

Pirsonov sistem raspodela i njegova primena u hidrologiji

Marija Radičević¹
Vesna Jevremović²
Jasna Plavšić³
Dragutin Pavlović⁴

APSTRAKT: Nalaženje raspodele koja pravilno prati često kompleksno ponašanje podataka je značajan problem u statistici jer je potrebno naći raspodelu koja će zadovoljiti širok raspon asimetrije i spljoštenosti. Pirsonov sistem raspodela odgovara podacima koji su vidno asimetrični. Ovaj sistem obuhvata 12 neprekidnih raspodela koji proističu iz zajedničke opšte diferencijalne jednačine sa četiri parametra, dok se pojedini tipovi dobijaju za određene karakteristične vrednosti prarametara. U ovom radu prikazane su teorijske osnove Pirsonovog sistema raspodela, date su praktične preporuke za njegovu primenu i prikazani su primeri primene u statističkoj analizi velikih voda u hidrologiji na većem broju hidroloških stanica u Srbiji. U ovoj analizi akcenat je stavljen na poređenje rezultata dobijenih iz opšteg Pirsonovog sistema i primenom raspodele tipa III koja se standardno koristi u hidrološkoj praksi.

Ključne reči: Pirsonov sistem raspodela, statistička analiza, velike vode, maksimalni godišnji protoci, Pirsonova raspodela III tipa

The Pearson distribution system and its application in hydrology

ABSTRACT: Finding a distribution that adequately describes a complex data behaviour is an important problem in statistics because it is necessary to find a distribution applicable to a range of skews and kurtoses. The Pearson distribution system is suitable for highly skewed data. This system comprises 12 continuous distributions that originate from a joint general differential equation with four parameters, while the individual types are obtained for some characteristic parameter values. This paper presents the theoretical basis of the Pearson distribution system, gives recommendations for its application and also presents the examples of its application in statistical analysis of floods in hydrology at a number of hydrologic stations in Serbia. The emphasis is put on comparison of the results obtained from the general Pearson system to the results of the standard procedure in hydrology in which the type III distribution is often applied.

Keywords: Pearson distribution system, statistical analysis, floods, annual maximum floods, Pearson type III distribution

¹ Marija Radičević, diplomirani matematičar, student master studija, Univerzitet u Beogradu – Matematički fakultet, marijar@matf.bg.ac.rs

² Dr Vesna Jevremović, v. prof., Univerzitet u Beogradu – Matematički fakultet, v_jevremovic@matf.bg.ac.rs

³ Dr Jasna Plavšić, v. prof., Univerzitet u Beogradu – Građevinski fakultet, jplavsic@grf.bg.ac.rs

⁴ Dr Dragutin Pavlović, docent, Univerzitet u Beogradu – Građevinski fakultet, epavlovd@hikom.grf.bg.ac.rs

1 Uvod

Značajan problem u statistici i u oblastima koje je koriste kao alat jeste nalaženje raspodele koja na prihvatljiv način prati karakteristike podataka. Još krajem XIX veka bilo je poznato kako modelirati posmatrane podatke tako da im odgovaraju prva dva teorijska momenta neke raspodele, ali pretpostavljajući simetriju te raspodele. Mnogi naučnici su došli do zaključka da raspodela nije dovoljno objašnjena na taj način. Međutim, nije bilo poznato kako konstruisati raspodelu verovatnoća u kojoj su asimetrija i spljoštenost netipični. Potrebno je bilo naći raspodelu koja će zadovoljiti širok raspon osobina. Sa ovim problemom se istraživači i danas sreću – kako u praktičnim, tako i u teorijskim razmatranjima.

Mnoge asimetrične raspodele i raspodele sa netipičnom spljoštenošću bile su poznate još u XIX veku, samo pod drugačijim nazivima nego danas. Pirsonov sistem raspodela upravo odgovara podacima koji su vidno asimetrični. Ovaj sistem raspodela predstavlja sistem od 12 neprekidnih raspodela. Prvi radovi u vezi sa ovim sistemom su objavljeni od strane Karla Pirsona⁵ 1895. godine i tada su uvedena 4 tipa ovih raspodela kao dopuna normalnoj raspodeli, koja je prvobitno numerisana brojem V. Ovaj sistem je proširen između 1901. i 1916. godine u mnogobrojnim radovima vezanim za biostatistiku.

U prvobitnom radu, Pirson je tipove klasifikovao na osnovu oblasti definisanosti raspodela, u zavisnosti od toga da li je raspodela definisana na ograničenom intervalu, pozitivnom delu realne prave ili na celoj realnoj pravoj. U drugom radu, 1901. godine, Pirson je predefinisao tip V, koji je prvobitno obuhvatao samo normalnu raspodelu, i uveo još jedan tip raspodele u sistem. Nakon toga, Pirson uvodi specijalne tipove i podtipove raspodela čime je dobijeno ukupno 12 tipova raspodela.

U ovom radu prikazane su teorijske osnove Pirsonovog sistema raspodela, date su praktične preporuke za njegovu primenu i prikazani su primeri primene u statističkoj analizi velikih voda u hidrologiji, na podacima sa izabranih hidrometrijskih stanica u republici Srbiji.

2 Pirsonov sistem raspodela

Pirsonov sistem je parametarska familija raspodela čija se funkcija gustine može lako izraziti u eksplicitnom obliku. Osobine ovog sistema čine ga pogodnim i za teorijske i za empirijske probleme. Gustine svih tipova raspodela Pirsonovog sistema se dobijaju kao rešenje diferencijalne jednačine:

$$\frac{p'(x)}{p(x)} = \frac{x - a}{b_0 + b_1x + b_2x^2}, \quad (1)$$

gde je $p(x)$ funkcija gustine, a a, b_0, b_1 i b_2 parametri raspodele. Oblik raspodele zavisi od vrednosti parametara.

Iz formule (1) transformacijama i integraljenjem mogu se dobiti jednačine koje važe za sve tipove raspodela iz Pirsonovog sistema. Te jednačine povezuju momente i parametre raspodele. Odatle sledi da se mogu odrediti momenti nižih redova u zavisnosti od parametara a, b_0, b_1, b_2 raspodele i momenata $m_0 = 1$ i m_1 , ali takođe, mogu se izraziti i parametri raspo-

⁵ Carl Pearson, 1857-1936, engleski matematičar i biostatističar

dele u odnosu na momente raspodele. Pošto se centralni momenti⁶ mogu izraziti preko momenta raspodele i pošto je $\mu_1 = 0$, parametre raspodele možemo izraziti i preko centralnih momenta raspodele. Zamenom $n = 0, 1, 2, 3$ u tako dobijenu jednačinu i sređivanjem dobijenog sistema, dobija se da su parametri Pirsonovog sistema u direktnoj vezi sa centralnim momentima (μ_2, μ_3, μ_4) raspodele na sledeći način:

$$b_1 = a = -\frac{\sqrt{\mu_2\beta_1(\beta_2+3)}}{10\beta_2-12\beta_1-18}, \quad b_0 = -\frac{\mu_2(4\beta_2-3\beta_1)}{10\beta_2-12\beta_1-18}, \quad b_2 = -\frac{2\beta_2-3\beta_1-6}{10\beta_2-12\beta_1-18} \quad (2)$$

gde su $\beta_1 = \frac{\mu_3^2}{\mu_2^2}$ i $\beta_2 = \frac{\mu_4}{\mu_2^2}$ kvadrat koeficijenta asimetrije i koeficijent spljoštenosti, respektivno⁷. Ova veza se može koristiti u konkretnim situacijama.

2.1 Tipovi Pirsonove raspodele

Oblik Pirsonove raspodele zavisi od vrednosti parametara a, b_0, b_1, b_2 . Ako a nije rešenje jednačine $b_0 + b_1x + b_2x^2 = 0$ i ako je $x = a$, gustina je konačna i $p'(x)$ je nula. Ako je $x \neq a$ i $p(x) \neq 0$, onda je i $p'(x) \neq 0$.

Pošto je $p(x)$ gustina raspodele, onda moraju biti zadovoljeni uslovi da je $p(x) \geq 0$ i $\int_{-\infty}^{+\infty} p(x)dx = 1$. Iz (1) sledi da $p(x) \rightarrow 0$ kada $x \rightarrow \pm\infty$, pa i $p'(x) \rightarrow 0$ kad $x \rightarrow \pm\infty$. Ukoliko prethodno ne važi, onda nije zadovoljeno da je $p(x) \geq 0$ i potrebno je ograničiti skup vrednosti x -a na one gde je $p(x) > 0$, a inače definisati da je $p(x) = 0$.

Rešenje jednačine (1) zavisi od prirode korena jednačine $b_0 + b_1x + b_2x^2 = 0$, a različiti koreni odgovaraju različitim tipovima raspodele.

Ako je $b_1 = b_2 = 0$, onda se iz jednačine (1) i iz odgovarajućih uslova za gustinu raspodele, dobija da je u pitanju normalna ili Gausova raspodela sa očekivanjem a i disperzijom $-b_0$. Normalna raspodela je granična raspodela svih tipova Pirsonovog sistema raspodele.

U nastavku će biti prikazano svih 12 tipova Pirsonovog sistema, sa formulama za njihovu gustinu i nekim osnovnim karakteristikama. Formule za gustine koje će biti prikazane se dobijaju rešavanjem jednačine (1) i sređivanjem tako dobijenog izraza uvođenjem novih oznaka radi kraćeg zapisa.

Tip I. Pirsonovoj raspodeli tipa I odgovaraju realni koreni jednačine $b_0 + b_1x + b_2x^2 = 0$ koji su različitog znaka⁸. Rešavanjem jednačine (1) u tom slučaju i korišćenjem osobina gustine raspodele, dobija se da je gustina Pirsonove raspodele tipa I:

$$p(x) = \frac{\alpha^{m_1}\beta^{m_2}}{(\alpha + \beta)^{m_1+m_2+1}B(m_1 + 1, m_2 + 1)} \left(1 + \frac{x}{\alpha}\right)^{m_1} \left(1 - \frac{x}{\beta}\right)^{m_2}, \quad -\alpha \leq x \leq \beta. \quad (3)$$

Ako su m_1 i m_2 strogo pozitivni, gustina raspodele je unimodalna⁹ i uzima vrednost 0 na krajevima oblasti definisanosti. Ako je jedan od m -ova između 0 i 1, kriva zaklapa oštar

⁶ Ako je X slučajna veličina, njen n -ti centralni momenat je $\mu_n = E(X - E(X))^n, (\forall n)$.

⁷ Pri ovako navedenim formulama koeficijenta asimetrije i koeficijenta spljoštenosti, vrednosti ovih koeficijenata za normalnu raspodelu su redom 0 i 3.

⁸ Koreni navedene jednačine se najčešće označavaju sa α i β .

ugao sa x -osom. Ako je tačno jedno m manje od 0, onda kriva ima beskonačnu ordinatu i ima oblik slova J. Ako su oba m -a manja od nule, onda kriva ima oblik slova U. Ovaj tip raspodele predstavlja generalizaciju beta raspodele.

Tip II. Pirsonova raspodela tipa II je specijalan slučaj raspodele tipa I. Ovaj tip raspodele se dobija kada se u jednačini (1) pretpostavi da je parametar $b_1 = 0$, a parametri b_0 i b_2 su različitog znaka. Na sličan način kao u prethodnom slučaju dobija se da je funkcija gustine Pirsonove raspodele tipa II:

$$p(x) = \frac{1}{\alpha B\left(m+1, \frac{1}{2}\right)} \left(1 - \frac{x^2}{\alpha^2}\right)^m, \quad -\alpha \leq x \leq \alpha. \quad (4)$$

Ova raspodela je simetrična. Ako je $m > 0$ raspodela je unimodalna, a ako je $m < 0$ onda raspodela ima U oblik. Ako je $m = 0$, dobija se gustina koja odgovara uniformnoj raspodeli na segmentu $[-\alpha, \alpha]$.

Raspodela tipa II je korišćena u formiranju tabele Spirmanovog koeficijenta korelacije rangova¹⁰ kada je broj elemenata serije manji od 100.

Tip III. Ako je u jednačini (1) $b_2 = 0$ (i $b_1 \neq 0$), dobija se Pirsonova raspodela tipa III. Funkcija gustine raspodele tipa III je:

$$p(x) = \frac{\beta^{m+1} \alpha^m}{e^{\alpha\beta} \Gamma(m+1)} \left(1 + \frac{x}{\alpha}\right)^m e^{-\beta x}, \quad (5)$$

za $x \in (-\alpha, +\infty)$ ako je $\alpha > 0$, odnosno za $x \in (-\infty, -\alpha)$ ako je $\alpha < 0$.

Raspodela je unimodalna ako je $m > 0$, a ima J oblik ako je $m < 0$. Gama raspodela spada u Pirsonovu raspodelu tipa III.

Tip IV. Ako je diskriminanta jednačine $b_0 + b_1x + b_2x^2 = 0$ negativna, onda ta jednačina nema realne korene. Jedan od takvih slučajeva odgovara Pirsonovoj raspodeli tipa IV. Ovaj tip raspodele se najčešće koristi u sledećem obliku:

$$p(x) = K \left(1 + \frac{x^2}{\omega^2}\right)^m e^{-c \cdot \arctan \frac{x}{\omega}}, \quad x \in \mathbb{R} \quad (6)$$

gde je K normirajuća konstanta koja se dobija iz uslova da je $p(x)$ gustina raspodele. U opštem slučaju teško je odrediti konstantu K numeričkim putem. Oblast definisanosti ove raspodele je cela realna prava i raspodela je unimodalna. Jednačina (6) dovodi do komplikovanog računa ako se pokuša izračunavanje funkcije raspodele. Najlakše je u jednačini (2) teorijske momente zameniti stvarnim momentima iz konkretnih podataka i tako dobiti parametre raspodele.

⁹ Unimodalna raspodela je raspodela koja ima tačno jedan maksimum.

¹⁰ Ako postoji značajno odstupanje raspodele od normalne ili ako podaci nisu numerički, onda se ne može primeniti Pirsonov koeficijent korelacije. Tada se primenjuje Spirmanov koeficijent korelacije koji se bazira na rangovima podataka.

17. Savetovanje SDHI i SDH - Vršac, Srbija 2015. Conference SDHI & SDH - Vršac, Serbia 2015.

Tip V. Pirsonova raspodela tipa V se dobija kada su koreni jednačine $b_0 + b_1x + b_2x^2 = 0$ međusobno jednaki, tj. kada je $b_1^2 = 4b_0b_2$. Gustina Pirsonove raspodele tipa V je:

$$p(x) = \frac{(ma)^{m-1}}{\Gamma(m-1)} x^{-m} e^{-\frac{ma}{x}}, 0 \leq x < +\infty. \quad (7)$$

Slično se dobija za $-\infty < x \leq 0$. Ova raspodela je unimodalna.

Tip VI. Ako su koreni jednačine $b_0 + b_1x + b_2x^2 = 0$ realni i istog znaka onda je u pitanju Pirsonova raspodela tipa VI. Gustina Pirsonove raspodele tipa VI je:

$$p(x) = \frac{\omega^{m_1-m_2-1}}{B(m_1-m_2-1, m_2-1)} x^{-m_1} (x-\omega)^{m_2}, x \in (\omega, +\infty). \quad (8)$$

Slično se dobija za $x \in (-\infty, \omega)$. Jednostavnom transformacijom $y = \frac{a}{x}$ se iz ove raspodele dobija Pirsonova raspodela tipa I. Ako je $m_2 > 0$, onda je raspodela unimodalna. Ako je $m_2 < 0$, onda raspodela ima J oblik.

Tip VII. Ako su parametri u jednačini (1) takvi da $b_1 = a = 0$, $b_0 > 0$ i $b_2 > 0$, onda je u pitanju Pirsonova raspodela tipa VII. Ovaj tip raspodele se razlikuje od raspodele tipa II jer su kod nje parametri b_0 i b_2 bili različitog znaka. Gustina raspodele tipa VII je:

$$p(x) = \frac{1}{\alpha B\left(\beta - \frac{1}{2}, \frac{1}{2}\right)} \left(1 + \frac{x^2}{\alpha^2}\right)^{-\beta}, x \in R. \quad (9)$$

Ovom tipu raspodela pripada Studentova t-raspodela. Kada je $\alpha = n^2$, a $\beta = \frac{n+1}{2}$, dobija se Studentova raspodela sa n stepeni slobode.

Tip VIII. Ako su u jednačini (1) parametri takvi da važi da je $b_0 = 0$ i $b_1 > 0$ dobija se Pirsonova raspodela tipa VIII. Funkcija gustine Pirsonove raspodele tipa VIII je:

$$p(x) = \frac{(-\alpha)^{\frac{a}{b_1}-1}}{B\left(1 - \frac{a}{b_1}, 2 + \frac{a}{\alpha}\right)} \left(1 + \frac{x}{\alpha}\right)^{1+\frac{a}{\alpha}}, 0 \leq x \leq \alpha. \quad (10)$$

Tip IX. Ako se u jednačini (1) zameni da je $b_0 = 0$ i $b_1 < 0$, dobija se raspodela tipa IX. Jednačina (1) će imati isti oblik kao i kod prethodnog tipa raspodele (imati u vidu da je znak parametra b_1 suprotan u odnosu na prethodni slučaj). Samim tim i funkcija gustine će biti istog oblika.

Tip X. Ako su u jednačini (1) b_1 i b_2 jednaki nuli, dobija se da je gustina raspodele:

$$p(x) = \frac{1}{\omega^{1+\frac{a}{\omega}} \Gamma\left(1 + \frac{a}{\omega}\right)} x^{\frac{a}{\omega}} e^{-\frac{x}{\omega}}, 0 \leq x < +\infty. \quad (11)$$

Tip XI. Ako su u jednačini (1) b_0 i b_1 jednaki nuli, dobija se Pirosonova raspodela tipa XI čija je gustina:

$$p(x) = (\omega - 1)b^{\omega-1}x^{-\omega}, b \leq x < +\infty, \quad (12)$$

gde je $b > 0$ parametar položaja i zavisi od konkretnih podataka.

Tip XII. Raspodela tipa XII je specijalan slučaj raspodele tipa I, kada je $10\beta_2 - 12\beta_1 - 18$. Tada se ne mogu koristiti izrazi koje povezuju momente i parametre raspodele. Jednačina (3) se može napisati na sledeći način:

$$p(x) = \frac{\left(\frac{\alpha}{\beta}\right)^m}{(\alpha + \beta)B(m + 1, 1 - m)} \left(\frac{1 + \frac{x}{\alpha}}{1 - \frac{x}{\beta}}\right)^m, -\alpha \leq x \leq \beta, |m| < 1.$$

3 Praktična primena Pirsonove raspodele

Za praktičnu primenu Pirsonove raspodele od koristi može da bude dijagram odnosa koeficijenata asimetrije i spljoštenosti dat na slici 1, koji prikazuje oblasti koje odgovaraju određenim tipovima Pirsonove raspodele. Prikazano je samo prvih sedam raspodela pošto preostalih pet raspodela predstavljaju podtipove ovih raspodela. Može se primetiti da Pirsonovim raspodelama tipa I, IV i VI odgovaraju oblasti, dok ostalim raspodelama odgovaraju linije na dijagramu i one se zbog toga nazivaju prelaznim tipovima (na x -osi je kvadrat koeficijenta asimetrije da bi slika bila preglednija).

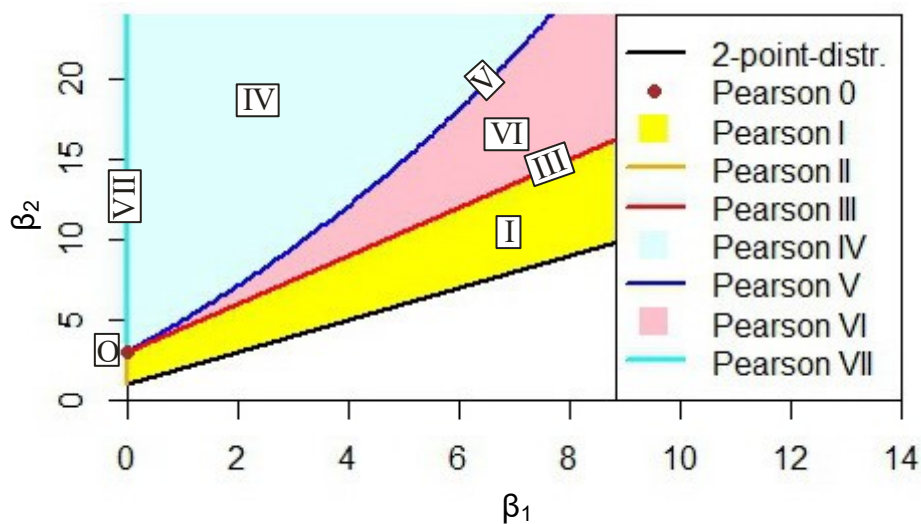
Da bi se nekim podacima dodelio neki od tipova Pirsonove raspodele, potrebno je proći kroz sledeće korake:

- 1) određivanje numeričkih vrednosti prva 4 momenta posmatranih podataka,
- 2) određivanje numeričkih vrednosti parametara Pirsonove raspodele koristeći vezu između momenata i parametara raspodele u jednačini (2),
- 3) koristeći dobijene parametre (i/ili dijagram na slici 1) odrediti kojem tipu Pirsonove raspodele podaci pripadaju,
- 4) zamenom vrednosti parametara u jednačini konkretne raspodele dobija se tražena raspodela.

Korisne informacije prilikom povezivanja podataka sa Pirsonovim sistemom:

- Tipovi Pirsonovih raspodela se spajaju u geometrijskom smislu. Na primer, tip V može biti posmatran kao prelaz između tipova IV i VI;
- Podaci moraju imati konačni momenti bar četvrtog reda;
- Kada se ne može jasno odrediti koji tip raspodele odgovara podacima potrebno je primeniti neku transformaciju momenata, kao što je na primer Šepardova korekcija;

- Često se integral Pirsonove krive ne može rešiti numeričkim putem, pa se tada primenjuju aproksimacije.



Slika 1. Oblasti primene pojedinih tipova Pirsonove raspodele

Radi lakšeg rada sa Pirsonovim sistemom raspodela u konkretnim situacijama mogu se koristiti razni programski paketi koji imaju ugrađene funkcije za rad sa ovim sistemom. U ovom radu korišćeno je programsko okruženje *R* i paket *PearsonDS* za rad sa Pirsonovim sistemom raspodela. Ovaj paket omogućava implementaciju Pirsonovog sistema raspodela, uključujući gustinu, funkciju raspodele, kvantile i slučajni uzorak iz određene raspodele, kao i prilagođavanje metodom momenata i metodom maksimalne verodostojnosti za prvih 7 tipova Pirsonovog sistema.

Za razmatrane podatke, potrebno je najpre testirati njihovu slučajnost¹¹. Postoji veliki broj testova za testiranje slučajnosti. Trebalo bi na podatke primeniti više od jednog testa, a primenjeni testovi bi trebalo da se baziraju na različitim svojstvima uzorka. Ako su *p*-vrednosti svih testova velike, znači da se prihvata nulta hipoteza o slučajnosti uzorka.

Kod Pirsonovog sistema raspodela parametri raspodele se određuju na osnovu prva četiri momenta: srednje vrednosti, varijanse, koeficijenta asimetrije i koeficijenta spljoštenosti. Koristeći funkciju *empMoments()* iz paketa *PearsonDS* mogu se izračunati empirijski momenti u konkretnom primeru. Primenom dobijenih momenata, metodom momenata može se odrediti tip i parametri Pirsonove raspodele uz pomoć funkcije *pearsonFitM()*.

Drugi način za određivanje parametara raspodele je metodom maksimalne verodostojnosti. Funkcija *pearsonMSC()* iz paketa *PearsonDS* vrši procenu parametara svih tipova Pirsonove raspodele numeričkom optimizacijom. Pored računanja logaritma funkcije verodostojnosti, ova funkcija daje i vrednosti drugih kriterijuma selekcije koji koriste logaritam funkcije verodostojnosti, kao i odgovarajuće parametre za svaki tip Pirsonove raspodele. Poslednji deo izlaza ove funkcije prikazuje koji tip raspodele je najbolji prema svakom od kriteriju-

¹¹ Ako je u pitanju prost slučajni uzorak redosled elemenata se može menjati, a ako uzorak nije slučajni, bitan je redosled elemenata u uzorku i u pitanju je vremenska serija.

ma¹². Kada su poznati tip raspodele i njeni parametri može se odrediti funkcija gustine, izrazi za funkciju raspodele i proračun kvantila kao i crtati odgovarajući dijagrami verovatnoće.

Na kraju je potrebno testirati saglasnost empirijske raspodele i dobijene raspodele. U ovom radu primenjen je test Kolmogorov-Smirnova, kojim se utvrđuje da li se raspodela populacije iz koje potiče uzorak (podaci) može opisati teorijskom raspodelom sa prethodno dobijenim parametrima.

4 Analiza velikih voda u Srbiji pomoću Pirsonovog sistema

Hidrološka praksa u analizi velikih voda na našim prostorima preporučuje upotrebu Gumbelove, log-normalne i log-Pirson III raspodele za analizu maksimalnih godišnjih protoka. Osvrt na ovu praksu i preporuke najnovijeg datuma dali su Blagojević i sar. (2014). Cilj ovog rada je da se ispita mogućnost primene Pirsonovog sistema raspodela za analizu velikih voda u Republici Srbiji.

Imajući u vidu osnovna svojstva Pirsonovog sistema, očekuje se da se njegova četvoro-parametarska forma može prilagoditi raznovrsnim skupovima podataka. Podaci koji se koriste moraju ispuniti pretpostavke o slučajnosti i homogenosti. Pored toga, u hidrologiji se smatra da raspodela najvećih ekstremnih vrednosti ne bi trebala da bude ograničena sa gornje strane. U Pirsonovom sistemu raspodele tipa III, IV, V i VI nisu ograničene sa gornje strane (za pozitivne vrednosti promenljive i pozitivne parametre razmere).

4.1 Podaci i metodologija

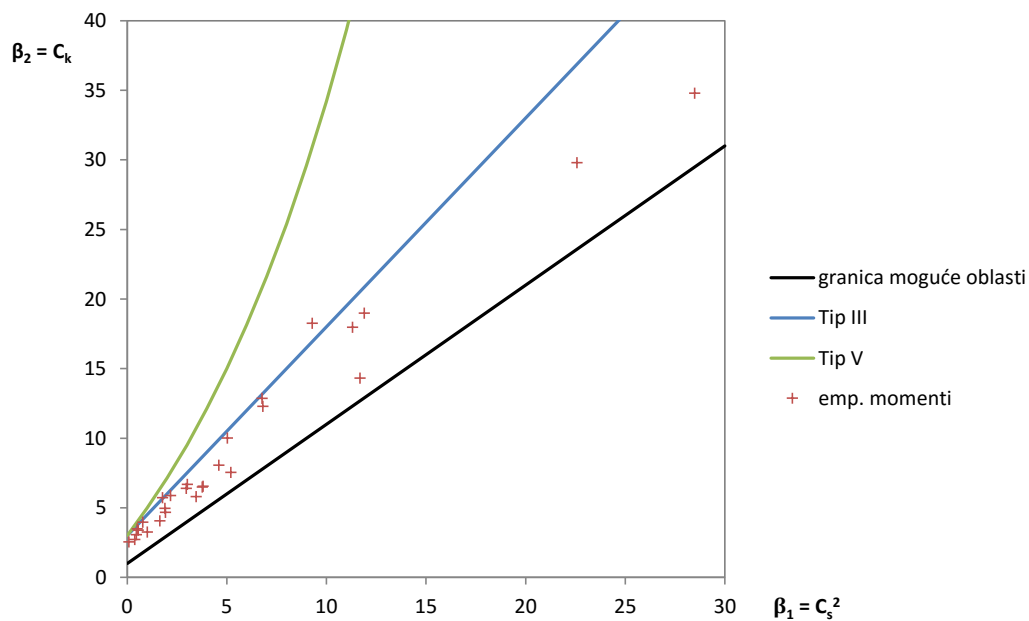
Početni skup podataka čine nizovi maksimalnih godišnjih protoka sa 67 hidrometrijskih stanica u periodu do 2012. godine, koje je obezbedio Republički hidrometeorološki zavod Srbije. Za sve stanice je testirana slučajnost nizova testovima Wold-Wolfovitz i testom tačaka zaokreta. Neparametarskim testovima Mann-Whitney i Mann-Kendall testirana je homogenost i prisustvo monotonog trenda. Sve stanice za koje su nulte hipoteze o slučajnosti i homogenosti odbačene na pragu značajnosti 5% izuzete su iz dalje analize, tako da je za analizu ostala ukupno 28 stanica. Izabrane stanice su nabrojane u tabeli 1, gde su date i vrednosti prva četiri empirijska momenta nizova osmotrenih podataka.

Proračun parametara i izbor tipa Pirsonove raspodele sproveden je pomoću funkcija iz paketa *PearsonDS* u programskom okruženju R, kao što je opisano u prethodnom odeljku. U ovom paketu tip Pirsonove raspodele se bira na osnovu vrednosti empirijskih koeficijenta asimetrije i spljoštenosti, odnosno na osnovu oblasti pojedinih tipova raspodela na slici 1. Na slici 2 prikazan je isti dijagram odnosa ovih momenata za razmatrane nizove maksimalnih godišnjih protoka. Na slici se može videti da najveći broj stanica pripada oblasti tipa I, dok samo dve stanice pripadaju oblasti tipa VI. Podseća se da je tip III granični tip između tipova I i VI.

¹² Kao alternativa može se koristiti funkcija *pearsonFitML()* koja kao rezultat daje samo tip i parametre raspodele prema kriterijumu maksimalne verodostojnosti.

Tabela 1. Odabrane hidrološke stanice sa statistikama (prva četiri uzoračka momenta).

Reka/Stаница	Površina sliva (km ²)	Empirijski momenti			
		\bar{x}	C_v	C_s	C_k
Dunav/Bezdan	210250	4935	0.24	0.75	3.38
Dunav/Bogojevo	251593	5543	0.23	0.71	3.49
Tisa/Senta	141715	2185	0.27	0.69	3.07
Lim/Brodarevo	2762	471	0.36	1.33	5.72
Lim/Prijepolje	3160	485	0.40	1.48	5.87
Beli Timok/Vratarnica	1771	142	0.55	0.89	3.96
Kolubara/Slovac	995	175	0.41	0.32	2.55
Moravica/Arilje	830	124	0.65	1.74	6.68
Skrapč/Požega	630	116	0.88	2.24	10.01
Veliki Rzav/Arilje	564	100	0.48	1.38	4.96
Kosanica/Visoka	370	70.3	0.83	2.60	12.85
Vlasina/Vlasotince	879	152	0.85	2.61	12.27
Sava/S. Mitrovica	87996	4147	0.18	0.63	2.71
Kolubara/Valjevo	340	87.8	0.70	1.39	4.67
Obnica/Belo Polje	185	59.7	0.61	1.01	3.26
Visočica/Bračevci	227	40.6	0.65	1.28	4.05
Bjelica/Guča	239	70.9	0.67	1.72	6.38
Rasina/Brus	213	31.7	0.74	1.95	6.54
Dičina/Brđani	208	45.9	0.81	2.15	8.05
Belica/Jagodina	193	19.6	0.93	1.94	6.48
Jablanica/Sedlari	140	41.9	0.89	1.86	5.79
Crnajka/Crnajka	96	29.2	1.14	2.28	7.54
Studenica/Ušće	540	59.4	0.65	3.36	17.96
Moravica/Ivanjica	475	76.5	0.79	3.45	18.97
Toplica/Magovo	180	29.7	1.23	3.42	14.30
Vlasina/Svođe	350	54.4	1.45	5.34	34.77
Lužnica/Svođe	319	61.0	1.12	4.75	29.78
Z. Morava/K. Stena	3077	310	0.50	3.05	18.24



Slika 2. Dijagram odnosa momenata C_s^2 i C_k za izbor tipa raspodele iz Pirsonovog sistema.

U korišćenim R funkcijama, parametri izabranog tipa teorijske Pirsonove raspodele određuju se metodom momenata na osnovu sva prva četiri empirijska momenta pomoću sistema jednačina (2).

Da bi se uporedio ovaj pristup usvajanja tipa i parametara raspodele sa standardnim pristupom u hidrologiji koji podrazumeva primenu tipa III, sprovedene su tri varijante proračuna teorijske raspodele za svaki niz:

- 1) Izbor tipa raspodele je prepušten softveru, odnosno tip se bira prema oblastima na dijagramu na slici 2 u kojoj se nađu empirijski momenti. Parametri izabranog tipa raspodele se ocenjuju metodom momenta na osnovu prva četiri empirijska momenta i sistema (2). Ovako dobijena raspodela označena je sa Px_mom4 .
- 2) Pirsonova raspodela III tipa sa parametrima ocenjenim metodom momenata na osnovu prva četiri empirijska momenta i sistema (2). Oznaka za ovu raspodelu je $P3_mom4$.
- 3) Pirsonova raspodela III tipa sa parametrima određenim metodom momenata na „klasičan“ način, odnosno iz prvih tri empirijska momenta (npr. Rao i Hamed, 2000). Oznaka za ovu raspodelu je $P3_mom3$.

4.2 Rezultati

U tabeli 2 prikazani su, pored empirijskih koeficijenata asimetrije i spljoštenosti, tip Pirsonove raspodele prema vrednostima empirijskih momenata i rezultati (p -vrednosti) testiranja saglasnosti empirijske i teorijskih raspodela testom Kolmogorov-Smirnova (KS). Velika p -vrednost ukazuje da nema razloga za odbacivanje nulte hipoteze o saglasnosti, dok za p -vrednosti ispod praga značajnosti α nultu hipotezu treba odbaciti. Uobičajeni prag značajnosti je $\alpha = 0,05$. Oznaka „NA“ znači da softver nije mogao da nađe rešenje sistema (2) i odredi parametre Pirsonove raspodele III tipa iz prva četiri momenta.

Prema vrednostima empirijskih momenata, većini nizova odgovara tip I jer su ovi momenti u zoni ispod prave linije koja odgovara tipu III (slika 2), dok je tip VI dodeljen samo dvema stanicama.

Rezultati u tabeli 2 su grupisani u više grupa po stepenu slaganja teorijskih raspodela sa empirijskom. Pripadnost grupi je data u poslednjoj koloni tabele 2. Na slici 3 prikazane su p -vrednosti KS testa u zavisnosti od koeficijenta asimetrije, gde su stanice razvrstane u grupe. Dijagrami zavisnosti p -vrednosti KS testa od koeficijenta spljoštenosti izgledaju veoma slično.

Rezultati upućuju na više zaključaka. Najpre se uočava da se primenom metode momenata kroz rešavanje sistema jednačina (2) uvek dobijaju bolji rezultati ako se usvoji tip raspodele prema oblastima na slici 1 (prvi pristup) nego ako se pretpostavi tip III (drugi pristup). Međutim, tip III čiji se parametri ocene iz prva tri empirijska momenta (treći pristup) je superiorniji u odnosu na rezultate koji se zasnivaju na sva četiri empirijska momenta. Razlog verovatno leži u manje pouzdanoj oceni četvrtog empirijskog momenta zbog pristrasnosti i veće uzoračke disperzije u kraćim nizovima.

U svim pristupima najbolja slaganja se postižu za podatke sa umerenom asimetrijom tj. za koeficijente asimetrije $C_s < 2$. Primena klasičnog pristupa sa parametrima ocenjenim iz prva tri momenta na podatke sa $C_s < 2$ uvek daje dobre rezultate, dok za ocene iz sistema jednačina (2) to nije pravilo (stanice u grupi 4). Na slici 4 prikazana su dva primera rezultata za ovu grupu stanica. Na primeru stanice Jagodina na Belici na levom dijagramu može se videti da prvi pristup daje Pirsonovu raspodelu tipa I sa relativno niskom gornjom granicom raspo-

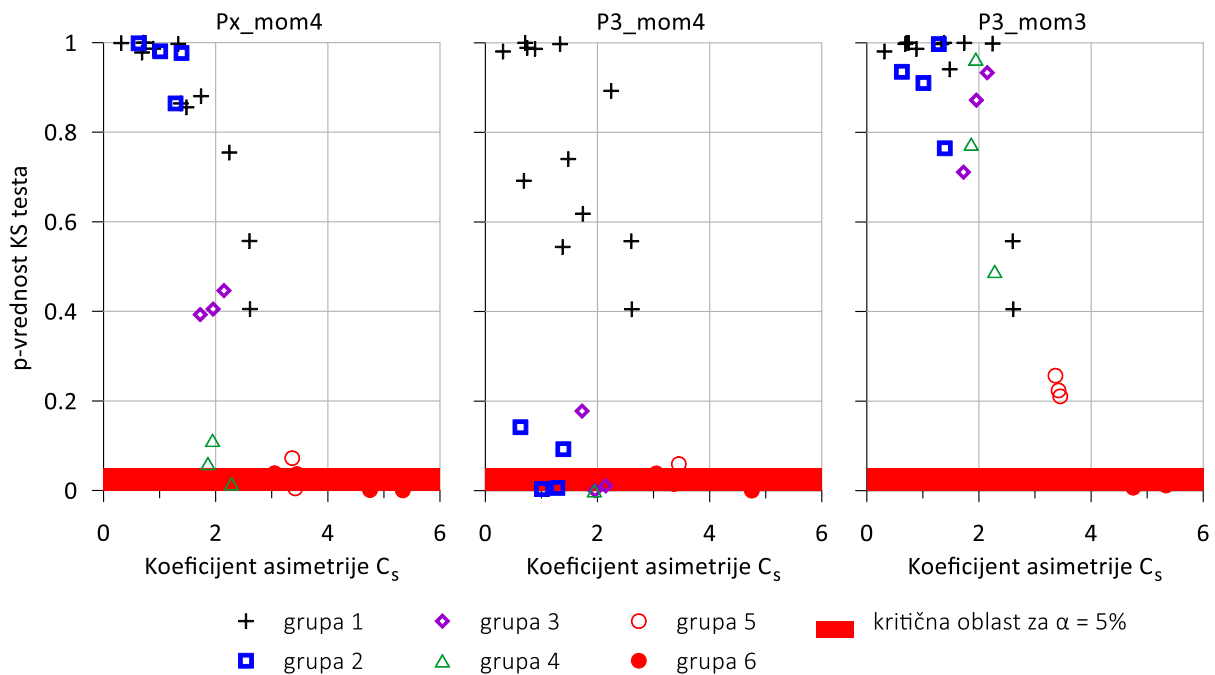
17. Savetovanje SDHI i SDH - Vršac, Srbija 2015.

Conference SDHI & SDH - Vršac, Serbia 2015.

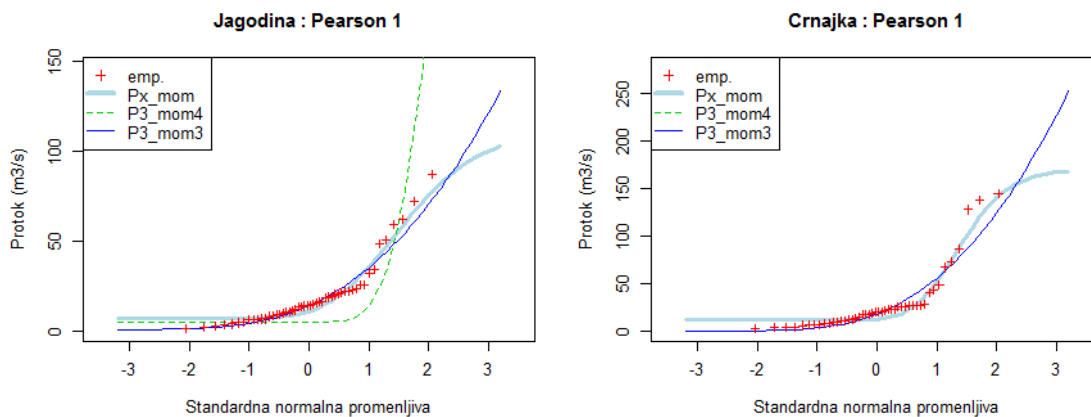
dele, čime se dobija funkcija raspodele oblika slova „S“ i oborenog gornjeg kraja. Takva raspodela nije pogodna za određivanje merodavnih velikih voda. Istovremeno, treći pristup daje razumno slaganje podataka i raspodele III tipa ocenjenjene iz prva tri momenta. Na ovoj slici se takođe može videti da su parametri raspodele III tipa po drugom pristupu nerealni u slučaju stanice Jagodina, dok u slučaju Crnajke nije postojalo rešenje sistema jednačina (2). Slični zaključci važe i za stanice iz grupe 3, kod kojih je slaganje sa tipom I u prvom pristupu bolje nego kod stanica iz grupe 4, kod kojih se po KS testu odbacuje hipoteza o saglasnosti raspodele.

Tabela 2. Rezultati (p -vrednosti) testa saglasnosti Kolmogorov-Smirnova (KS) za tri teorijske Pirsonove raspodele (Px_mom4 : izabrani tip sa ocenama parametara iz sistema jednačina (2), $P3_mom4$: tip III sa parametrima ocenjenim iz sistema jednačina (2), i $P3_mom3$: tip III sa parametrima ocenjenim iz prva tri momenta).

Stanica	C_s	C_k	Tip	p -vrednosti KS testa			Grupa
				Px_mom4	$P3_mom4$	$P3_mom3$	
Dunav/Bezdan	0.75	3.38	1	1.000	0.989	1.000	1
Dunav/Bogojevo	0.71	3.49	1	1.000	1.000	1.000	1
Tisa/Senta	0.69	3.07	1	0.978	0.692	0.998	1
Lim/Brodarevo	1.33	5.72	6	0.997	0.997	0.997	1
Lim/Prijepolje	1.48	5.87	1	0.856	0.740	0.941	1
Beli Timok/Vratarnica	0.89	3.96	1	0.987	0.987	0.987	1
Kolubara/Slovac	0.32	2.55	1	0.999	0.981	0.981	1
Moravica/Arilje	1.74	6.68	1	0.880	0.618	1.000	1
Skrapež/Požega	2.24	10.01	1	0.755	0.893	0.998	1
Veliki Rzav/Arilje	1.38	4.96	1	0.864	0.544	1.000	1
Kosanica/Visoka	2.60	12.85	1	0.557	0.557	0.557	1
Vlasina/Vlasotince	2.61	12.27	1	0.405	0.405	0.405	1
Sava/S. Mitrovica	0.63	2.71	1	0.999	0.142	0.935	2
Kolubara/Valjevo	1.39	4.67	1	0.977	0.093	0.765	2
Obnica/Belo Polje	1.01	3.26	1	0.981	0.004	0.910	2
Visočica/Bračevci	1.28	4.05	1	0.864	0.006	0.997	2
Bjelica/Guča	1.72	6.38	1	0.393	0.178	0.711	3
Rasina/Brus	1.95	6.54	1	0.405	0.000	0.872	3
Dičina/Brđani	2.15	8.05	1	0.446	0.011	0.933	3
Belica/Jagodina	1.94	6.48	1	0.112	0.000	0.964	4
Jablanica/Sedlari	1.86	5.79	1	0.060	NA	0.774	4
Crnajka/Crnajka	2.28	7.54	1	0.015	NA	0.490	4
Studenica/Ušće	3.36	17.96	1	0.072	0.015	0.257	5
Moravica/Ivanjica	3.45	18.97	1	0.037	0.059	0.211	5
Toplica/Magovo	3.42	14.30	1	0.005	NA	0.224	5
Vlasina/Svođe	5.34	34.77	1	0.000	NA	0.011	6
Lužnica/Svođe	4.75	29.78	1	0.001	0.000	0.007	6
Z. Morava/K. Stena	3.05	18.24	6	0.038	0.038	0.024	6



Slika 3. Saglasnost empirijske sa teorijskom raspodelom (p -vrednost KS testa) u zavisnosti od koeficijenta asimetrije.



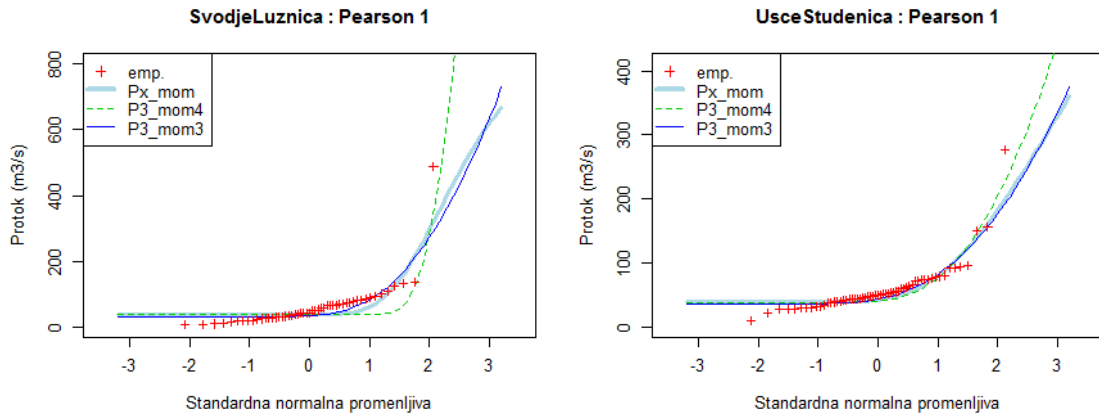
Slika 4. Primeri nizova iz grupe 4 sa koeficijentom asimetrije $C_s \approx 2$ sa izraženo oborenim gornjim krajem raspodele tipa I pod uticajem gornjeg parametra lokacije.

U domenu veoma izražene pozitivne asimetrije na stanicama u grupama 5 i 6, tj. za $C_s > 2$, ni tip I (prvi pristup) ni tip III (treći pristup) po pravilu ne može da da teorijsku raspodelu saglasnu sa empirijskom. U svih šest slučajeva iz ove dve grupe radi se stanicama sa izraženim gornjim izuzecima (Plavšić i sar., 2014). Neke od ovih stanica – a to su stanice iz grupe 5 – ipak „prolaze“ test saglasnosti KS po trećem pristupu (ali ne i test Kramer-fon Mizesa). Dva primera ovih rezultata data su na slici 5.

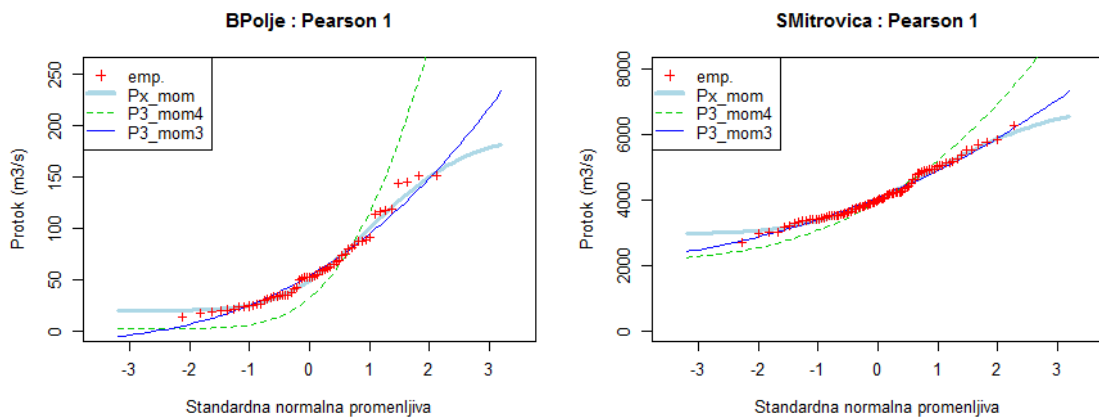
U grupu 2 svrstani su nizovi kod kojih je slaganje empirijske i teorijske raspodele dobro i u prvom i u trećem pristupu, ali veoma loše u drugom pristupu. Iako se radi o stanicama sa umerenom asimetrijom i spljoštenošću, ocena parametara tipa III iz formula (2) daje potpuno neprilagođene rezultate (slika 6).

17. Savetovanje SDHI i SDH - Vršac, Srbija 2015. Conference SDHI & SDH - Vršac, Serbia 2015.

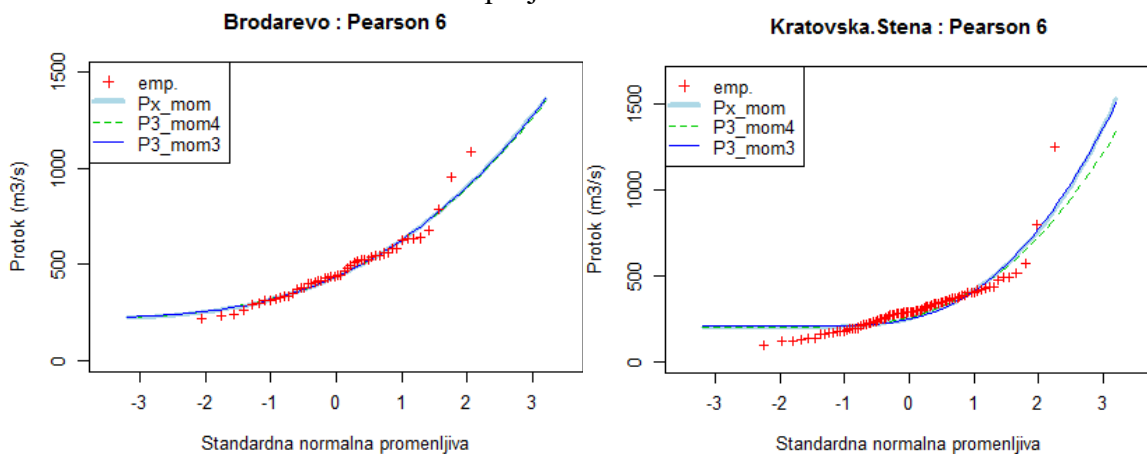
Na slici 7 prikazani su rezultati za dve stanice za koje momenti ukazuju na Pearsonovu raspodelu tipa VI. Podaci sa stanice Brodarevo na Limu imaju izuzetno dobro slaganje sa ovim tipom raspodele, što nije slučaj za stanicu Kratovska Stena na Zapadnoj Moravi na kojoj su koeficijenti asimetrije i spljoštenosti veoma veliki zbog prisustva izuzetaka.



Slika 5. Primeri nizova iz grupa 5 i 6 sa velikom asimetrijom i gornjim izuzecima koje nijedan tip raspodele ne opisuje dobro.



Slika 6. Primeri nizova iz grupe 2 sa lošom ocenom parametara tipa III iz prva četiri empirijska momenta.



Slika 7. Nizovi čiji momenti upućuju na Pearsonovu raspodelu tipa VI.

5 Zaključak

Pirsonov sistem raspodela ima prednost u primeni kada se rade istraživanja vezana za biomedicinu, demografiju, ekologiju, životni vek i slično. Zbog toga su se mnogi naučnici bavili ovom raspodelom i ispitivali raspodele koje se dobijaju za različite modifikacije diferencijalne jednačine (1). Iz toga su proistekle neke nove raspodele koje se koriste u praktičnim primenama. Kao alternativa osnovnom četvoroparametarskom sistemu mogu se koristiti polinomi višeg reda ili određene restrikcije parametara. Kako se parametrizacije modela kvadratnim i kubnim polinomima pokazuju kao suviše komplikovane za upotrebu, najčešće se primenjuju linearni brojlac i kvadratni imenilac.

U ovom radu namera je bila da se ispita primenljivost Pirsonovog sistema raspodela u hidrologiji za analizu nizova maksimalnih godišnjih protoka na rekama u Republici Srbiji. Za analizu su uzeti nizovi sa 28 stanica koji su zadovoljili testove slučajnosti i homogenosti. Razmatrana je primena opšteg četvoroparametarskog Pirsonovog sistema raspodela uz ocenu parametara metodom momenata na osnovu sistema jednačina koji povezuje parametre sa prva četiri momenta. Prema odnosu između momenata, većini stanica odgovara Pirsonova raspodela I tipa, a za dve stanice tipa VI. Raspodela tipa I je ograničena sa gornje strane, što generalno nije povoljno u analizi maksimalnih vrednosti kada se zahteva proračun merodavnih velikih voda male verovatnoće prevazilaženja, odnosno kada je potrebno ekstrapolirati usvojenu teorijsku raspodelu znatno izvan raspona osmotrenih vrednosti.

Rezultati su pokazali da se za nizove sa umerenom asimetrijom mogu dobiti solidni rezultati i uz navedeni pristup za prilagođavanje teorijske raspodele. Međutim, zbog korišćenja četvrtog uzoračkog momenta u tom pristupu i evidentnog unošenja u proračun veće neizvesnosti usled toga, rezultati mogu da budu i veoma loši ne samo za nizove sa većom asimetrijom i spljoštenošću, već i za ne tako asimetrične nizove. Klasičan pristup u hidrologiji, u kome se koristi tip III uz ocenu parametara na osnovu prva tri empirijska momenta, znatno je robusniji i po pravilu daje solidne rezultate osim kod ekstremne asimetrije i u prisustvu izuzetaka.

Analize sprovedene u ovom radu su bile ograničene na ocenu parametara metodom momenata, dok bi za donošenje celovitih zaključaka buduća istraživanja trebalo usmeriti i na metodu maksimalne verodostojnosti. Pored toga, izražena asimetrija hidroloških nizova ustoličila je primenu Pirsonove raspodele III tipa na logaritmovanim nizovima, pa buduća istraživanja svakako treba usmeriti i na primenu Pirsonovog sistema na logaritamski transformisane nizove.

I pored svega, ne može se poreći da je upotreba familije Pirsonovih raspodela za prethodne namene isključena. Treba naglasiti da je potrebno da se pridržava dobrih pravila statističke analize u hidrologiji. Kombinovanjem testiranja, grafičkog predstavljanja, poređenjem sa standardnim raspodelama u upotrebi (ovde je to bila Pirsonova raspodela tipa III), može se steći predstava o upotrebljivosti, dijapazonu primene kao i o specifičnostima softvera koji se koristi u pogledu izbegavanja mogućih pogrešnih zaključaka ako se softver koristi nekritički.

Zahvalnost

Autori se zahvaljuju Republičkom hidrometeorološkom zavodu za stavljanje podataka o maksimalnim godišnjim protocima na raspolaganje.

Literatura

1. Andreev A., Kanto A., Malo P., (2005), *Simple Approach For Distribution Selection In The Pearson System*, Helsinki School of Economics, Helsinki.
2. Lahcene, B. (2013), On Pearson families of distributions and its applications, *African Journal of Mathematics and Computer Science Research*, Vol. 6(5), pp. 108-117.
3. Đorić D., Mališić J., Jevremović V., Nikolić Đorić E., (2007) *Atlas Raspodela*, Građevinski fakultet, Beograd.
4. O'Connor J.J., Robertson E. F. (2003) *Karl Pearson*, elektronsko izdanje; <http://www-history.mcs.st-andrews.ac.uk/history/Biographies/Pearson.html>.
5. Becker M., Klossner S. (2015), *Package 'PearsonDS'*, dokumentacija R paketa, <https://cran.r-project.org/web/packages/PearsonDS/index.html>.
6. Kendall M.G. (1946), *The Advanced Theory of Statistics*, Volume 1, Charles Griffin & Company limited, London, 137-145.
7. Johnson N.L., Kotz S., Balakrishnan N. (1970), *Continuous Univariate Distribution*, Volume 1, John Wiley & Sons, INC, New York, 15-24.
8. Blagojević B., Mihailović V., Plavšić J. (2014) Statistička analiza velikih voda na profilima hidroloških stanica: potreba za promenom pristupa, *Vodoprivreda*, 46(1-6): 199-209.
9. Plavšić J., Mihailović V., Blagojević B. (2014) Assessment of methods for outlier detection and treatment in flood frequency analysis. Proc. Mediterranean Meeting on Monitoring, modelling, early warning of extreme events triggered by heavy rainfall, University of Calabria, Cosenza, Italy, 26-28 June 2014. E. Ferrari and P. Versace (eds.), ISBN 978-88-6822-268-0, pp. 181-192.
10. Rao A.R. and Hamed K.H. (2000) *Flood Frequency Analysis*. CRC Press, New York.