UDK: 532.562
Originalni naučni rad

NEIZVESNOSTI U ANALIZI VELIKIH VODA METODOM PARCIJALNIH SERIJA

Jasna PLAVŠIĆ
Građevinski fakultet u Beogradu

REZIME

U hidrološkoj praksi statistička analiza velikih voda zasniva se prevashodno na prilagođavanju teorijske raspodele verovatnoće osmotrenim podacima, što znači da su rezultati analize velikih voda uvek povezani sa određenim neizvesnostima. Neizvesnosti u statističkoj analizi velikih voda koja se zasniva na nizovima godišnjih maksimuma protoka su detaljno izučene i opisane su u stručnoj literaturi iz hidrološke statistike. S druge strane, to nije slučaj sa analizom velikih voda koja se zasniva na metodi pikova, odnosno na parcijalnim serijama. U ovom radu prikazana je ukratko metoda parcijalnih serija, a zatim su razmatrane neizvesnosti u primeni ove metode. Opisani su faktori koji utiču na neizvesnosti, uključujući uticaje dužine niza, vrednosti parametara i vrste metode za ocenu parametara.

Ključne reči: velike vode, statistička analiza, metoda parcijalnih serija, neizvesnost

UVOD

Statistička analiza velikih voda tradicionalno se zasniva na nizovima maksimalnih godišnjih protoka, dakle na nizovima sa po jednom vrednošću iz svake godine osmatranja. Ovaj pristup je uobičajen u hidrološkoj praksi jer ne zahteva veliku količinu podataka, a i matematički aparat koje se primenjuje je jednostavan. Nedostatak nizova godišnjih maksimuma je u tome što u njih ulazi samo po jedan protok iz svake godine, što znači da će u niz ući relativno mali protoci iz sušnih godina dok se „drugi najveći“ protok iz neke vodne godine neće naći u nizu. Alternativa nizovima godišnjih maksumuma su tzv. parcijalne serije, koje se sastoje od “pikova” koji prevazilaze određeni izabrani prag, ili prekoračenja. Smatra se da parcijalne serije sadrže više informacija o velikim vodama od nizova godišnjih maksimuma.

Statistička analiza parcijalnih serija podrazumeva nešto složeniji matematički aparat u odnosu na analizu nizova godišnjih maksimuma, s obzirom da se zapravo zasniva na slučajnim procesima. Celokupan postupak se često naziva i metoda pikova iznad praga (engl. *peaks over threshold method*), za koji je teorijske osnove postavio Todorović [13], [17]. On je formulisao slučajni proces velikih voda u kome vreme pojave prekoračenja (a time i broj prekoračenja u nekom vremenskom intervalu) i veličina prekoračenja predstavljaju slučajne promenljive. Todorović je prvobitno razvio svoj model za niz nezavisnih i jednako raspoređenih prekoračenja, ali je dao dovoljno opštu matematičku formulaciju koja se može koristiti i u složenijim slučajevima. Todorović i Zelenhasić [17] su izabrali Poasonov proces za opisivanje pojave prekoračenja i eksponencijalnu raspodelu kao raspodelu veličine prekoračenja.

Metoda pikova iznad praga je dalje razvijana u nekoliko pravaca. Todorović i Rousselle [15] su proširili model mogućnošću za uvođenje sezonskih neravnomernosti u analizu velikih voda. Ashkar i Rousselle [2], [3], [4] su se često bavili uslovima za ispunjavanje pretpostavki o Poasonovom procesu i eksponencijalnoj raspodeli. Negativnu binomnu raspodelu za broj prekoračenja, kao alternativu Poasonovoj, predložili su Calenda i sar. [5], a kasnije i Cunnane [7]. Binomna raspodela za broj prekoračenja je manje korišćena. Pored eksponencijalne raspodele za veličinu pikova, korišćena je dvoparametarska gama raspodela [20], opšta Pareto raspodela (npr. [11], [19]) i Vejbulova raspodela [18]. Todorović je sličnu metodologiju primenjivao i u oblasti padavina, procesa disperzije u rekama i u poroznoj sredini, kao i u oblasti pronosa nanosa (npr. [14], [16]).

Neizvesnostima pri oceni kvantila iz parcijalnih serija prvi se bavio Cunnane [6] za kombinaciju Poasonove i eksponencijalne raspodele, odredivši izraz za varijansu ocene kvantila. Kasnije se više autora bavilo istom temom, i uz različite pristupe manje ili veće složenosti.

O METODI PARCIJALNIH SERIJA

U slučajnom procesu koji su definisali Todorović i Zelenhasić [17] posmatraju se protoci velikih voda X_i u vremenskom intervalu (0, t] koji prevazilaze određenu baznu vrednost ili prag x_0 (slika 1). Tada veličine:

$$Z_i = X_i - x_0, \quad i = 1,2,\ldots$$

predstavljaju prekoračenja ili pikove iznad praga. Svako Z_i je slučajna promenljiva. Broj prekoračenja u intervalu (0, t] takođe je slučajna promenljiva $\eta(t)$, koja zapravo predstavlja diskretan slučajni proces. Najveće od svih prekoračenja Z_i u intervalu (0, t] je slučajni proces $\chi(t)$:

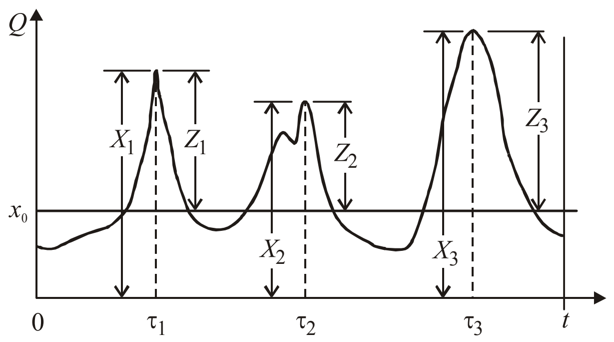
$$\chi(t) = \max\{Z_1, Z_2, \ldots, Z_{\eta(t)}\}, \quad t \geq 0$$

Drugim rečima, $\chi(t)$ je najveće prekoračenje iznad izabranog praga u proizvoljnom, ali fiksiranom, vremenskom intervalu. Pretpostavljajući da su prekoračenja nezavisna i jednako raspoređena, kao i pod pretpostavkom da su prekoračenja nezavisna od trenutka njihove pojave, raspodela najvećeg prekoračenja jednaka je [17]:

$$F(z;t) = P\{\chi(t) \leq z\} = $$
$$= P\{\eta(t) = 0\} + \sum_{n=1}^{\infty}[H(z)]^n \cdot P\{\eta(t) = n\}$$

gde je $H(z)$ funkcija raspodele prekoračenja.

Ako interval (0, t] predstavlja jednu godinu, tada je $\chi(t)$ najveće godišnje prekoračenje iznad x_0. Takođe, ako se maksimalni godišnji protok označi sa $X(t) = \chi(t) + x_0$, tada raspodela godišnjih maksimuma glasi:



Slika 1. Protoci i prekoračenja preko praga kao slučajni proces.

$$F(x) = P\{X(t) \leq x\} = P\{\chi(t) \leq x - x_0\} = $$
$$= \sum_{n=0}^{\infty}[H(x - x_0)]^n \cdot P\{\eta(t) = n\}, \quad x > x_0 \qquad (1)$$

Ova raspodela je definisana samo za $x > x_0$, odnosno za protoke veće od praga, dok za protoke ispod praga nije definisana. Vrednost funkcije raspodele godišnjih maksimuma za $x = x_0$ jednaka je verovatnoći da tokom godine nema nijednog prekoračenja preko praga x_0:

$$F(x_0) = P\{X(t) \leq x_0\} = P\{\eta(t) = 0\}$$

Iz opšteg izraza (1) mogu se dobiti izrazi za pojedine modele parcijalnih serija (pod modelima se podrazumevaju pojedine kombinacije raspodele broja prekoračenja i raspodele samih prekoračenja). U tabeli 1 prikazane su raspodele nekoliko modela parcijalnih serija.

U praksi, statistička analiza parcijalnih serija za izabrani prag odvija se kroz tri koraka: (a) prilagođavanje raspodele broja prekoračenja $p_n = P\{\eta = n\}$ (izbor između Poasonove, binomne i negativne binomne raspodele); (b) prilagođavanje raspodele prekoračenja $H(z)$ (eksponencijalna, Vejbulova i sl.), i (c) proračun raspodele godišnjih maksimuma $F(x)$ korišćenjem parametara raspodela određenih u prethodnim koracima.

Tabela 1. Funkcije raspodele godišnjih maksimuma za nekoliko najčešćih modela parcijalnih serija.

Model	Funkcija raspodele i inverzni oblik (kvantil)
Poason / ekspon. **P+E** (Λ,β)	$F(x) = \exp\{-\Lambda \exp[-(x - x_0)/\beta]\}$ $x(F) = x_0 + \beta[-\ln(-\ln F) + \ln \Lambda]$
Poason / opšta Pareto **P+GP** (Λ,k,b)	$F(x) = \exp\{-\Lambda[1 - k(x - x_0)/b]^{1/k}\}$ $x(F) = x_0 + \frac{b}{k}\left[1 - \left(\frac{-\ln F}{\Lambda}\right)^k\right]$
Poason / Vejbulova **P+W** (Λ,γ,β)	$F(x) = \exp\left\{-\Lambda \exp\left[-\left(\frac{x - x_0}{\beta}\right)^\gamma\right]\right\}$ $x(F) = x_0 + \beta[-\ln(-\ln F) + \ln \Lambda]^{1/\gamma}$
Negativna binomna / ekspon. **NB+E** (Λ,a,β)	$F(x) = \left\{1 + \frac{\Lambda}{a}\exp\left(-\frac{x - x_0}{\beta}\right)\right\}^{-a}$ $x(F) = x_0 + \beta \ln\left[\frac{\Lambda}{a(F^{-1/a} - 1)}\right]$

Ako je definisana funkcija raspodele godišnjeg maksimuma $F(x)$, tada se može koristiti klasična definicija povratnog perioda:

$$T(x) = \frac{1}{P\{X > x\}} = \frac{1}{1 - F(x)}, \quad x > x_0$$

Sa ovakvom formulacijom povratnog perioda rezultati analize parcijalnih serija su uporedivi sa rezultatima analize nizova godišnjih maksimuma. Može se pokazati [17] da ovaj izraz predstavlja prosečan broj godina u kome će se prvi put prevazići protok x. S druge strane, neki autori (npr. [9]) su skloni da pretpostave da broj prekoračenja praga predstavlja Poasonov proces i da zatim sprovedu samo drugi korak u analizi parcijalnih serija od tri gore navedena. U tom slučaju, povratni period se određuje na osnovu raspodele prekoračenja H:

$$T_{PS}(x) = \frac{1}{\lambda[1 - H(x - x_0)]}, \quad x > x_0$$

gde je λ intenzitet Poasonovog procesa, odnosno prosečan broj prekoračenja godišnje. Odnos između dve definicije povratnog perioda je dobro poznata relacija:

$$T_{PS}(x) = \frac{1}{\ln T(x) - \ln[T(x) - 1]}$$

Razlike između dve definicije povratnog perioda su značajne samo za kraće povratne periode (manje od 10 godina). Osnovna razlika je u tome što klasičan povratni period $T(x)$ ima vrednosti veće od 1 godine, dok $T_{PS}(x)$ može biti i kraći od 1 godine. Međutim, druga definicija nije pogodna za modele parcijalnih serija sa brojem prekoračenja koji ne prati Poasonovu raspodelu.

NEIZVESNOSTI U PARCIJALNIM SERIJAMA

Pokazatelji neizvesnosti

Uobičajeni zadatak statističke analize velikih voda je da se proceni merodavna velika voda, odnosno protok određene verovatnoće prevazilaženja (najčešće izražene kroz povratni period). Drugi zadatak statističke analize može biti da se proceni verovatnoća prevazilaženja neke vrednosti protoka. Ova dva zadatka statističke analize velikih voda podrazumevaju zapravo isti postupak izbora najboljeg modela parcijalnih serija i ocenu parametara odgovarajućih raspodela. Naravno, “tačna” raspodela populacije velikih voda na razmatranom profilu uvek ostaje nepoznata jer se njeni parametri jedino mogu oceniti iz uzorka. Na primer, za model parcijalnih serija sa tri parametra (a_1, a_2 i a_3) može se u opštem slučaju reći da imaju nepoznatu tačnu raspodelu $F(x; a_1, a_2, a_3)$ i raspodelu $F(x; \hat{a}_1, \hat{a}_2, \hat{a}_3)$ ocenjenu iz uzorka, kao što je pokazano na slici 2.

Iz ovih razloga, ocena kvantila $\hat{x}_T$ (ili ocena verovatnoće $\hat{p}_x$ zadate vrednosti x) pomoću izabranog modela parcijalnih serija zavisi od ocena parametara, a koji opet zavise od svojstava uzorka. Kao i svaka ocena iz uzorka, ocena kvantila $\hat{x}_T$ je slučajna promenljiva koja ima gustinu raspodele i svojstva kao što su očekivanje $E[\hat{x}_T]$ i varijansa $\mathrm{var}[\hat{x}_T]$. Ova svojstva zapravo opisuju neizvesnosti u ocenjivanju $\hat{x}_T$.

Veličine koje se koriste kao uobičajeni pokazatelji neizvesnosti su pristrasnost *B*, standardna greška *SE* i koren srednje kvadratne greške *RMSE*, koji su ovde definisani za ocenu kvantila:

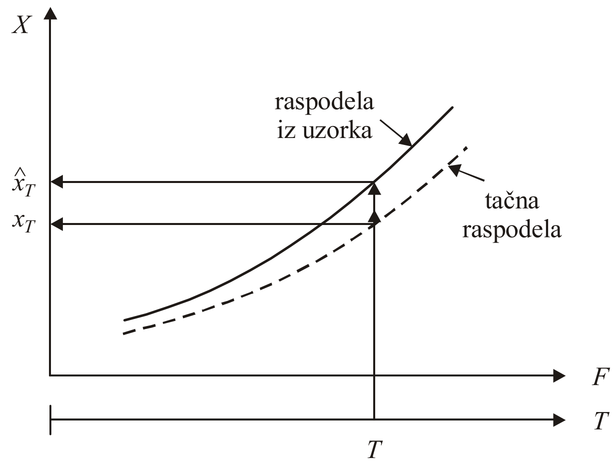
$$B[\hat{x}_T] = E[\hat{x}_T] - x_T, \quad SE[\hat{x}_T] = \sqrt{\mathrm{var}[\hat{x}_T]},$$

$$RMSE[\hat{x}_T] = \sqrt{\mathrm{var}[\hat{x}_T] + B^2[\hat{x}_T]}$$

U razmatranjima grešaka u oceni kvantila jednostavnije je posmatrati kvantil prekoračenja $\hat{z}_T$ umesto kvantila velikih voda $\hat{x}_T$, čime se eliminiše uticaj praga x_0. Time se ništa ne menja sa gledišta ovih grešaka, s obzirom da za pristrasnost i varijansu ova dva kvantila važi:

$$B[\hat{x}_T] = B[\hat{z}_T], \quad \mathrm{var}[\hat{x}_T] = \mathrm{var}[\hat{z}_T]$$

Tada je moguće posmatrati relativnu neizvesnost, kao na primer $B[\hat{z}_T]/z_T$ ili $SE[\hat{z}_T]/z_T$.



Slika 2. Odnos ocene kvantila i njegove nepoznate tačne vrednosti.

Pristupi određivanju neizvesnosti

U opštem slučaju postoji nekoliko načina da se odrede neizvesnosti. Ukoliko je poznata funkcija gustine raspodele ocene kvantila, onda je moguće da se $E[\hat{x}_T]$ ili $\mathrm{var}[\hat{x}_T]$ ocene direktno. Ashkar i Rousselle [1] su odredili izraze za gustinu raspodele ocene kvantila za model P+E, kao i Rasmussen i Rosbjerg [9] koji su na osnovu te gustine došli i do $E[\hat{x}_T]$ i $\mathrm{var}[\hat{x}_T]$. Međutim, za druge modele parcijalnih serija praktično je nemoguće da se ta funkcija gustine odredi. Uobičajeni pristup se zasniva na razvoju ocene kvantila $\hat{x}_T$ u Tejlorov red u kome se zanemaruju izvodi višeg reda, dok se ocena kvantila tretira kao funkcija ocena parametara, kao što je $\hat{x}_T = g(\hat{a}_1, \hat{a}_2, \hat{a}_3)$. Ovaj pristup daje dobro poznate izraze za varijansu i očekivanje funkcija slučajnih promenljivih, koje za slučaj ocene kvantila glase:

$$\mathrm{var}[\hat{x}_T] \approx \sum_i \left(\frac{\partial g}{\partial \hat{a}_i}\right)^2 \mathrm{var}[\hat{a}_i] + 2\sum_{i \neq j} \frac{\partial g}{\partial \hat{a}_i}\frac{\partial g}{\partial \hat{a}_j} \mathrm{cov}[\hat{a}_i, \hat{a}_j]$$

$$E[\hat{x}_T] \approx g + \frac{1}{2}\sum_i \frac{\partial^2 g}{\partial \hat{a}_i^2} \mathrm{var}[\hat{a}_i] + \sum_{i \neq j} \frac{\partial^2 g}{\partial \hat{a}_i \partial \hat{a}_j} \mathrm{cov}[\hat{a}_i, \hat{a}_j]$$

gde se funkcija g i njeni izvodi računaju za vrednosti $\mu_{\hat{a}_1}, \mu_{\hat{a}_2}, \mu_{\hat{a}_3}$. Ovaj pristup je korišćen u radovima [6] i [10] za model P+E, i u radu [11] za model P+GP. Prethodni izrazi se nazivaju i asimptotskim izrazima jer opisuju svojstva ocena iz velikih uzoraka.

Takođe je moguće da se očekivanje i varijansa ocene kvantila odrede simulacijama metodom Monte Karlo. Prednost te metode je što omogućava da se odrede neizvesnosti i za male dužine nizova. Međutim, ne može se očekivati da se ovaj pristup primenjuje u svakodnevnoj praksi. Monte Karlo simulacije su korišćene u radu [12] za model P+E i u radu [11] za model P+GP.

Ako je zadatak da se oceni verovatnoća pojave zadatog protoka x, pristupi za određivanje neizvesnosti su isti kao i za ocenu kvantila. Naravno, svejedno je kojom veličinom se opisuje rizik od velikih voda, jer se radi o jednoznačno povezanim veličinama. To može biti sama funkcija raspodele F_x, verovatnoća neprevazilaženja $p_x = 1 - F_x$ ili povratni period $T_x = 1 / p_x$; ponekad se koristi i rizik u užem smislu reči, $R_x = 1 - F^L$, koji se definiše kao verovatnoća prevazilaženja protoka x tokom L godina. U radu [1] korišćen je direktan pristup i određena je funkcija gustine raspodele za ocenu povratnog perioda $\hat{T}_x$ u modelu P+E, a u radu [9] isti pristup korišćen je za ocenu povratnog perioda parcijalnih serija $\hat{T}_{PS,x}$ i ocenu rizika $\hat{R}_x$. Međutim, očekivanje i varijansa mogu se odrediti samo za $\hat{R}_x$, ali ne za $\hat{T}_x$ ili $\hat{T}_{PS,x}$. U radu [9] korišćen je i pristup sa razvojem u Tejlorov red, a njegovi autori nazvali su rezultate tog pristupa aproksimativnim formulama. Određivanje gustine raspodele za druge modele parcijalnih serija praktično je nemoguće.

Rezultati određivanja neizvesnosti za sve modele parcijalnih serija iz tabele 1 prikazani su detaljno u radu [8], kako za ocene kvantila, tako i za ocene verovatnoće. Za modele sa dvoparametarskim raspodelama prekoračenja (Vejbulova i opšta Pareto) razmatrane su dve najčešće korišćene metode za ocenu parametara, a to su metode običnih i verovatnosnih težinskih momenata. U tom radu izvršeno je sistematsko poređenje rezultata različitih autora za model P+E. Neizvesnosti u oceni kvantila u modelu P+GP su prilagođene prema rezultatima iz [11]. Neizvesnosti u oceni verovatnoće za model P+GP, kao i neizvesnosti u oceni kvantila i verovatnoće u ostalim modelima (P+W, NB+E i B+E) određene su pomoću pristupa sa razvojem u Tejlorov red.

U radu [8] mogu se naći izrazi za proračun svih pomenutih neizvesnosti, dok će u narednim odeljcima biti prikazani glavni zaključci tog rada ilustrovani na nekoliko primera hidroloških nizova na stanicama u Srbiji.

Faktori koji utiču na neizvesnosti

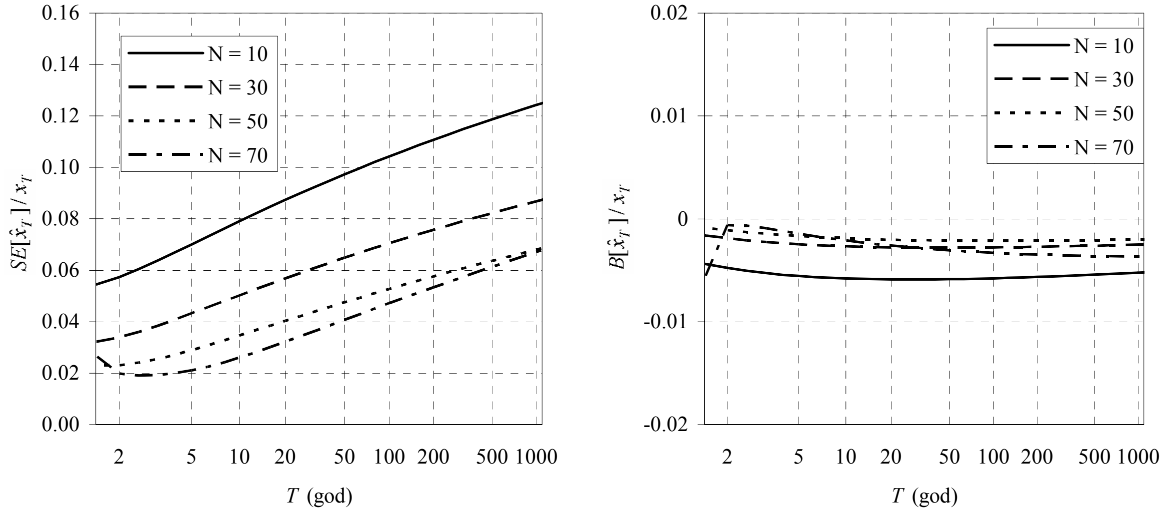
Uticaj dužine niza

Neizvesnosti su, naravno, najveće kada se statistička analiza sprovodi sa najkraćim nizovima. Kada su u pitanju parcijalne serije, dužina niza se ne ogleda samo kroz broj godina osmatranja N, već i kroz ukupan broj prekoračenja M ili kroz prosečan godišnji broj prekoračenja $\Lambda = M/N$. U opštem slučaju, za neizvesnosti u oceni kvantila i verovatnoće po svim modelima parcijalnih serija može se reći da se smanjuju sa povećanjem kako ukupne dužine niza M, tako i N i Λ.

Da bi se pokazao efekat dužine niza na neizvesnosti, izabrana je stanica Bezdan na Dunavu sa 70 godina hidroloških osmatranja. Od tako dugačkog niza formirani su kraći nizovi (od 10, 30 i 50 godina) kao što je pokazano u tabeli 2. Pri tome je za mali broj godina osmatranja uzet veliki broj podataka godišnje Λ, i obrnuto za veći broj godina. Na slici 3 prikazane su relativna standardna greška i relativna pristranost ocene kvantila za model P+W. Može se jasno videti da izuzetno kratki nizovi kao što je 10 ili 30 godina daju znatno veće neizvesnosti, čak i sa velikim Λ. To pokazuje da je pogrešno verovanje da metoda parcijalnih serija može da "kompenzuje" mali broj godina osmatranja velikim prosečnim brojem prekoračenja, pa je u tom smislu ne treba zloupotrebljavati.

Tabela 2. Nizovi protoka sa stanice Bezdan na Dunavu korišćeni za analizu uticaja dužine niza.

Niz	*N*	Λ
1961-1970	10	5.9
1951-1980	30	4.2
1941-1990	50	2.5
1931-2001	70	1.04



Slika 3. Relativna standardna greška (levo) i relativna pristrasnost (desno) ocene kvantila po modelu P+W za nizove različite dužine na stanici Bezdan na Dunavu.

Uticaj parametara oblika raspodela prekoračenja

Na neizvesnosti modela parcijalnih serija koje uključuju dvoparametarske raspodele za visinu prekoračenja najviše utiče vrednost parametra oblika te raspodele (parametar *k* opšte Pareto raspodele i parametar γ Vejbulove raspodele; videti tabelu 1). Ovaj zaključak je logičan, s obzirom da i same ove raspodele mogu biti veoma različite u zavisnosti od parametra oblika. Na primer, opšta Pareto raspodela je ograničena sa gornje strane za pozitivne vrednosti njenog parametra oblika, dok za negativne vrednosti nije organičena. U okolini nekih vrednosti ovih parametara neizvesnosti mogu biti veoma velike. Poznavanje neizvesnosti u zavisnosti od vrednosti parametara raspodela može biti korisno u postupku izbora raspodele pri modeliranju nekog niza.

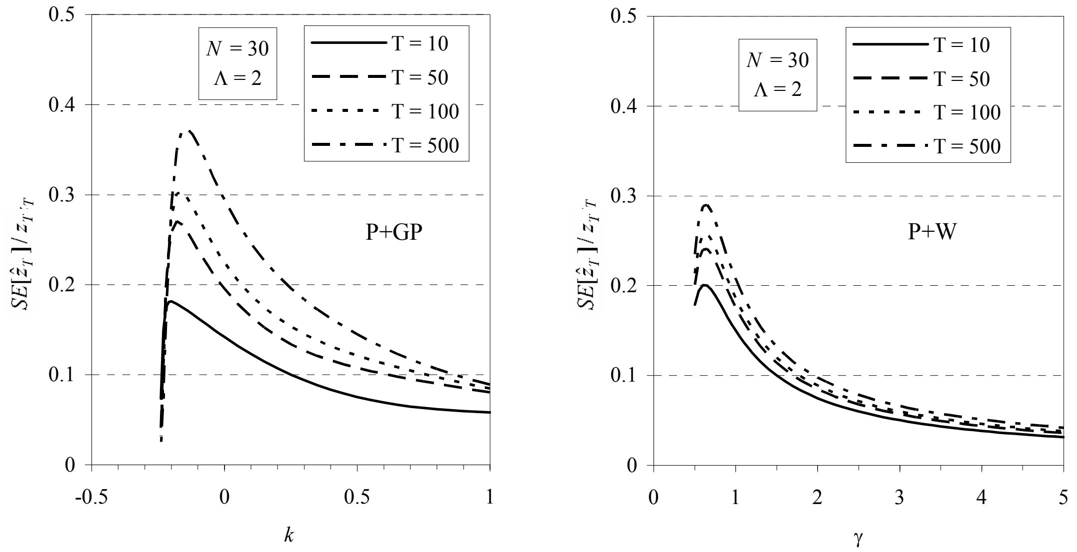
Na slici 4 prikazan je primer kako parametri oblika raspodela prekoračenja utiču na neizvesnosti. Ova slika predstavlja teorijske rezultate za relativnu standardnu grešku kvantila modela P+GP i P+W za uobičajenu dužinu niza (N = 30 godina i Λ = 2). Može se videti da su neizvesnosti prilično velike kada je parametar *k* opšte Pareto raspodele manji od 0, i kada je parametar γ Vejbulove raspodele manji od 1. Ukoliko želimo male neizvesnosti, preporučljivo je da se ove dve raspodele sa takvim vrednostima parametara izbegavaju.

Uticaj metode ocene parametara

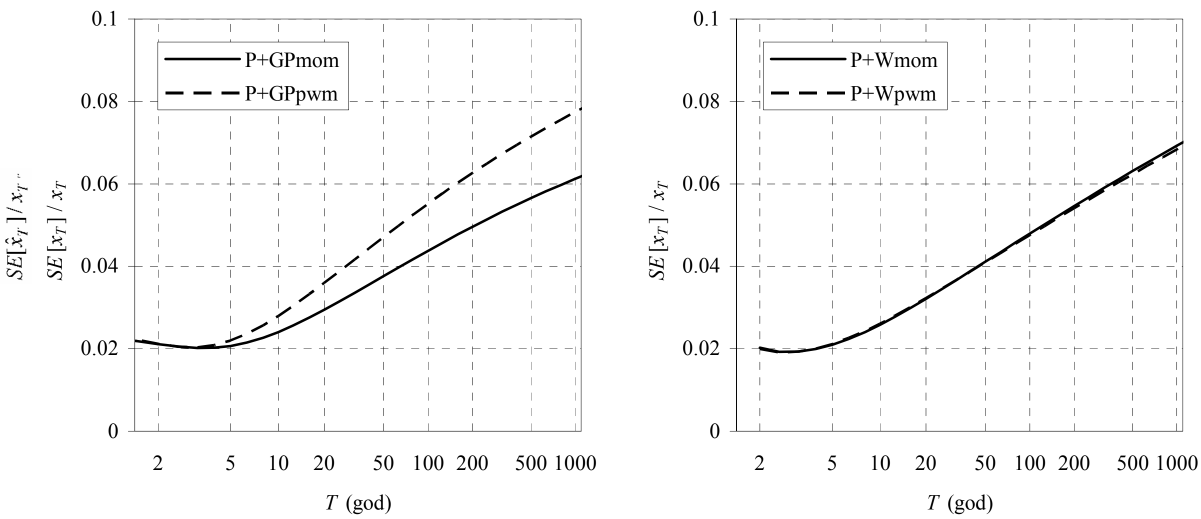
Metoda ocene parametara kod dvoparametarskih raspodela prekoračenja nije bez značaja za neizvesnosti. Kada su u pitanju Vejbulova i opšta Pareto raspodela, najčešće primenjivane metode za ocenu parametara su metoda momenata i metoda verovatnosnih težinskih momenata. Nema posebnih pravilnosti u načinu uticaja izabrane metode na neizvesnosti. Ponekad ove dve metode daju značajno različite neizvesnosti, a nekad praktično iste, kao što je pokazano na slici 5.

Kod opšte Pareto raspodele metoda momenata nije primenljiva za vrednost parametra $k \leq -0.5$; s druge strane, ocene parametara su u domenu $-0.5 < k < -0.2$ opterećene velikom neizvesnošću dok se za $k \leq -0.25$ čak ne mogu ni odrediti asimptotskim formulama. Sa gledišta neizvesnosti, metoda momenata ima prednost nad metodom težinskih momenata kada je $k > -0.2$ u najvećem broju slučajeva (za razumne dužine nizova). Metoda težinskih momenata ima prednost za vrednosti $k \leq -0.2$.

Kod Vejbulove raspodele uticaj izbora metode je u principu nešto manji. Pristranost ocene kvantila kod metode momenata naglo raste u odnosu na drugu metodu u domenu $0.5 < \gamma < 1$. Relativna standardna greška ocene kvantila ima veoma slične vrednosti za obe metode ocene parametara.



Slika 4. Relativna standardna greška ocene kvantila modela P+GP (levo) i P+W (desno) u zavisnosti od vrednosti parametara oblika i za nekoliko povratnih perioda T.

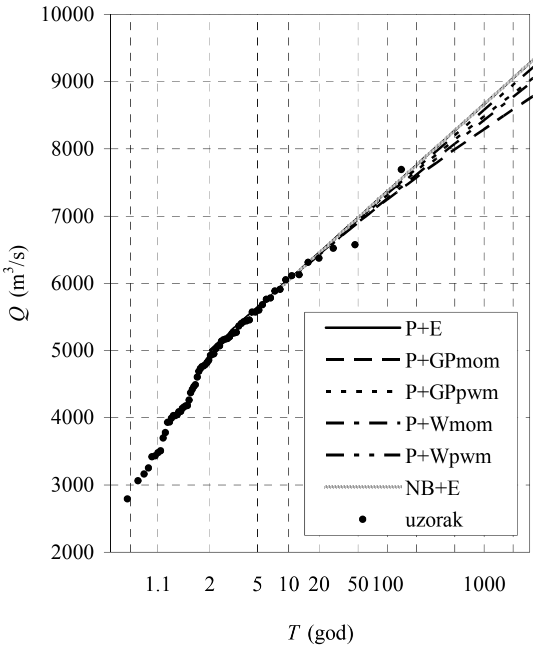


Slika 5. Relativna standardna greška ocene kvantila za dve metode ocene parametara na stanici Bezdan na Dunavu u modelima P+GP (levo) i P+W (desno); mom: metoda momenata; pwm: metoda težinskih momenata.

Neizvesnosti kao dodatni kriterijum za izbor raspodele verovatnoće

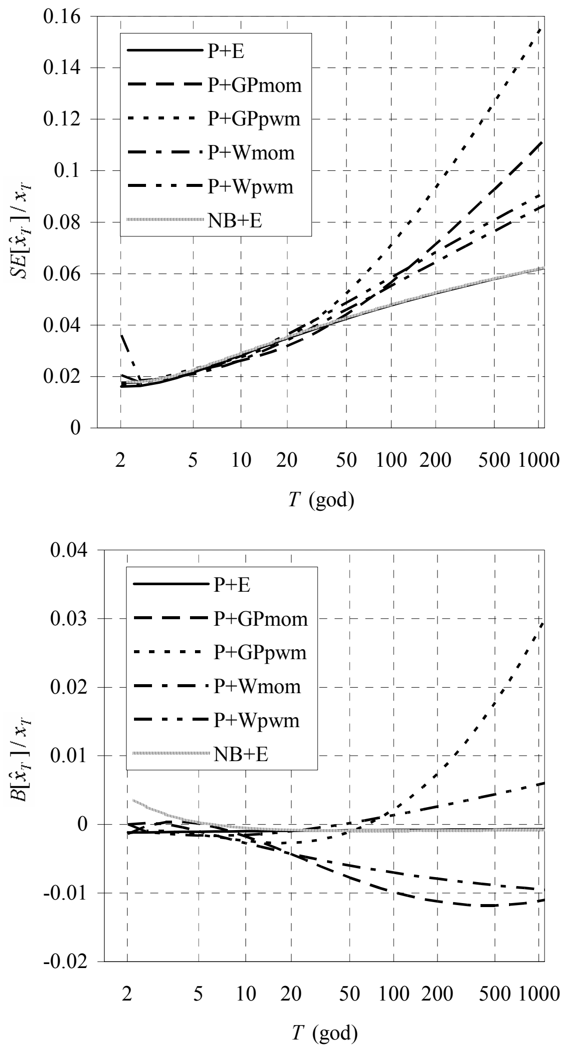
Neizvesnosti mogu biti dobar dodatni kriterijum za izbor najpogodnijeg modela parcijalni serija. U nastavku će biti prikazano nekoliko primera gde neizvesnosti mogu da pomognu pri takvom odlučivanju.

Kod nekih nizova, najčešće kod onih bez izuzetaka i sa malom asimetrijom, dešava se da sve pretpostavljene raspodele daju slične rezultate tako da ni klasični testovi saglasnosti ne pomažu pri izboru raspodele. U primeru niza protoka na stanici Bezdan na Dunavu, za vrednost praga od 5000 m^3/s svi primenjeni modeli parcijalnih serija se skoro preklapaju na dijagramu verovatnoće (slika 6).

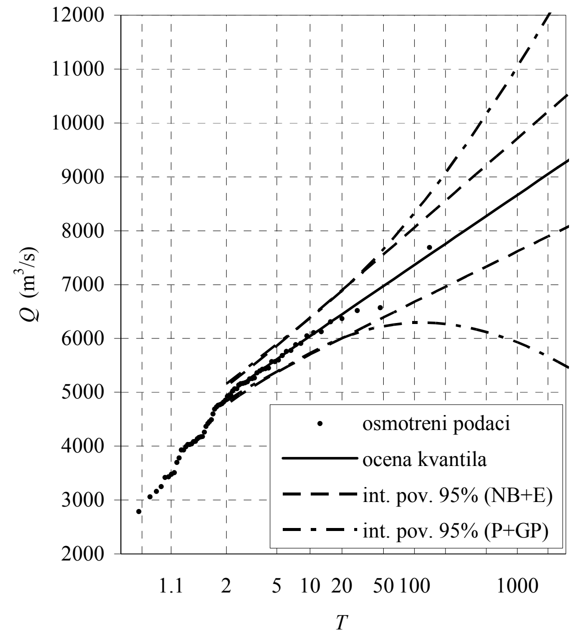


Slika 6. Raspodele velikih voda za stanicu Bezdan na Dunavu za bazu od 5000 m^3/s prema različitim modelima parcijalnih serija.

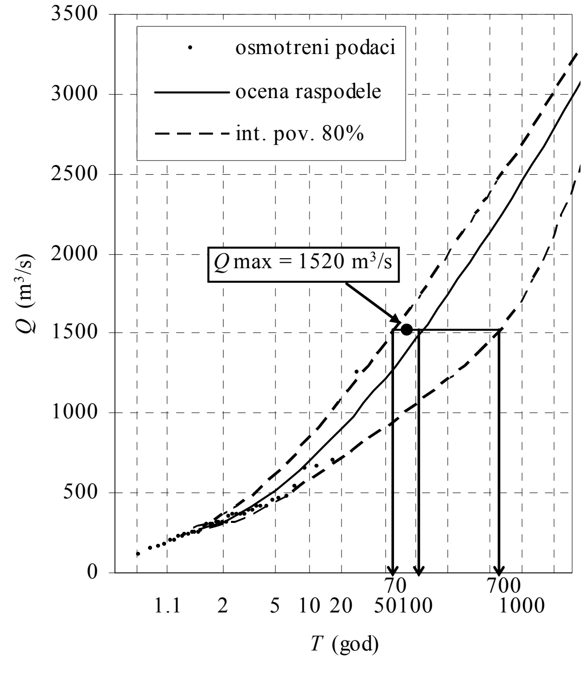
Međutim, ako se pogledaju neizvesnosti u oceni kvantila za sve ove modele na slici 7, tada se može uočiti da se one veoma razlikuju za različite modele. Na primer, relativna standardna greška ocene 1000-godišnje velike vode kreće se od 6% do 16%. Na dijagramima na slici 7 neizvesnosti kvantila u modelima P+E i NB+E praktično se preklapaju i imaju najmanje vrednosti, dok ostali modeli imaju veće neizvesnosti u domenu većih povratnih perioda. U ovakvom slučaju, kada drugi kriterijumi za izbor najbolje raspodele ne pomažu, može se izabrati ona raspodela odnosno model sa najmanjom neizvesnošću (u ovom primeru to je model P+E). Izborom modela sa najmanjom standardnom greškom dobija se najuži interval poverenja (slika 8). Ovaj primer takođe pokazuje da nema potrebe za uvođenjem složenijih raspodela u slučaju kada se prekoračenja dobro modeliraju eksponencijalnom raspodelom.



Slika 7. Relativna standardna greška (gore) i relativna pristrasnost (dole) ocene kvantila prema različitim modelima parcijalnih serija za stanicu Bezdan na Dunavu za bazu od 5000 m^3/s.



Slika 8. Intervali poverenja od 95% za velike vode na stanici Bezdan na Dunavu za bazu od 5000 m^3/s, za modele sa najmanjom i najvećom neizvesnošću.



Slika 9. Interval poverenja od 80% za verovatnoću fiksiranog protoka na stanici Lopatnica Lakat na Ibru (model P+W za bazu od 280 m^3/s); za najveći osmotreni protok od 1520 m^3/s ovaj interval poverenja iznosi od 70 do 700 godina.

Neizvesnosti u oceni verovatnoće

Kada su u pitanju neizvesnosti u oceni verovatnoće pojave neke osmotrene velike vode, opšti zaključak je da su one veće što su protoci veći. Preciznije rečeno, neizvesnosti postaju veoma velike za protoke koji bi prema tačnoj raspodeli odgovarali povratnim periodima većim od broja godina osmatranja.

Na slici 9 prikazan je primer raspodele velikih voda na stanici Lopatnica Lakat na Ibru, na kojoj je osmotren najveći protok od 1520 m^3/s. Na dijagramu je prikazan i interval poverenja od 80% za ocenu verovatnoće. Sa slike se može videti da povratni period najvećeg osmotrenog protoka prema izabranoj raspodeli iznosi oko 100 godina, dok interval poverenja ukazuje da stvarni povratni period ovog protoka iznosi negde između 70 i 700 godina. Iz ovog primera se može zaključiti da je ocena povratnog perioda najvećih voda, naročito izvan opsega osmotrenih vrednosti, veoma nezahvalan zadatak.

ZAKLJUČCI

Opšti zaključak o neizvesnostima u modelima parcijalnih serija je teško formirati, s obzirom da na ove neizvesnosti utiče veliki broj faktora. Ono što se najlakše uočava za sve modele jeste da su neizvesnosti najveće za najkraće nizove, i to ne samo za mali ukupan broj prekoračenja, već i za mali broj godina osmatranja sa velikim prosečnim godišnjim brojem prekoračenja. To pokazuje da metodu parcijalnih serija ne treba zloupotrebljavati na kratkim nizovima (ovakva tendencija se može sresti i u radovima viđenijih autora). Na neizvesnosti utiču i vrednosti parametara oblika raspodela prekoračenja, kao i metoda za ocenu parametara. Neizvesnosti su takođe dobar dodatni kriterijum pri izboru najpogodnije raspodele odnosno modela parcijalnih serija.

Analiza velikih voda je samo jedan korak u hidrotehničkim i vodoprivrednim studijama i projektima koji se bave procenom rizika i šteta, planiranjem sistema i projektovanjem objekata i mera za zaštitu od poplava i drugih srodnih nepogoda kao što su erozija i klizišta. Ona treba da obezbedi neophodne merodavne veličine za projektovanje ili ocenu verovatnoće pojave kritičnih događaja. Stoga je jasno da projektovana rešenja u velikoj meri zavise od rezultata analize velikih voda. U analizi rizika od velikih voda neizvesnosti su uvek postojale i dalje će postojati. One se neminovno prenose i na dalje rezultate, kao što su dimenzije objekata ili procenjene štete. Ipak, u hidrološkoj i hidrotehničkoj praksi preovlađuje tendencija da se postojanje neizvesnosti zanemaruje, i to tako što se ocene velikih voda smatraju kao izvesne (pa čak i kao deterministički podaci) ili tako što se sve neizvesnosti – one koje potiču od samog fizičkog procesa i one koje proističu iz njegovog modeliranja – kompenzuju npr. usvajanjem većeg projektnog povratnog perioda. Iako se neizvesnosti ne mogu izbeći, njihovo razmatranje je važno kako bi se omogućilo da se sagledaju ograničenja projektovanih objekata ili da se, uzimanjem neizvesnosti u obzir, postignu bolja tehnička rešenja i doprinese kvalitetnijem procesu odlučivanja o prihvatljivom riziku u zaštiti od velikih voda.

LITERATURA

[1] Ashkar, F. i Rousselle, J.: Design discharge as a random variable: a risk study, *Water Resour. Res.*, 17(3): 577-591, 1981.

[2] Ashkar, F. i Rousselle, J.: Some remarks on the truncation used in partial flood series models, *Water Resour. Res.*, 19(2): 477-480, 1983.

[3] Ashkar, F. i Rousselle, J.: The effect of certain restrictions imposed on the interrarival times of flood events on the Poisson distribution used for modeling flood counts, *Water Resour. Res.*, 19(2): 481-485, 1983.

[4] Ashkar, F. i Rousselle, J.: Partial duration series modeling under the assumption of a Poissonian flood count, *J. Hydrology*, 90: 135-144, 1987.

[5] Calenda, G., Petaccia, A. i Togna, A.: Theoretical probability distribution of critical hydrologic events by the partial-duration series method, *J. Hydrology*, 33: 233-245, 1977.

[6] Cunnane, C.: A particular comparison of annual maxima and partial duration series methods of flood frequency prediction, *J. Hydrology*, 18: 257-271, 1973.

[7] Cunnane, C.: A note on the Poisson assumption in partial duration series models, *Water Resour. Res.*, 15(2): 489-494, 1979.

[8] Plavšić, J.: *Analiza rizika od velikih voda pomoću prekidnih slučajnih procesa*, doktorska disertacija, Građevinski fakultet u Beogradu, 2005.

[9] Rasmussen, P.F. i Rosbjerg, D.: Risk estimation in partial duration series, *Water Resour. Res.*, 25: 2319-2330, 1989.

[10] Rosbjerg, D.: Estimation in partial duration series with independent and dependent peak values, *J. Hydrology*, 76: 183-195, 1985.

[11] Rosbjerg, D., Madsen, H. i Rasmussen, P.F.: Prediction in partial duration series with generlized Pareto-distributed exceedances, *Water Resour. Res.*, 28: 3001-3010, 1992.

[12] Tavares, L.V. i Da Silva, J.E.: Partial duration series method revisited, *J. Hydrology*, 64: 1-14, 1983.

[13] Todorović, P.: On some problems involving random number of random variables, *Ann. Math. Stat.*, 41(3): 1059-1063, 1970.

[14] Todorović, P.: A stochastic model of dspersion of sediment particles released from a continuous source, *Water Resour. Res.*, 11(6): 919-925, 1975.

[15] Todorovic, P. i Rousselle, J.: Some problems of flood analysis, *Water Resour. Res.*, 7(5): 1144-1150, 1971.

[16] Todorović, P. i Yevjevich, V.: Stochastic processes of precipitation, *Hydrology Paper* No. 35, Colorado State University, Fort Collins, 1969.

[17] Todorović, P. i Zelenhasić, E.: A stochastic model for flood analysis, *Water Resour. Res.*, 6(6): 1641-1648, 1970.

[18] Vukmirović, V. i Petrović, J.: Flood flow analysis using renewal processes, UNESCO-IHP V *Tech. Documents in Hydrology* No. 11 (Annual FRIEND AMHY meeting, Thessaloniki, 1995), 159-169, 1997.

[19] Wang, Q.J.: The POT model described by the generalized Pareto distribution with Poisson arrival rate, *J. Hydrology*, 129: 263-280, 1991.

[20] Zelenhasić, E.: Theoretical probability distributions for flood peaks, *Hydrology Paper* No. 42, Colorado State University, Fort Collins, 1970.

UNCERTAINTIES IN FLOOD FREQUENCY ESTIMATION BY PARTIAL DURATION SERIES

by

Jasna PLAVŠIĆ
Faculty of Civil Engineering, Belgrade

Summary

In hydrologic practice, flood frequency analysis is usually based on fitting a theoretical distribution to observed data series. Flood frequency analysis in hydrologic practice is essentially based on fitting of a theoretical probability distribution to observed flood data, meaning that the results of the flood frequency analysis is always associated with some amount of uncertainty. For flood frequency analysis that is based on series of annual maximum floods, uncertainties have been investigated in detail and can be readily found in literature on hydrologic statistics. On the other hand, that is not the case with uncertainties in flood frequency analysis based on partial duration series (or peaks over threshold method). The paper describes briefly the peaks over threshold method, and then discusses uncertainties in its application. Uncertainties in quantile estimates and in probability estimates for fixed flood value are considered. Factors affecting these uncertainties are described, including sample size, shape parameters of distributions for exceedances, and parameter estimation method. Several examples showing uncertainties in flood series on Serbian rivers are presented.

Key words: floods, frequency analysis, partial duration series, peaks over threshold method, uncertainty

Redigovano 31.08.2006.