med porazdelitvenimi funkcijami se povečujejo z večanjem povratne dobe.

5.2 Rezultati statističnih testov

Predstavljeni grafični testi služijo prvim ocenam o primernosti posameznih porazdelitev in metod za oceno parametrov. Za določitev najustreznejše porazdelitvene funkcije in za primerjavo posameznih porazdelitev in metod ocenjevanja parametrov pa smo uporabili tudi več statističnih testov, ki so predstavljeni v poglavju 3. Pri vseh uporabljenih testih je bila predpostavljena ničelna domneva oziroma hipoteza, da obravnavana slučajna spremenljivka sledi testirani porazdelitveni funkciji. V primeru, da je rezultat testa padel izven predpisanega območja mejnih vrednosti, smo ničelno domnevo zavrnil in sprejeli alternativno ob izbrani stopnji tveganja 0,05. Določitev najustreznejše porazdelitvene funkcije je predpogoj za izvedbo verjetnostne analize s funkcijo kopula. Pri kopuli moramo namreč določiti oziroma izbrati tudi robne porazdelitvene funkcije (angl. *marginal distribution functions*), ki so lahko različne za obravnavane odvisne spremenljivke, ki jih želimo modelirati.

Ugotovili smo, da večina uporabljenih testov pokaže najboljše ujemanje podatkov o pretokih z logaritemsko Pearsonovo porazdelitvijo tipa III. Najslabše sta se s podatki o pretokih ujemali normalna in generalizirana logistična porazdelitev (GL). Ostale porazdelitvene funkcije (logaritemsko normalna, Pearsonova III, Gumbelova in GEV) pa so izkazale le nekoliko slabše ujemanje kot logaritemsko Pearsonova III porazdelitev. Poleg

tega smo ugotovili, da je metoda momentov L v veliki večini primerov (različne porazdelitve in različni testi) dala nižje (pri testu PPCC pa višje) rezultate testnih statistik kot metoda momentov ali metoda največjega verjetja. Med metodo momentov in metodo največjega verjetja pa ni bilo tako enoznačne in izrazite razlike.

Preglednica 6: Vrednosti testa Anderson-Darling in kritične vrednosti (stopnja značilnosti 0,05) za vse tri načine ocenjevanja parametrov (konic pretokov).

Table 6: Anderson-Darling test results and critical values (significance level 0,05) for all parameter estimation methods (discharges).

Porazdelitev/ Distribution function	MOM (KV)	LM (KV)	MLE (KV)
Normalna	0,545 (0,762)	0,531 (0,762)	0,545 (0,762)
Logaritemsko normalna	0,311 (0,762)	0,277 (0,762)	0,311 (0,762)
Pearsonova III	0,275 (0,752)	0,239 (0,752)	0,288 (0,752)
Logaritemsko Pearsonova III	0,289 (0,752)	0,251 (0,752)	0,285 (0,752)
Gumbelova	0,531 (0,777)	0,348 (0,777)	0,335 (0,777)
GEV	0,277 (0,56)	0,254 (0,56)	0,303 (0,56)
GL	0,545 (0,698)	0,389 (0,698)	0,389 (0,698)

V preglednici 6 so prikazani rezultati testa Anderson-Darling za vse tri načine ocenjevanja parametrov porazdelitev. Poleg tega so prikazane tudi kritične vrednosti (KV) s stopnjo značilnosti 0,05. Iz preglednice 6 lahko vidimo, da je metoda momentov L dala nižje vrednosti testnih statistik kot metoda momentov ali metoda največjega verjetja. Pri testu Anderson-Darling sicer logaritemsko Pearsonova porazdelitev tipa III ni dala najnižjih vrednosti testnih statistik, vendar pa je ta porazdelitvena funkcija izkazala najboljše rezultate pri drugih uporabljenih testih, ki pa niso tako zanimivi za grafični prikaz. Iz preglednice 6 lahko vidimo, da nobene izmed porazdelitvenih funkcij (Pearsonova III, logaritemsko Pearsonova III, logaritemsko normalna in GEV) ne moremo

zavriniti z izbrano stopnjo značilnosti 0,05 pri nobenem izmed načinov ocenjevanja parametrov. Podobne ugotovitve veljajo tudi za druge uporabljene statistične teste. S stopnjo značilnosti 0,05 je bila pri testu Hi-kvadrat zavrnjena le normalna porazdelitev.

Preglednica 7: Ocenjene vrednosti pretokov za vse porazdelitvene funkcije z uporabo metode momentov L.

Table 7: Estimated discharge values for all distributions using the method of L-moments.

Porazdelitvena funkcija/Distribution function	Q 10 [m ³ /s]	Q 100 [m ³ /s]
Normalna	1739	2150
Logaritemsko normalna	1784	2509
Pearsonova III	1770	2523
Logaritemsko Pearsonova III	1768	2368
Gumbelova	1775	2423
GEV	1768	2376
GL	1736	2513

V preglednici 7 so prikazane končne ocenjene vrednosti pretokov z 10-letno in s 100-letno povratno dobo. Prikazani so rezultati v primeru ocenjevanja parametrov z metodo momentov L, ki je pri večini uporabljenih testov dala boljše rezultate kot metoda momentov ali metoda največjega verjetja.

6. Verjetnostna analiza volumnov visokovodnih valov

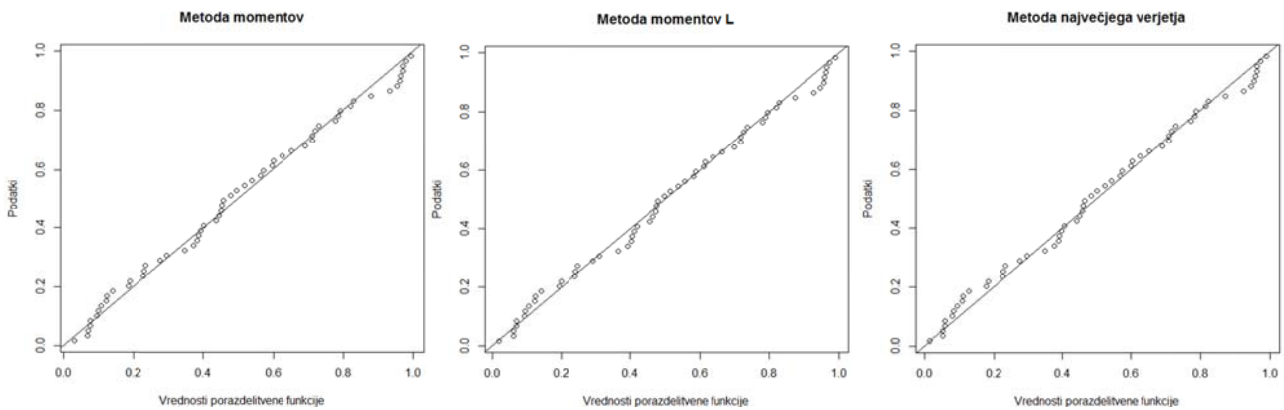
6.1. Rezultati analiz z uporabo različnih teoretičnih porazdelitev in enačb za oceno parametrov

Tudi pri verjetnostni analizi volumnov visokovodnih valov smo uporabili enake porazdelitvene funkcije in metode ocenjevanja parametrov kot pri verjetnostnih analizah konic. Za primerjavo metod in določitev najustreznejše porazdelitvene funkcije, ki bi jo lahko potem

uporabili kot robno porazdelitev pri funkciji kopula, smo ponovno uporabili grafične in statistične teste. Za prikaz merjenih podatkov je bila uporabljena Weibullova enačba (sliki 7 in 9).

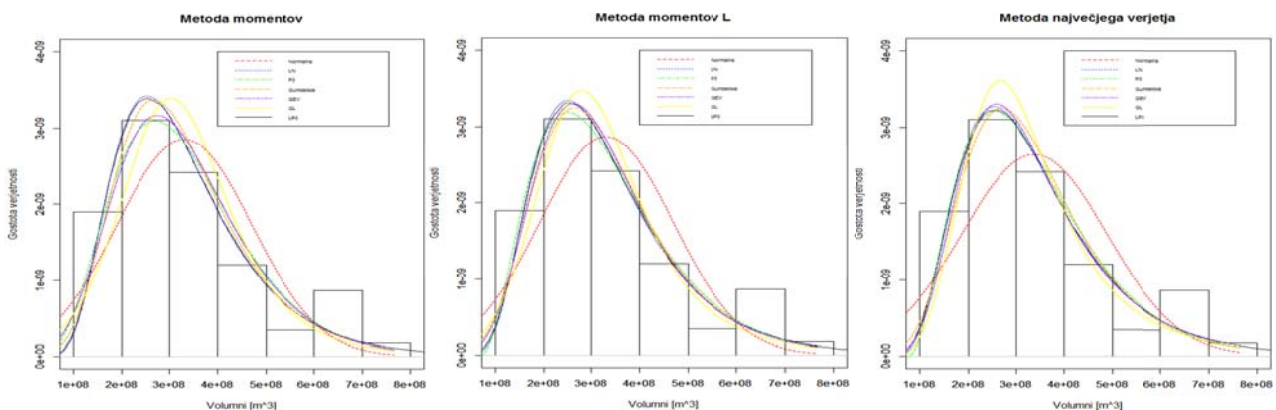
Na sliki 7 so prikazani diagrami P-P za Pearsonovo porazdelitev tipa III. Vidimo lahko, da dobimo v primeru uporabe metode momentov L nekoliko boljše ujemanje med podatki in teoretično porazdelitveno funkcijo kot pri ocenjevanju parametrov z metodo momentov ali metodo največjega verjetja, vendar razlike med posameznimi načini niso zelo izrazite. Grafična predstavitev histograma podatkov (volumni letnih

maksimumov) in gostot verjetnosti uporabljenih porazdelitvenih funkcij kaže, da se tudi pri analizah volumnov s podatki najslabše ujema normalna porazdelitev (slika 8). Nekoliko slabše prileganje podatkom lahko zasledimo tudi pri generalizirani logistični porazdelitvi. Opazimo lahko, da z oceno parametrov porazdelitev po metodi momentov L ali metodi največjega verjetja dobimo boljše ujemanje podatkov s posameznimi porazdelitvami kot z uporabo metode navadnih momentov. Logaritemsko normalna porazdelitev, Pearsonova III porazdelitev, logaritemsko Pearsonova III porazdelitev in porazdelitev ekstremnih vrednosti dajo podobne rezultate.



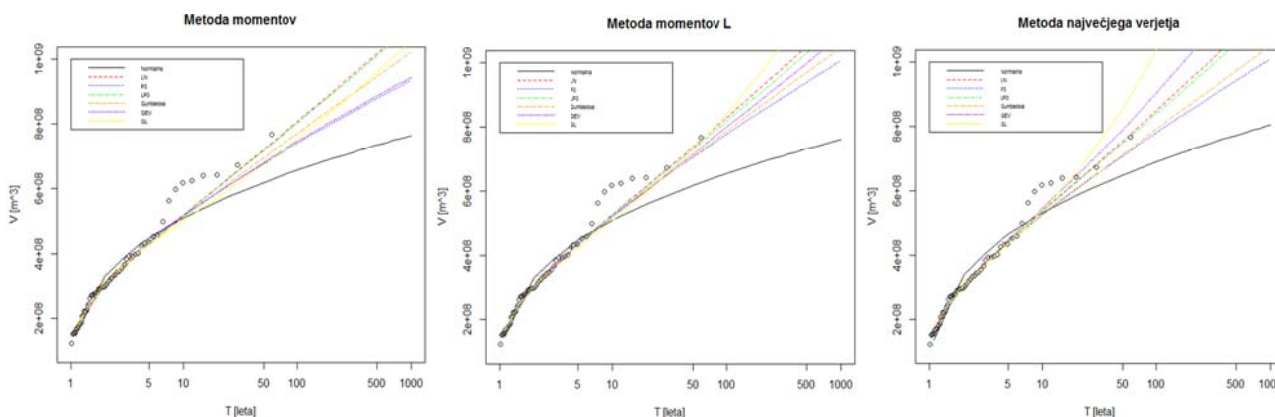
Slika 7: Diagrami P-P za podatke o volumnih in Pearsonovo III porazdelitev pri ocenjevanju parametrov po vseh treh uporabljenih metodah (metoda momentov, metoda momentov L, metoda največjega verjetja).

Figure 7: P-P plots for volume data and Pearson III distribution when parameters are estimated with method of moments, method of L-moments and maximum likelihood method, respectively.



Slika 8: Primerjava med histogrami podatkov o volumnih in gostotami verjetnosti za vse uporabljene porazdelitvene funkcije in vse tri načine ocenjevanja parametrov (metoda momentov, metoda momentov L, metoda največjega verjetja).

Figure 8: Comparison between histograms (volume data) and probability density functions (PDF) for all considered distribution functions for method of moments, method of L-moments and maximum likelihood method, respectively.



Slika 9: Rezultati verjetnostne analize volumnov visokovodnih valov (vzorec letnih maksimumov) z uporabo vseh porazdelitvenih funkcij in treh načinov ocenjevanja parametrov.

Figure 9: Flood frequency analysis results for volume values (annual maximum series sample) for all used distribution functions and parameter estimation techniques.

Na sliki 9 so prikazane krivulje, ki povezujejo ocenjene vrednosti volumnov s pripadajočimi povratnimi dobami. Normalna porazdelitev je pri vseh treh načinih ocenjevanja parametrov porazdelitev dala najnižje ocenjene vrednosti volumnov. Generalizirana logistična porazdelitev je pri metodi momentov L in metodi največjega verjetja dala najvišje ocenjene vrednosti volumnov, medtem ko je pri metodi momentov najvišje ocene izkazala logaritemsko normalna porazdelitev. Pri metodi momentov L so razlike med uporabljenimi porazdelitvenimi funkcijami najmanjše, medtem ko je odstopanje med različnimi porazdelitvami največje pri metodi največjega verjetja. Samo z uporabo grafičnih testov je težko določiti, katera porazdelitvena funkcija se najboljše prilaga merjenim podatkom, zato je za bolj zanesljive zaključke o ustreznosti posameznih metod potrebno uporabiti tudi statistične teste.

6.2 Rezultati statističnih testov

Tudi pri verjetnostnih analizah volumnov visokovodnih valov smo za določitev najustreznejše porazdelitvene funkcije uporabili statistične teste, ki so prikazani v poglavju 3. Ugotovili smo, da je najboljše ujemanje s podatki pri večini uporabljenih testov izkazala Pearsonova porazdelitev tipa III. Tudi pri analizah volumnov smo ugotovili, da je ocena parametrov po metodi momentov L pri večini porazdelitev dala boljše

ujemanje s podatki kot metoda navadnih momentov ali metoda največjega verjetja. Tudi tokrat je najslabše ujemanje s podatki dala normalna porazdelitev, slabše ujemanje pa je bilo ugotovljeno tudi pri uporabi generalizirane logistične porazdelitve. V preglednici 8 so prikazani rezultati testa Anderson-Darling za vse tri načine ocenjevanja parametrov in vseh sedem uporabljenih porazdelitvenih funkcij. Poleg tega so prikazane tudi mejne vrednosti (stopnja značilnosti 0,05) za posamezne porazdelitvene funkcije. S pomočjo preglednice 8 lahko potrdimo ugotovitve, ki smo jih dobili z uporabo grafičnih in ostalih statističnih testov. Vidimo pa lahko tudi, da Pearsonove porazdelitve tipa III, ki je pri večini primerov izkazala najboljše rezultate, ne moremo zavrniti z izbrano stopnjo značilnosti 0,05. S statističnimi testi smo ugotovili, da je bila pri nekaterih testih zavrnjena le normalna porazdelitev (Anderson-Darling, PPCC, Hi-kvadrat), ostalih porazdelitvenih funkcij (Pearsonova III, logaritemsko normalna, logaritemsko Pearsonova III, GEV), pa nismo mogli zavrniti s stopnjo značilnosti 0,05.

V preglednici 9 so prikazane končne ocenjene vrednosti volumnov visokovodnih valov z 10-letno in s 100-letno povratno dobo pri ocenjevanju parametrov z metodo momentov L, ki je dala najboljše rezultate statističnih in grafičnih testov.

Preglednica 8: Vrednosti testa Anderson-Darling in kritične vrednosti (stopnja značilnosti 0,05) za vse tri načine ocenjevanja parametrov (volumni valov).

Table 8: Anderson-Darling test results and critical values (significance level 0,05) for all parameter estimation methods (wave volumes).

Porazdelitev/ Distribution function	MOM (KV)	LM (KV)	MLE (KV)
Normalna	1,224 (0,762)	1,231 (0,762)	1,454 (0,762)
Logaritemsko normalna	0,319 (0,762)	0,287 (0,762)	0,303 (0,762)
Pearsonova III	0,360 (0,752)	0,282 (0,752)	0,322 (0,752)
Logaritemsko Pearsonova III	0,318 (0,752)	0,286 (0,752)	0,304 (0,752)
Gumbelova	0,355 (0,777)	0,303 (0,777)	0,349 (0,777)
GEV	0,374 (0,56)	0,296 (0,56)	0,371 (0,56)
GL	0,611 (0,698)	0,356 (0,698)	0,507 (0,698)

Preglednica 9: Ocenjene vrednosti volumnov za vse porazdelitvene funkcije z uporabo metode momentov L.

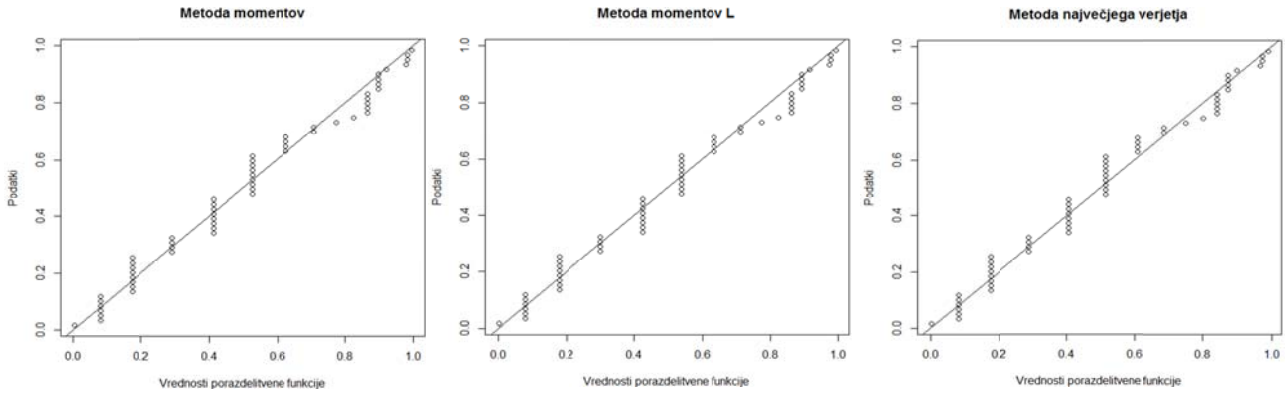
Table 9: Estimated volume values for all distributions using the method of L-moments.

Porazdelitvena funkcija/Distribution function	V 10 [10 ⁶ m ³]	V 100 [10 ⁶ m ³]
Normalna	528	684
Logaritemsko normalna	546	872
Pearsonova III	546	835
Logaritemsko Pearsonova III	547	892
Gumbelova	540	825
GEV	540	868
GL	527	910

7. Verjetnostna analiza trajanj visokovodnih valov

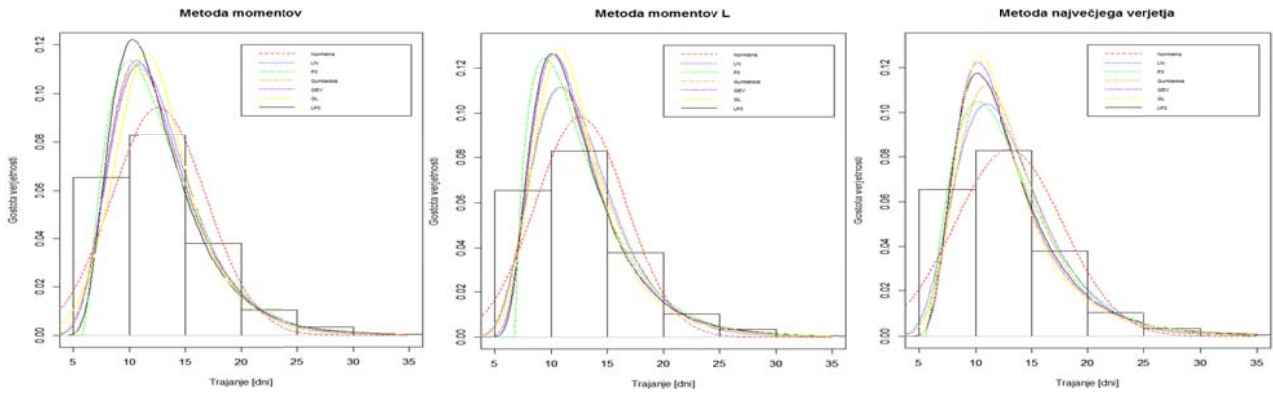
7.1. Rezultati analiz z uporabo različnih teoretičnih porazdelitev in enačb za oceno parametrov

Tudi pri verjetnostnih analizah trajanj visokovodnih valov smo uporabili enake porazdelitvene funkcije, metode ocenjevanja parametrov porazdelitev ter grafične in statistične teste kot pri analizah konic in volumnov valov. Diagrami P-P v primeru trajanj visokovodnih valov ne kažejo očitnih razlik med metodami za oceno parametrov porazdelitev (slika 10). Ker smo pri analizah uporabili dnevne vrednosti pretokov z vključenimi maksimalnimi konicami, smo dobili več valov z enakim časom trajanja. Z uporabo diagramov P-P je zato težko določiti razliko med posameznimi metodami ocenjevanja parametrov porazdelitve. Podobno se izkaže tudi pri histogramih podatkov (časi trajanja visokovodnih valov) in gostotah verjetnosti uporabljenih porazdelitvenih funkcij (slika 11). Ta grafični test nam služi kot dodatna kontrola izvedenih statističnih testov. Samo na osnovi slike 11 je težko določiti razlike med porazdelitvenimi funkcijami in metodami ocenjevanja parametrov. Iz krivulj na sliki 12, ki povezujejo ocenjene vrednosti časov trajanja in pripadajoče povratne dobe, lahko vidimo, da je ponovno pri vseh treh načinih ocenjevanja parametrov najnižje ocenjene vrednosti časov trajanja valov dala normalna porazdelitev. Porazdelitev GL pa je pri metodi momentov L in metodi največjega verjetja izkazala najvišje ocenjene vrednosti časov trajanja. Logaritemsko Pearsonova III porazdelitev je dala najvišje ocenjene vrednosti pri metodi momentov. Iz slike 12 lahko vidimo, da se pri nižjih vrednostih ponovitvenega intervala (manj kot 10 let) vse porazdelitvene funkcije dobro ujemajo s podatki, ki so prikazani z Weibullovo enačbo, pri višjih vrednostih povratnih dob pa nobena porazdelitvena funkcija ne uspe dobro slediti podatkom.



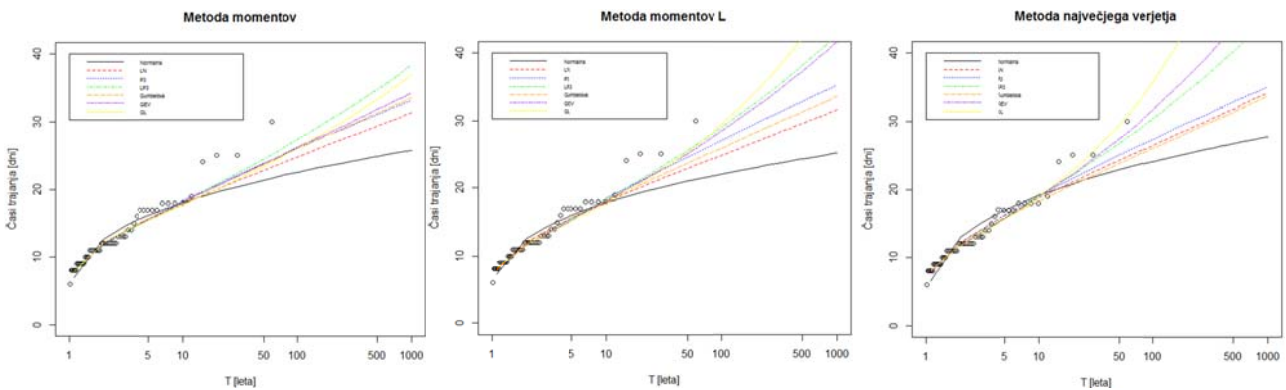
Slika 10: Diagrami P-P za podatke o časih trajanja in logaritemsko Pearsonovo III porazdelitev pri ocenjevanju parametrov po vseh treh uporabljenih metodah (metoda momentov, metoda momentov L, metoda največjega verjetja).

Figure 10: P-P plots for duration data and log-Pearson III distribution when parameters are estimated with method of moments, method of L-moments and maximum likelihood method, respectively.



Slika 11: Primerjava med histogrami podatkov o časih trajanja in gostotami verjetnosti za vse uporabljene porazdelitvene funkcije in vse tri načine ocenjevanja parametrov (metoda momentov, metoda momentov L, metoda največjega verjetja).

Figure 11: Comparison between histograms (duration data) and probability density functions (PDF) for all considered distribution functions for method of moments, method of L-moments and maximum likelihood method, respectively.



Slika 12: Rezultati verjetnostne analize časov trajanja visokovodnih valov (vzorec letnih maksimumov) z uporabo vseh porazdelitvenih funkcij in treh načinov ocenjevanja parametrov.

Figure 12: Flood frequency analysis results for duration values (annual maximum series sample) for all used distribution functions and parameter estimation techniques.

7.2 Rezultati statističnih testov

Tudi pri analizah trajanj visokovodnih valov smo uporabili statistične teste, ki so prikazani v poglavju 3. Pri večini uporabljenih testov je najboljše ujemanje s podatki izkazala logaritemsko Pearsonova porazdelitev tipa III. V povprečju je metoda največjega verjetja pri testih RMSE, MAE, RMAE, AIC in 5 testih, ki primerjajo empirične in teoretične verjetnosti, dala nekoliko boljše ujemanje kot metoda momentov L, vendar je bila kombinacija logaritemsko Pearsonove III porazdelitve in metode največjega verjetja zavrnjena pri testu Kolmogorov-Smirnov s stopnjo značilnosti 0,05. Pearsonova III porazdelitev in logaritemsko Pearsonova III porazdelitev (metoda momentov L) sta torej dali nižje vrednosti testnih statistik (višje vrednosti pri testu PPCC) pri večini izvedenih testov in obenem nista bili zavrjeni pri nobenem izmed testov, s katerimi lahko testiramo hipoteze. Pri vsaj enem testu so bile zavrjene (vsi trije načini ocenjevanja parametrov) tako normalna porazdelitev, logaritemsko normalna porazdelitev, Gumbelova porazdelitev, porazdelitev GEV in porazdelitev GL. V preglednici 10 so prikazani rezultati testa Anderson-Darling za vse uporabljene porazdelitvene funkcije in vse tri načine ocenjevanja parametrov. Vidimo lahko, da je pri tem testu najboljše ujemanje izkazala kombinacija logaritemsko Pearsonove III porazdelitve in metode momentov L.

Kot že rečeno, je bilo le malce slabše ujemanje dobljeno pri uporabi Pearsonove porazdelitve tipa III. Vidimo lahko tudi, da je bilo veliko porazdelitvenih funkcij zavrjenih s stopnjo značilnosti 0,05. Tako smo npr. zavrnilo normalno, logaritemsko normalno, Gumbelovo in porazdelitev GEV pri vseh treh načinih ocenjevanja parametrov. Zavrjene pa so bile tudi ostale porazdelitvene funkcije, vendar le pri nekaterih metodah ocenjevanja parametrov (Pearsonova III pri metodi največjega verjetja, GL pri metodi momentov in metodi momentov L), logaritemsko Pearsonova III porazdelitev pa ni bila zavrnjena pri nobenem načinu ocenjevanja parametrov. V preglednici 11 so prikazane končne ocenjene vrednosti trajanj visokovodnih valov z 10-letno in s 100-letno povratno dobo. Prikazani so

rezultati za vseh sedem uporabljenih porazdelitvenih funkcij in za metodo momentov L.

Preglednica 10: Vrednosti testa Anderson-Darling in kritične vrednosti (stopnja značilnosti 0,05) za vse tri načine ocenjevanja parametrov (časi trajanja valov)

Table 10: Anderson-Darling test results and critical values (significance level 0,05) for all parameter estimation methods (wave durations)

Porazdelitev/ Distribution function	MOM (KV)	LM (KV)	MLE (KV)
Normalna	2,095 (0,762)	2,153 (0,762)	2,653 (0,762)
Logaritemsko normalna	0,904 (0,762)	0,896 (0,762)	1,065 (0,762)
Pearsonova III	0,711 (0,752)	0,611 (0,752)	0,951 (0,752)
Logaritemsko Pearsonova III	0,642 (0,752)	0,604 (0,752)	0,667 (0,752)
Gumbelova	0,798 (0,777)	0,798 (0,777)	0,943 (0,777)
GEV	0,762 (0,563)	0,617 (0,563)	0,633 (0,563)
GL	1,134 (0,698)	0,708 (0,698)	0,630 (0,698)

Preglednica 11: Ocenjene vrednosti časov trajanja visokovodnih valov za vse porazdelitvene funkcije z uporabo metode momentov L

Table 11: Estimated duration values for all distributions using the method of L-moments

Porazdelitvena funkcija/Distribution function	T 10 [dni/days]	T 100 [dni/days]
Normalna	18,7	23,4
Logaritemsko normalna	18,7	26,5
Pearsonova III	19,5	29,8
Logaritemsko Pearsonova III	19,2	32,3
Gumbelova	19,1	27,7
GEV	18,9	31,7
GL	18,6	32,6

8. Razprava in zaključki

V raziskavi smo pri verjetnostni analizi vseh treh spremenljivk uporabili več vrst teoretičnih porazdelitev in metod za oceno parametrov porazdelitev. Na ta način smo naredili natančne analize in primerjave rezultatov posameznih kombinacij metod. Tudi mnogi drugi avtorji so v podobnih raziskavah uporabili več različnih porazdelitev (Karmakar in Simonovic, 2008; Genest in Favre, 2007) in metod za oceno parametrov (Hosking, 1990). Za določitev najustreznejše porazdelitvene funkcije in za primerjavo posameznih porazdelitev in metod ocenjevanja parametrov smo uporabili več grafičnih in statističnih testov.

Pri verjetnostni analizi visokovodnih konic so rezultati analiz in testov pokazali, da se s podatki z vodomerne postaje Litija 1 najbolje ujema logaritemsko Pearsonova III porazdelitev. Ta porazdelitvena funkcija je bila kot robna porazdelitev uporabljena tudi pri analizah Zhanga in Singha (2007). Nekoliko slabše rezultate je dala Pearsonova III porazdelitev, tej pa sledita generalizirana porazdelitev ekstremnih vrednosti in logaritemsko normalna porazdelitev. Podobno sta Karmakar in Simonovic (2007) v svoji raziskavi ugotovila, da je porazdelitev gama, med katere spadata tudi Pearsonova III in logaritemsko Pearsonova III porazdelitev, izkazala najboljše ujemanje s podatki in sta jo tako izbrala za robno porazdelitev pri analizah konic. Najslabše ujemanje je v našem primeru pričakovano dala normalna porazdelitev. Pri testiranju normalne porazdelitve (stopnja tveganja 0,05) s testom Hi-kvadrat je prišlo tudi do edine zavrnitve ničelne hipoteze. Ostalih porazdelitvenih funkcij pri nobenem testu nismo mogli zavrniti z izbrano stopnjo značilnosti 0,05. Po našem mnenju je poleg testov iz skupine 1, 2 ali 4 smiselno uporabiti vsaj en test iz skupine 3. S temi testi lahko testiramo hipoteze (Kolmogorov-Smirnov, Anderson-Darling, PPCC) in tako preverimo, ali je izbrana porazdelitvena funkcija sploh primerna za izvedbo verjetnostne analize. Tudi grafični testi v našem primeru kažejo na to, da se z merjenimi podatki najslabše ujema normalna porazdelitev, najbolje pa se z merjenimi visokovodnimi konicami ujema

logaritemsko Pearsonova III in Pearsonova III porazdelitev. Za oceno parametrov se je v večini primerov za najbolj učinkovito izkazala metoda momentov L, nekoliko slabše rezultate sta dali metoda momentov in metoda največjega verjetja. Tudi Hosking in Wallis (1997) sta v svoji raziskavi ugotovila, da je metoda momentov L velikokrat učinkovitejša kot najpogosteje uporabljena metoda navadnih momentov. Hosking (1990) pa je zapisal, da je metoda momentov L v primeru normalno velikih vzorcev boljša kot metoda največjega verjetja. Metodi momentov in momentov L je v svoji diplomski nalogi obravnaval tudi Kučič (2007). Ugotovil je, da je metoda momentov L veliko bolj stabilna in da se razlike med metodama večajo s povratno dobo.

Pri verjetnostni analizi volumnov smo ugotovili, da se z vzorcem merjenih podatkov najbolje ujema Pearsonova porazdelitev tipa III z oceno parametrov po metodi momentov L. Sledita ji logaritemsko normalna porazdelitev in logaritemsko Pearsonova porazdelitev tipa III z oceno parametrov po metodi največjega verjetja. Nekoliko slabše rezultate pa sta dali Gumbelova porazdelitev in generalizirana porazdelitev ekstremnih vrednosti (metoda momentov L). Tudi pri grafičnih testih je Pearsonova III porazdelitev izkazala dobre rezultate. Kot robne porazdelitve za modeliranje volumnov visokovodnih valov so bile v nekaterih drugih raziskavah izbrane naslednje porazdelitvene funkcije: Gumbelova (Zhang in Singh, 2006), gama (Genest in Favre, 2007) in logaritemsko Pearsonova III (Zhang in Singh, 2006; Zhang in Singh, 2007). Karmakar in Simonovic (2008) pa sta za podatke o volumnih ugotovila, da je med parametričnimi porazdelitvami najboljše ujemanje s podatki izkazala logaritemsko normalna porazdelitev. Z uporabo statističnih testov iz preglednice 4 smo v naši raziskavi ugotovili, da je bila v enem primeru ponovno zavrnjena le normalna porazdelitev, medtem ko ostalih porazdelitvenih funkcij nismo mogli zavrniti z izbrano stopnjo značilnosti 0,05. Z uporabo statističnih testov iz skupin 1, 2, 3 in 4 smo ugotovili, da je ocena parametrov z metodo momentov L ponovno izkazala boljše ujemanje posameznih porazdelitev s podatki kot ocena

parametrov z metodo momentov in metodo največjega verjetja pri večini primerov. Ta dognanja so v skladu z ugotovitvami drugih avtorjev (Hosking, 1990; Hosking in Wallis, 1997), vendar razlika med metodo momentov L in metodo največjega verjetja ni bila tako izrazita kot pri analizah konic.

Verjetnostna analiza trajanja valov je pokazala, da se s podatki najbolje ujema logaritemsko Pearsonova porazdelitev tipa III, sledi pa ji Pearsonova III porazdelitev. Tudi pri analizah trajanja valov se je s podatki najslabše ujemala normalna porazdelitev. Karmakar in Simonovic (2008) sta pri analizah trajanja valov ugotovila, da je med parametričnimi metodami najboljša ujemanje izkazala porazdelitev gama. V družino porazdelitev gama spadata tako Pearsonova III porazdelitev kot logaritemsko Pearsonova III porazdelitev, zato lahko rečemo, da sta raziskovalca prišla do podobnih ugotovitev. Tudi Zhang in Singh (2007) sta kot robno porazdelitev časov trajanja uporabila logaritemsko Pearsonovo III porazdelitev. Pri analizah časov trajanja je v naši raziskavi prišlo do zavrnitve večine uporabljenih kombinacij pri vsaj enem statističnem testu. S stopnjo značilnosti 0,05 smo pri vsaj enem testu in pri vseh treh načinih ocenjevanja parametrov zavrnili normalno, logaritemsko normalno, Gumbelovo, GEV in porazdelitev GL. Do številnih zavrnitev porazdelitvenih funkcij pri analizah časov trajanja je prišlo zaradi tega, ker je imelo zaradi uporabe dnevni vrednosti pretokov več letnih maksimumov enako trajanje vala. Ocena parametrov z metodo momentov L je pri analizah časov trajanja izkazala boljše ujemanje podatkov z izbranimi porazdelitvami kot uporaba metode momentov in metode največjega verjetja pri večini uporabljenih kombinacij porazdelitev in metod ocenjevanja parametrov. Te ugotovitve so v skladu z ugotovitvami pri analizah konic in volumnov visokovodnih valov.

S statističnimi in grafičnimi testi smo določili porazdelitvene funkcije, ki so najprimernejše za multivariatno analizo visokovodnih valov Save v Litiji s funkcijo kopula. Tako smo za robne porazdelitvene funkcije izbrali logaritemsko Pearsonovo III porazdelitev pri konicah in časih

trajanja ter Pearsonovo III porazdelitev pri volumnih. Te porazdelitvene funkcije so uporabili tudi drugi raziskovalci (Zhang in Singh, 2006; Genest in Favre, 2007; Karmakar in Simonovic, 2008). Poleg tega smo ugotovili, da uporaba metode momentov L za oceno parametrov porazdelitev pri večini uporabljenih statističnih in grafičnih testov daje boljše ujemanje posameznih porazdelitev s podatki kot metoda največjega verjetja in metoda momentov, zato metodo momentov L priporočamo za praktično uporabo, kljub nekoliko zahtevnejši uporabi.

Zahvala

Zahvaljujemo se Agenciji RS za okolje za posredovane podatke z vodomerne postaje Litija 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

- Bezak, 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). Unpublished Thesis, Univerza v Ljubljani, FGG, 106 p. (in Slovenian).
- 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).
- 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.
- Frantar, P., Hrvatina, M. (2008). "Pretočni režimi" v P. Frantar (ur.), *Vodna bilanca Slovenije 1971–2000*. MOP ARSO, Ljubljana, 43–50.
- Favre, A. C., El Adlouni, S., Perreault, L., Thiemonge, N., Bobee, B. (2004). Multivariate hydrological frequency analysis using copulas, *Water Resources Research*, 40, W01101, doi:10.1029/2003WR002456.

- 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.
- Grimaldi, S., Serinaldi, F. (2006). Asymmetric copula in multivariate flood frequency analysis, *Advances in Water Resources*, **29**, 1155–1167.
- Haan, C. T. (2002). *Statistical methods in hydrology*. Iowa State Press, 496 p.
- Haddad, K., Rahman, A. (2010). Selection of the best fit flood frequency distribution and parameter estimation procedure: a case study for Tasmania in Australia. *Stochastic Environmental Research and Risk Assessment*, **25**(3), 415–428.
- Hosking, J. R. M. (1990). L-moments: analysis and estimation of distributions using linear combinations of order statistic, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **52**(1), 105–124.
- Hosking, J. R. M., Wallis, J. R. (1997). *Regional frequency analysis: an approach based on L-moments*. Cambridge University Press, Cambridge, 224 p.
- Karmakar, S., Simonovic, S. P. (2008). Bivariate flood frequency analysis: Part 1. Determination of marginals by parametric and nonparametric techniques, *Journal Flood Risk Management*, **1**(4), 190–200.
- 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.
- Kottegoda, N. T., Rosso, R. (2008). *Applied statistics for civil and environmental engineers*. Qxford, Blackwell Publishing, 718 p.
- Kučič, K. 2007. Metoda momentov L pri analizi visokih vod (Method of L-moments for flood frequency analysis). Unpublished Thesis, Univerza v Ljubljani, FGG, 85 p. (in Slovenian).
- Maidment, D. (1993). *Handbook of hydrology*. McGraw-Hill, New York, 1424 p.
- Mikoš, M., Brilly, M., Ribičič, M. (2004). Poplave in zemeljski plazovi v Sloveniji = Floods and landslides in Slovenia, *Acta hydrotechnica*, **22**(37), 113–133.
- Pugelj, A. 2012. Analiza visokovodnih valov Save v Šentjakobu (Water wave analysis of Sava in Šentjakob). Unpublished Thesis, Univerza v Ljubljani, FGG, 98 p. (in Slovenian).
- 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.
- Ricci, V. (2005). Fitting distributions with R. <http://cran.r-project.org/doc/contrib/Ricci-distributions-en.pdf> (Pridobljeno 1. 3. 2012.).
- Salvadori, G., De Michele, C., (2004). Frequency analysis via copulas: Theory aspects and applications to hydrological events, *Water Resources Research*, **40**, W12511, doi:10.1029/2004WR003133.
- Salvadori, G., De Michele, C. (2007). On the use of Copulas in hydrology: Theory and practice, *Journal of Hydrologic Engineering*, ASCE, **12**(4), 369–380.
- Swanson, D. A., Tayman, J., Bryan, T. M. (2011). MAPE-R: a rescaled measure of accuracy for cross-sectional subnational population forecasts. *Journal of Population Research*, **28**(2-3), 225–243.
- Šraj, M., Dirnbek, L., Brilly, M. (2010). The influence of effective rainfall on modeled runoff hydrograph, *Journal of Hydrology and Hydromechanics*, **58**(1), 3–14.
- Turk, G. (2012). *Verjetnostni račun in statistika*. Ljubljana, Univerza v Ljubljani, Fakulteta za gradbeništvo in geodezijo, 264 p. (in Slovenian).
- Ulaga, F. (2011). Hidrološka postaja Litija na Savi, *Naše okolje*, **18**(8), 81–85 (in Slovenian).
- 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.