

Sveučilište u Zagrebu

Građevinski fakultet

Lovro Štefan, Luka Đud

Klasifikacija hidrograma prema obliku na slivu rijeke Kupe i Krapine pomoću algoritma
strojnog učenja

Zagreb, 2024.

Ovaj rad izrađen je u Hidrotehničkom zavodu Građevinskog fakulteta (GF) i u suradnji s Fakultetom elektrotehnike i računarstva pod vodstvom izv. prof. dr. sc. Kristine Potočki (GF), prof. dr. sc. Damira Pintara (FER) i predan je na natječaj za dodjelu Rektorove nagrade u akademskoj godini 2023./2024.

Istraživanja u ovom radu predstavljaju nastavak preliminarnih istraživanja provedenih u okviru doktorskog rada "Influence of climate change and flood wave characteristics on local scour around bridge piers" koji je financiran od Hrvatske zaklade za znanost u okviru projekta R3PEAT - "Remote Real-time Riprap Protection Erosion Assessment on large rivers" (UIP-2019-04-4046) i natječaja „Projekt razvoja karijera mladih istraživača – Izobrazba novih doktora znanosti“ (DOK-2020-01-5354).

POPIS KRATICA

Kratica	Objašnjenje
AM	Metoda godišnjih maksimuma (engl. <i>Annual Maxima</i>)
POT	Metoda premašenja praga (engl. <i>Peak Over Threshold</i>)
DHMZ	Državni hidrometeorološki zavod
GMP	Glavna meteorološka postaja
AMP	Automatska meteorološka postaja

SADRŽAJ

1	UVOD.....	1
2	OPĆI I SPECIFIČNI CILJEVI RADA	3
3	METODE I TIJEK RADA	3
4	REZULTATI	17
5	RASPRAVA.....	42
6	ZAKLJUČCI	44
7	ZAHVALE.....	46
8	POPIS LITERATURE.....	46
9	SAŽETAK	50
10	SUMMARY	51
	PRILOG	53

1 UVOD

Pri projektiranju većine hidrotehničkih građevina ili osmišljanju mjera obrane od poplava jedan od ključnih problema je određivanje mjerodavnih projektnih vodostaja, odnosno projektnih protoka na koje se dimenzioniraju hidrotehničke građevine. Na području Europe, ali i u ostatku svijeta svake godine događaju se poplave koje često imaju razorne posljedice na ljude, prirodu i građevine, a njihov učinak sve je naglašeniji klimatskim promjenama [1]. Izbor mjerodavnog projektnog protoka prema tome izravno utječe na sigurnost ljudi i građevina, a istovremeno uvjetuje troškove i racionalnost građevine ili mjera zaštite od voda. Stoga je u hidrološkoj i hidrotehničkoj praksi vrlo važno pitanje definiranja novih ili aktualizacije postojećih metoda određivanja projektnog protoka koje bi na optimalan način ispunile sve zahtjeve. Osim vršnog protoka, važne (a često zanemarivane) su i informacije o volumenu, trajanju i obliku vodnog vala, a objedinjene su u projektnom hidrogramu [2].

Metode definiranja projektnog protoka i hidrograma mogu se podijeliti u dvije glavne skupine, probabilističke i determinističke metode, ovisno o tome radi li se na izučenim ili neizučenim slivovima [3]. Izučeni slivovi su oni za koje postoje dugi nizovi izmjerениh protoka, a neizučeni su oni na kojima se mjerena ne provode ili su kraćeg vremenskog niza. Determinističke metode koriste se kod nedovoljno izučenih ili neizučenih slivova. Ove metode podatke o oborinama koje koristeći razne parametre sliva ili klime modeliraju u protok na izlazu iz sliva. Često su u obliku iskustvenih jednadžbi, poput racionalne metode korištene diljem svijeta ili Srebrenovićeve metode koju često primjenjuju inženjeri u Hrvatskoj. Sljedeću skupinu determinističkih metoda predstavljaju jedinični hidrogrami iz kojih se na temelju količine i prostornog rasporeda kiše mogu odrediti maksimalni protok i volumen vodnog vala. Iz jediničnog hidrograma može se dobiti sintetički hidrogram, koji jednostavnim oblikom trokuta definiranim vremenom rasta i pada te vršnom ordinatom opisuje otjecanje na neizučenim ili nedovoljno izučenim slivovima [4]. Prednosti determinističkih metoda su uporabivost na bilo kojem poprečnom profilu vodotoka ili na neizučenim slivovima, a mane grubi rezultati i osjetljivost na ulazne parametre koji su subjektivan izbor projektanta. Druga skupina su probabilističke (vjerojatnosne) metode za izučene slivove. Baziraju se na definiranju uzorka vodnih valova - izborom po jednog najvećeg vala po godini u slučaju dužeg kontinuiranog niza podataka (metoda godišnjih maksimuma, *annual maxima*, AM) ili svih valova iznad određenog praga za kraće nizove podataka (niz premašenja, *peak over threshold*, POT). Na tako formirane nizove vodnih valova može se prilagoditi funkcija distribucije vjerojatnosti, poput normalne, log-normalne, Gumbelove raspodjele i drugih. Ovdje je važan

ispravan izbor krivulje vjerojatnosti jer za velika povratna razdoblja, odnosno male vjerojatnosti premašenja, različite funkcije raspodjele vjerojatnosti za isti ulazni niz vodnih valova mogu dati drastično različite protoke [5]. Uz determinističke metode i probabilističke analize mogu se koristiti i složenije metode kao kombinacije spomenutih, poput računalnih modela oborine-otjecanje kod kojih je potrebna kalibracija, stohastičko generiranje nizova oborina te vjerojatnosna analiza. Kombinirane metode rjeđe su u praksi, kao i metode regionalnih analiza, koje se kod slivova s nedovoljno podataka koriste za određivanje protoka i hidrograma usporedbom s izučenim slivovima sličnih svojstava [6].

Rezultat određivanja projektnog protoka i projektnog hidrograma često je samo jedna vrijednost protoka i jedan oblik hidrograma, koji ne moraju nužno dobro predstavljati različitost otjecanja na promatranom slivu, a događaju s najvećim protokom ne mora odgovarati i maksimalni volumen vodnog vala [5]. Stoga se nameće pitanje prikaza ponašanja sliva većim brojem projektnih hidrograma različitih oblika [7], a sintetičke dijagrame grubo pojednostavljenih oblika danas sve više zamjenjuju složenije aproksimacijske krivulje. Brunner i sur. [2] predlažu metodu definiranja projektnog hidrograma na način da se vodni valovi aproksimiraju kontinuiranim funkcijama i klasificiraju po obliku u određeni broj klastera. Metoda omogućava bolje obuhvaćanje varijabilnosti oblika hidrograma i olakšava analizu te predviđanje poplavnog ponašanja na regionalnoj razini. Slična metoda u literaturi je primijenjena i za tipologiju poplavnih događaja korištenjem parametarskog klasteriranja [8]. Preliminarna istraživanja klasifikacije hidrograma na jednoj vodomjernoj postaji na rijeci Savi, pokazala su potencijal za primjenu metoda strojnog učenja za primjenu ovih metoda za analizu hidrograma u Hrvatskoj [9], međutim rezultati nisu bili statistički validirani, niti primjenjeni na izučene slivove različite veličine.

Ovim se radom istražuje mogućnost primjene slične metodologije - klasteriranja na temelju oblika hidrograma aproksimiranih kontinuiranim funkcijama (funkcijska analiza) - na slivove rijeka Krapine i Kupe. Slivovi su odabrani prema kriteriju površine, sliv rijeke Krapine kao mali sliv, a sliv rijeke Kupe kao srednje velik sliv te predstavljaju ogledne kategorije slivova za koje se u Republici Hrvatskoj može primijeniti navedena metodologija. Osim toga, radi se o slivovima na kojima postoji velika opasnost od poplava. Nadalje, projektni hidrogrami s ovih slivova predstavljaju dio ulaznih podataka za hidrološko-hidraulički model sustava za obranu od poplava Srednje Posavljte su za određivanje njihovog oblika korišteni ranije spomenuti pristupi bazirani na empirijskim izrazima ili njihove varijacije [10]. Sliv rijeke Krapine zbog svojih karakteristika sklon je plavljenju potencijalno uzrokujući velike štete na

gusto naseljenom području velike gospodarske aktivnosti [11]. Sliv rijeke Kupe smatra se područjem iznimne strateške vrijednosti [12], a zbog geografskog položaja njegova zaštita od poplava ima međunarodni značaj [13].

2 OPĆI I SPECIFIČNI CILJEVI RADA

Cilj rada je testirati metodu predstavljenu u Brunner i sur. [2], prema potrebi je dodatno razviti i prilagoditi, te provesti postupak klasificiranja vodnih valova na hidrološkim postajama Kupljenovo na rijeci Krapini i Farkašić na Kupi u svrhu definiranja normaliziranih oblika projektnih hidrograma, a kroz rad iskoristiti u hidrologiji često korišten programski jezik R te kroz njega primijeniti strojno učenje za obradu i klasteriranje podataka - vodnih valova. Dodatni ciljevi su analizirati podatke o oborinama na promatranim slivovima u ovisnosti o dobivenim klasterima te identificirati i predložiti daljnje korake prema izradi projektnih dijagrama iz dobivenih klastera.

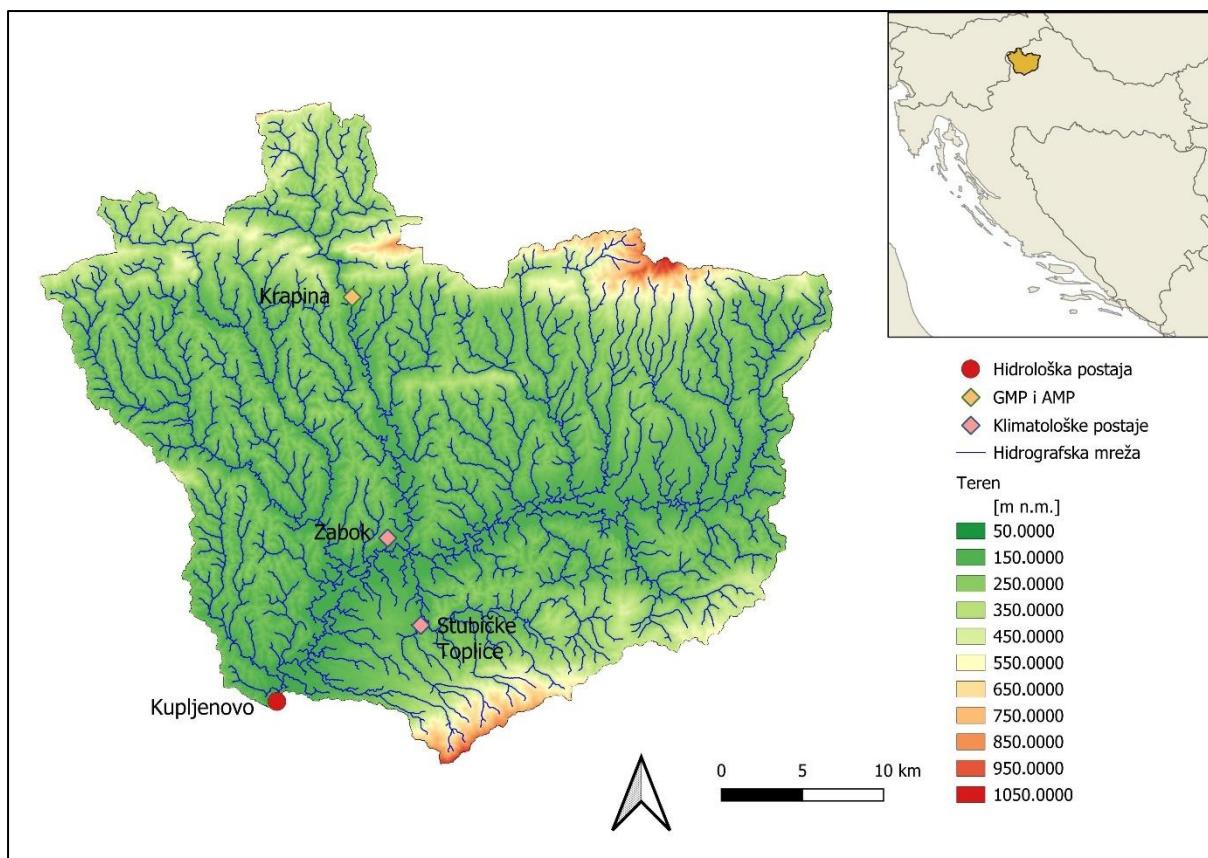
3 METODE I TIJEK RADA

3.1. Podaci i područje istraživanja

Glavni ulazni podaci za izradu rada su hidrološki i meteorološki podaci. Hidrološki podaci su dnevni podaci o protocima na hidrološkim postajama Kupljenovo na rijeci Krapini te postaji Farkašić na Kupi. Hidrološka postaja Kupljenovo (podaci o postaji u Tablica 1) nalazi se na geografskoj širini $45^{\circ} 56' 05''$ i geografskoj dužini $15^{\circ} 49' 04''$, 13,8 km uzvodno od ušća Krapine u Savu, a pripadno slivno područje površine je 1150 km^2 . Dnevni protoci dostupni su za razdoblje od 1. siječnja 1964. do 31. prosinca 2022. godine. Postaja Farkašić (Tablica 2) na Kupi smještena je 47,15 km uzvodno od ušća u Savu, na širini $45^{\circ} 29' 00''$ i dužini $16^{\circ} 09' 10''$. Površina sliva rijeke Kupe do postaje Farkašić iznosi 8992 km^2 . Podaci o dnevnim protocima dostupni su za razdoblja od 1. siječnja 1965. do 31. prosinca 1990. te od 1. siječnja 2000. do 31. prosinca 2022. godine. Podaci od 1991. do 1999. izostaju zbog ne provođenja mjerena uslijed ratnog stanja. Za dodatnu analizu uzeti su i podaci o satnim protocima koji su za Kupljenovo dostupni u razdoblju od 1.1.2001. do 31.12.2022., a za Farkašić od 1.1.2007. do 31.12.2022. Dodatno su prikupljeni i meteorološki podaci o dnevnim oborinama - na slivu rijeke Krapine odabrane su glavna i automatska meteorološka postaja (GMP i AMP) Krapina (dostupni podaci od 1994. do 2022.), klimatološka postaja Stubičke Toplice (1961.-1987.,

1990.-2015., 2016. - 2021.) i klimatološka postaja Zabok (1992.-2021.), a na slivu Kupe GMP i AMP Parg (1961.-2022.), klimatološka postaja Delnice (1981.-1994., 2007.-2021.), GMP i AMP Karlovac (1961.-2022.) i klimatološka postaja Pisarovina (1981.-2003., 2005.-2021.). Svi prethodno navedeni podaci zatraženi su od Državnog hidrometeorološkog zavoda te dobiveni u txt i csv formatu. Slika 1 i Slika 2 prikazuju karte slivova Kupe i Krapine do promatranih hidroloških postaja izrađene u programu QGIS 3.38 Grenoble. Na kartama su označeni topografija terena, hidrografska mreža sliva i položaj hidrološke, meteoroloških i klimatoloških postaja. Vidljiv je i položaj sliva u Republici Hrvatskoj.

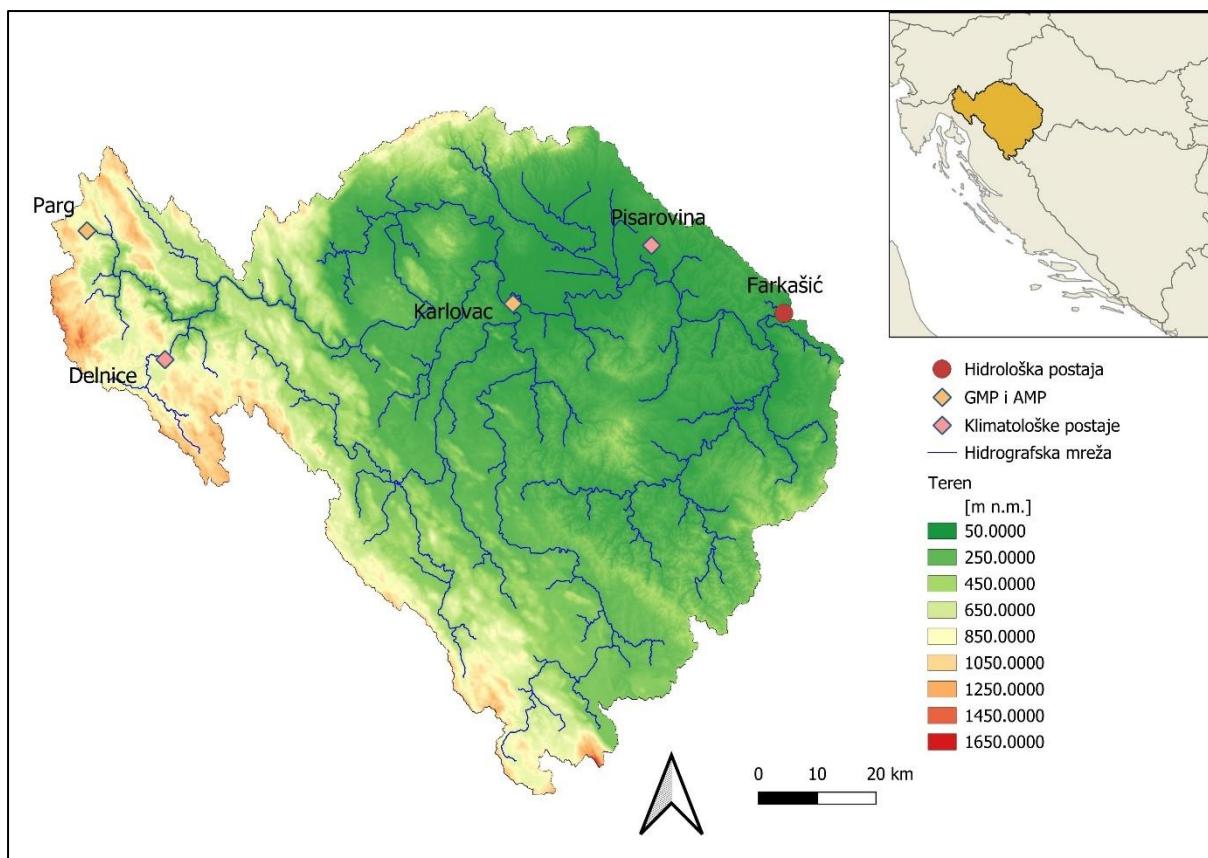
Tablica 3 prikazuje pregled statistike podataka o protocima dobivenih od DHMZ-a, redom su navedeni broj opservacija dnevnih protoka, prosječni protok, standardna devijacija, minimalni i maksimalni protok, protok 1. kvartila, medijalni protok te protok 3. kvartila. Tablica 4 daje popis meteoroloških i klimatoloških postaja za koje su dobiveni podaci o oborinama, a iz tablice su vidljivi periodi mjerjenja oborina i razdoblja prekida mjerjenja.



Slika 1: Karta sliva rijeke Krapine do hidrološke postaje Kupljenovo

Tablica 1: Podaci o hidrološkoj postaji Kupljenovo

Izvor podataka	Državni hidrometeorološki zavod	
Vodotok	Krapina	
Hidrološka postaja	Kupljenovo	
Šifra postaje	3054	
Početak rada	10. rujna 1963.	
Kota nule vodokaza [m.n.m]	128.877	
Geografske koordinate postaje	širina	dužina
	45° 56' 05"	15° 49' 04"
Gaus-Krueger	X	Y
	5088155	5563758
Površina sliva [km²]	1150	
Udaljenost od ušća [km]	13.8	



Slika 2: Karta sliva rijeke Kupe do hidrološke postaje Farkašić

Tablica 2: Podaci o hidrološkoj postaji Farkašić

Izvor podataka	Državni hidrometeorološki zavod	
Vodotok	Kupa	
Hidrološka postaja	Farkašić	
Šifra postaje	4010	
Početak rada	17. rujna 1964.	
Kota nule vodokaza [m.n.m]	93.820	
Geografske koordinate postaje	širina	dužina
	45° 29' 00"	16° 09' 10"
Gaus-Krueger	X	Y
	5038321	5590513
Površina sliva [km²]	8992	
Udaljenost od ušća [km]	47.15	

Tablica 3: Pregled podataka o dnevnim protocima na hidrološkim postajama Kupljenovo i Farkašić

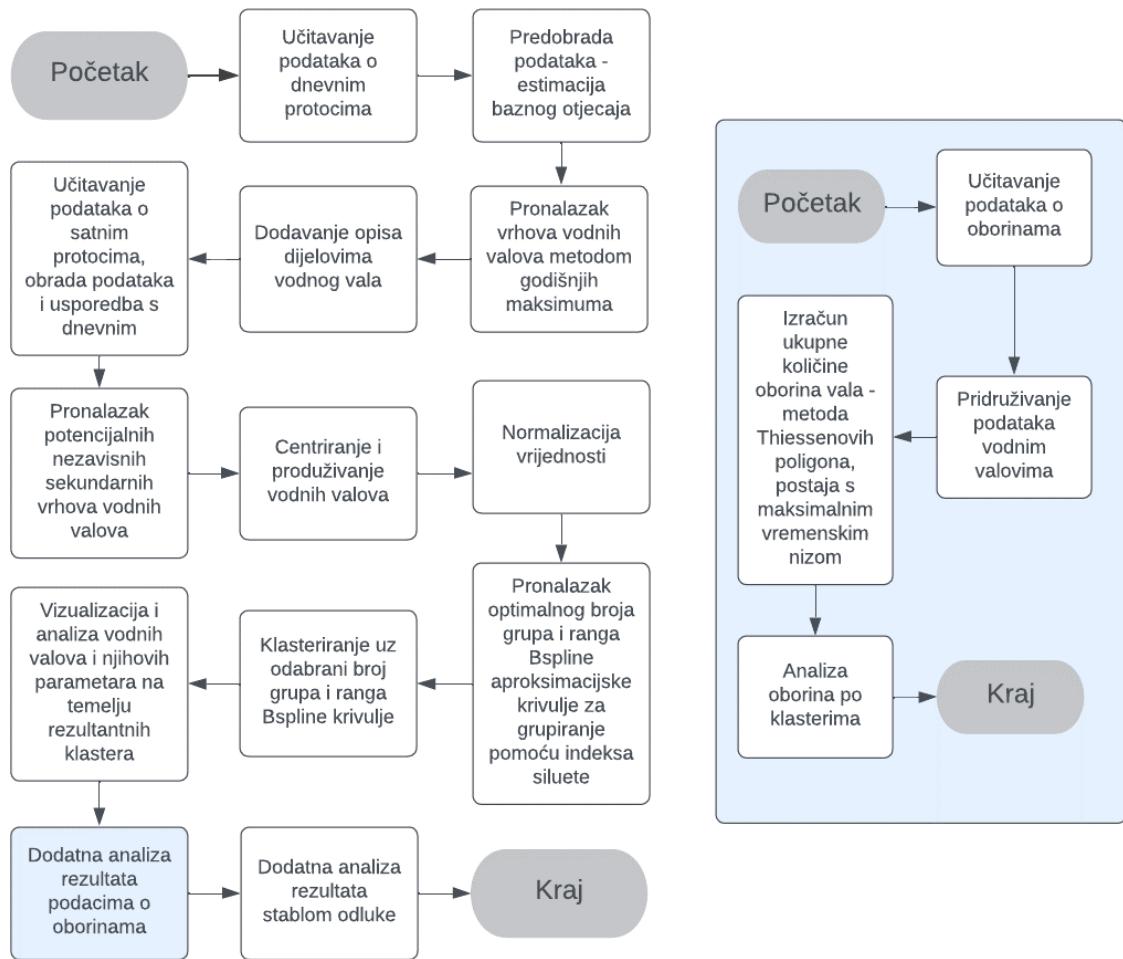
Hidrološka postaja	Broj dnevnih podataka o protoku	Q_{avg} [m ³ /s]	St. dev. [m ³ /s]	Q_{min} [m ³ /s]	Q_{max} [m ³ /s]	$Q_{1.\text{kvartil}}$ [m ³ /s]	Q_{med} [m ³ /s]	$Q_{3.\text{kvartil}}$ [m ³ /s]
Kupljenovo	21550	11.10	17.88	0.34	252.00	2.59	4.93	11.23
Farkašić	17897	193.55	202.57	13.60	1585.00	56.20	193.55	256.00

Tablica 4: Pregled podataka o meteorološkim i kišomjernim postajama

Sliv	Ime postaje	Vrsta postaje	Lokacija [N, E]	Nadmorska visina [m n.m.]	Mjerenje oborina	Prekidi u mjerjenjima
Krapina	Krapina	GMP i AMP	46° 8' 16", 15° 53' 18"	202	1994. - 2022.	-
	Zabok	klimatološka	46° 1' 11", 15° 53' 26"	150	1992. - 2001.	-
	Stubičke Toplice	klimatološka	45° 58' 31", 15° 55' 27"	180	1961. - 2021.	1988. - 1999., 2015.
Kupa	Parg	GMP i AMP	45° 35' 37", 14° 37' 50"	863	1961. - 2022.	-
	Delnice	Klimatološka	45° 24' 16", 14° 48' 29"	681	1981. - 2021.	1995. - 2006.
	Karlovac	GMP i AMP	45° 29' 37", 15° 33' 54"	110	1961. - 2022.	-
	Pisarovina	klimatološka	45° 34' 35", 15° 52' 57"	138	1981. - 2021.	2004.

3.2. Metodologija

Glavni koraci korištene metodologije temeljeni su na radu Brunner i sur. [2], a prikazani su dijagramom toka (Slika 3). Koraci su detaljnije objašnjeni u nastavku.

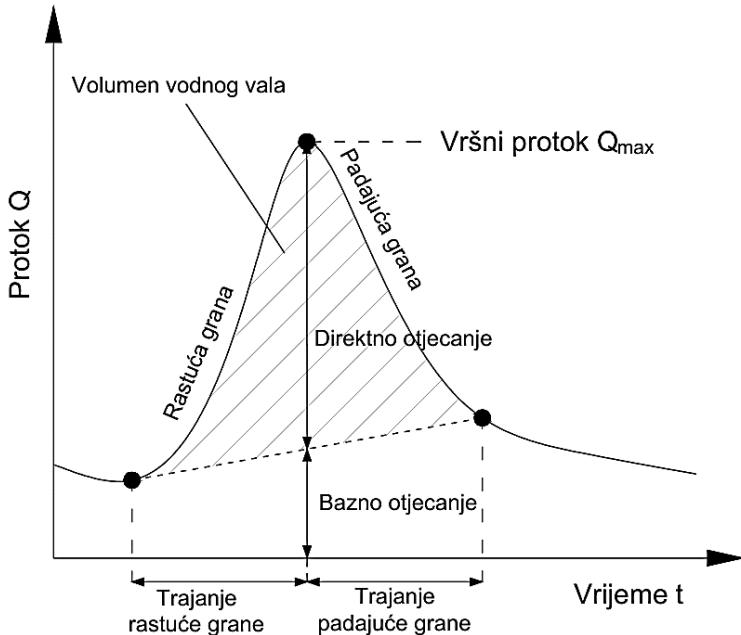


Slika 3: Pojednostavljeni prikaz glavnih koraka metodologije određivanja klastera oblika hidrograma

Obrada i analiza podataka provedeni su koristeći RStudio, razvojno okruženje za programski jezik R koji je pogodan za statističke analize. R se često koristi u području hidrologije za koju su razvijeni i brojni paketi koji omogućuju obradu, analizu i prikaz hidroloških podataka. Korišteni paketi uključuju tidyverse - zbirku paketa za obradu, uređivanje, čišćenje i prikaz podataka [14], lubridate - olakšava rad s vremenskim serijama

[15], fda - funkcionalna analiza podataka [16], writexl - ispis rezultata u xlsx obliku [17], zoo - obrada nepravilnih vremenskih serija [18], funData - statistička analiza funkcijskih podataka [19], cluster - klasifikacija podataka u klastere i njihova analiza [20], factoextra - vizualizacija i interpretacija rezultata multivarijantne analize podataka [21], rpart - stvaranje stabla odlučivanja [22], te caret - treniranje i prikaz klasifikacijskih i regresijskih modela. Također je korišten i lfstat – paket specijaliziran za analizu i modeliranje hidroloških podataka, uključujući određivanje baznog protoka [23].

Podaci dobiveni od DHMZ-a predstavljaju ukupni protok, odnosno otjecanje te je potrebno provesti odvajanje baznog otjecanja kako bi se u daljnje korake išlo s vodnim valovima formiranim na temelju protoka direktnog otjecanja. Bazno otjecanje je protok koji se sporo mijenja te protječe vodotokom i u sušnim razdobljima, dok je direktno otjecanje uzrokovano oborinama i očituje se skokovima u hidrogramu otjecanja [24]. Odvajanje baznog otjecanja te glavne značajke hidrograma koje definiraju njegov oblik prikazane su shematski (Slika 4). Glavne značajke hidrograma koje definiraju njegov oblik su trajanje rastuće i padajuće grane, vršni protok te volumen vodnog vala. Za odvajanje baznog otjecanja korištena je naredba *createlfobj* iz paketa *lfstat*. Ta funkcija je pogodna za stvaranje objekta baznog protoka za vremenski niz dnevnih podataka o protoku. Bazno otjecanje definira detektiranjem minimalnih lomnih točaka u hidrogramu kao početak i kraj baznog otjecanja, a vrijednosti između njih linearno interpolira. Direktni protok dobiven je oduzimanjem baznog otjecanja od dnevnih podataka o protoku te se za njega dalje određuje oblik.

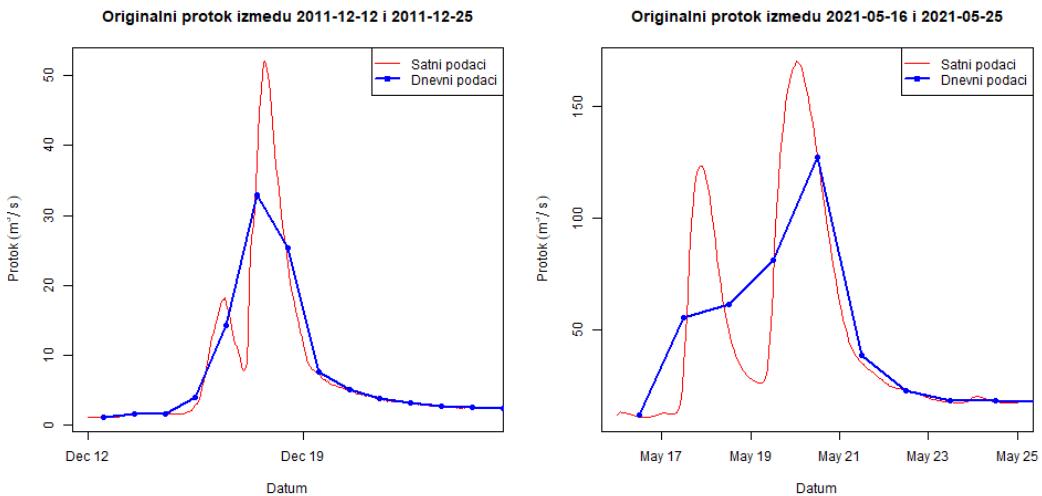


Slika 4: Shematski prikaz hidrograma

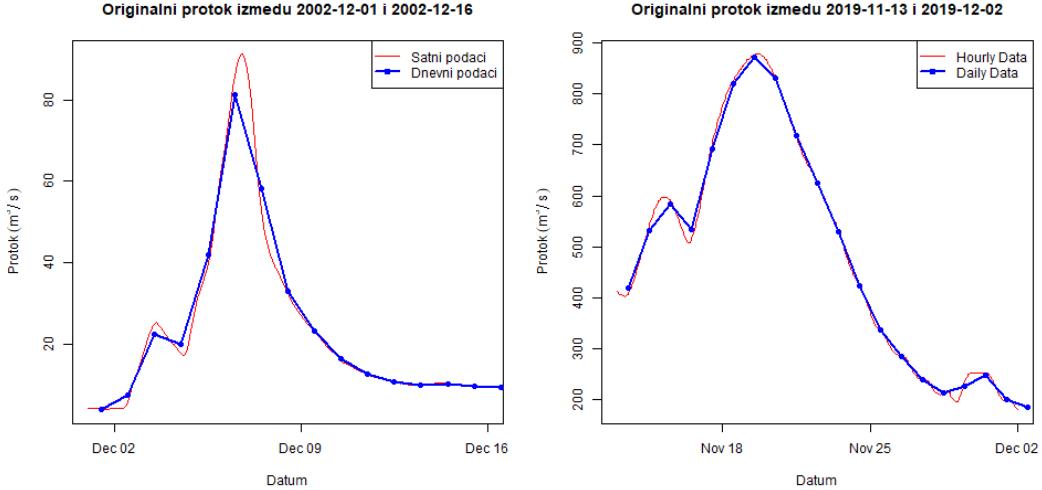
Nizovi dnevnih podataka dovoljno su dugi (58 godina za Kupljenovo, 47 godina za Farkašić) za formiranje uzorka vodnih valova pristupom godišnjih maksimuma - u literaturi se različito opisuje potrebna dužina vremenskog niza, od 14 do 30-ak godina [25, 26]. Maksimalni protoci su određeni na temelju hidrološke godine. Hidrološke godine su za oba sliva definirane s početkom 1. rujna i krajem 31. kolovoza, a nazvane su prema kalendarskoj godini u kojoj počinju. Ovakve hidrološke godine odabrane su jer je uočeno da se najveći protoci u godini u najvećem broju događaju u zimskom razdoblju te se njihovim korištenjem izbjegava protezanje maksimalnog vodnog vala kroz dvije godine, kakav slučaj bi se dogodio uporabom kalendarskih godina (ili hidroloških godina s početkom u siječnju).

Provedena je usporedba dnevnih i satnih podataka o protoku na obje hidrološke postaje za presjek razdoblja dostupnosti dnevnih i satnih mjerjenja. Dnevne i satne podatke važno je usporediti jer dnevni podaci daju samo jednu vrijednost protoka, prosječnu dnevnu, a satni daju 24 satnih prosječnih protoka. Prema tome satni podaci puno točnije opisuju protok kao kontinuiranu varijablu te vršni protok, a time je i oblik hidrograma definiran satnim podacima bliži stvarnom obliku vodnog vala nego hidrogram definiran dnevnim podacima. Određena je njihova prosječna razlika za valove definirane prema dnevnim podacima. Za Kupljenovo prosječno odstupanje iznosi $17.85 \text{ m}^3/\text{s}$ (13.8% razlike), a za Farkašić $15.19 \text{ m}^3/\text{s}$ (1.5% odstupanja). U određenim godinama za Kupljenovo se maksimalni protok iz dnevnih podataka ne poklapa s datumom pojave maksimalnog satnog protoka (npr. za hidrološku

godinu 2006. - dnevni maksimum: 21.3.2007, $65.7 \text{ m}^3/\text{s}$; satni maksimum: 13.2.2007, $91 \text{ m}^3/\text{s}$). Osim toga, u određenim godinama se jednak maksimalni protok pojavljuje više različitih datuma. Napravljena je i vizualna usporedba oblika hidrograma satnih i dnevnih podataka prema kojoj se može zaključiti da se uglavnom vrlo dobro poklapaju valovi određeni dnevnim i satnim podacima o protoku, uz par iznimaka znatnijeg odstupanja na postaji Kupljenovo (Slika 5). Primjer usporedbe dnevnog i satnog vala dan je na Slika 6. S obzirom na veliku sličnost oblika, a samim time i volumena hidrograma za satne i dnevne podatke, te zbog činjenice da su nizovi dnevnih podataka na raspolaganju za znatno dulji vremenski interval od satnih, odlučeno je da su dnevni podaci o protoku primjereni za daljnje korake.



Slika 5: Hidrogrami s najvećim odstupanjem satnih i dnevnih podataka o protoku



Slika 6: Primjer dobrog poklapanja satnih i dnevnih podataka za Kupljenovo (lijevo) i Farkašić (desno)

Nakon određivanja maksimalnih protoka po hidrološkim godinama, pristupa se definiranju pripadajućeg hidrograma. Definirani su početak i kraj hidrograma (kao dani najbliži vrhu u kojima je vrijednost direktnog otjecanja 0). Zatim je provjereno postoji li unutar tako definiranog vodnog vala još koji sekundarni vrh vodnog vala te su takvi sekundarni vrhovi odbačeni ako su zadovoljen jedan od sljedećih kriterija [27]:

$$t > 5 \text{ dana} + \log(A)$$

$$Q_{min} < \frac{3}{4} \min [Q_{max,1}, Q_{max,2}]$$

Ovdje je:

t - broj dana između maksimalnog protoka i detektiranog sekundarnog vrha

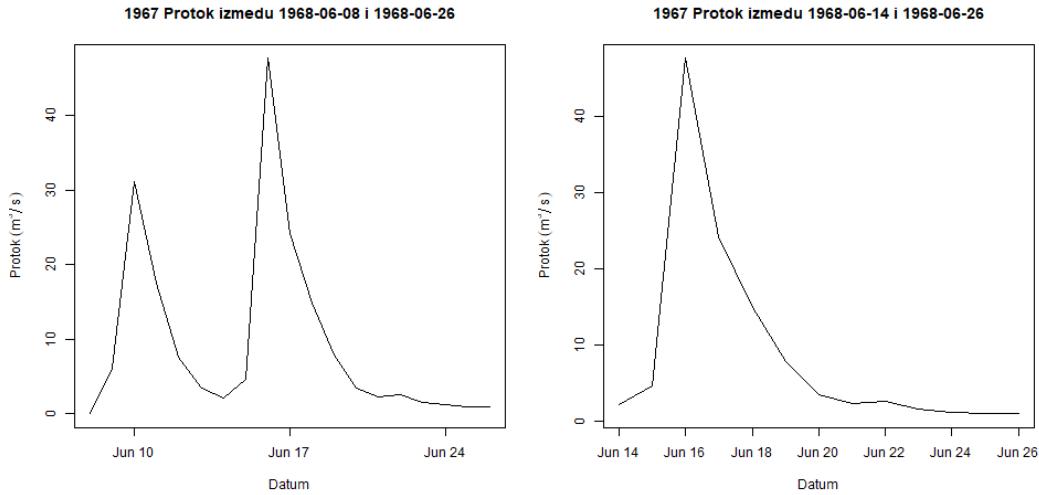
A - površina sliva u četvornim miljama

Q_{min} - minimalna vrijednost protoka između primarnog i sekundarnog vrha

$Q_{max,1}, Q_{max,2}$ - redom vrijednosti primarnog i sekundarnog vršnog protoka

Drugim riječima, ako je vrijeme proteklo između dvaju vršnih protoka veće od broja dana dobivenih prvim od gornjih kriterija ili ako je minimalni protok između njih manji od tri četvrtine vrijednosti vrha manjeg protoka, sekundarni vrh se može smatrati nezavisnim o primarnom vršnom protoku te se smatra vrhom zasebnog vodnog vala. U takvim slučajevima

sekundarni vrh se odbacuje iz početno definiranog vodnog vala. Njemu se za novi početni ili krajnji trenutak (ovisno je li sekundarni vrh bio smješten prije ili nakon primarnog vrha) postavlja dan koji odgovara postizanju minimalne vrijednosti između sekundarnog i primarnog vrha. Primjer „rezanja“ sekundarnog vala dan je na Slika 7.



Slika 7: Usporedba prikaza vodnog vala prije (lijevo) i nakon (desno) odbacivanja neovisnog sekundarnog vala

Valovi su zatim centrirani na način da se za svaki val dan postizanja vršnog protoka smatra nultim danom (prema tome dani rasta vala nalaze se na negativnoj strani x osi, a dani pada vala na pozitivnoj strani). Osim toga određeni su valovi s maksimalnim trajanjem grane rasta i grane pada te su prema tim vrijednostima produžene grane rasta i pada za preostale valove (produžene u smislu produženja datumi trajanja, no njima je pridružena vrijednost protoka nula, tako da zadržavaju svoj stvarni prijašnji oblik). Slijedi normalizacija valova - izračunana je površina ispod krivulje svakog vodnog vala te je svaka vrijednost ordinate hidrograma podijeljena s površinom ispod krivulje. Na taj način dobiveni su svojevrsni jedinični dijagrami koji nose informaciju o obliku svakog vala. Prethodni koraci potrebni su kako bi se mogli pristupiti klasteriranju vodnih valova prema obliku hidrograma.

3.2.1. K-means klasteriranje

K-means klasteriranje jedna je od najčešće korištenih metoda za klasteriranje u domeni znanosti o podacima. Ova metoda grupira skup podataka u klastera na način da je sličnost između objekata unutar istog klastera najveća, uz minimiziranje sličnosti objekata različitih klastera [28]. Ključni koncept K-meansa temelji se na traženju k centroida kojima se minimiziraju razlike unutar klastera.

Algoritam započinje inicijalizacijom k centroida, što može biti nasumično ili prema određenom kriteriju. Zatim se svaki podatak dodjeljuje najbližem centru, nakon čega se centri ažuriraju kao srednja vrijednost svih točaka unutar pojedinog klastera. Ovaj proces se ponavlja dok se centri ne prestanu mijenjati, odnosno dok sustav konvergira, ili dok se ne dostigne zadani broj iteracija [28].

Glavni izazov ovog algoritma je određivanje optimalnog broja klastera k , jer se taj broj često ne zna unaprijed. Također, K-means je osjetljiv na početne uvjete odnosno početni odabir centroida, što može dovesti do različitih rezultata za različite početne točke. Zbog toga su razvijena brojna proširenja i varijante algoritma, kao što je X-means, koji automatski prilagođava broj klastera tijekom iteracija [28].

Još jedna nadogradnja K-meansa je fuzzy K-means algoritam, koji omogućava djelomično pripadanje točaka u više klastera. Svaka točka ima svoj stupanj pripadnosti klasterima, umjesto da strogo pripada samo jednom. Novi centroidi izračunavaju se kao ponderirane srednje vrijednosti svih točaka, pri čemu je ponderiranje određeno stupnjem pripadnosti. Ovaj pristup često rezultira boljim performansama kod podataka koji imaju nejasne granice između klastera te je posebno koristan u područjima poput segmentacije slika [29].

U ovom istraživanju korišten je K-means algoritam za klasteriranje hidrograma, pri čemu su ulazni podaci aproksimirani pomoću Bspline krivulja. K-means je odabran zbog svoje jednostavnosti i efikasnosti u grupiranju podataka prema sličnosti, dok Bspline krivulje omogućuju fleksibilnu i preciznu aproksimaciju složenih oblika hidrograma. Bspline krivulje su vrsta spline funkcija koje koriste niz kontrolnih točaka za formiranje glatke krivulje koja prolazi kroz ili blizu ovih točaka, što omogućuje kontinuiranu i preciznu reprezentaciju hidroloških podataka zabilježenih na diskretnim vremenskim točkama. Ova metodologija je inspirirana radom Brunner i sur. [2], koji su koristili funkcionalne podatke za klasteriranje oblika hidrograma u cilju identifikacije regija sa sličnom reaktivnošću na poplave.

Bspline krivulje predstavljaju skupinu matematičkih funkcija koje se koriste za glatko interpoliranje skupa podataka. Svaka Bspline krivulja se definira pomoću kontrolnih točaka, čiji položaj utječe na oblik krivulje. Prednost Bspline krivulja je njihova fleksibilnost i sposobnost da precizno aproksimiraju složene oblike, što je posebno korisno kod reprezentacije hidrograma, gdje preciznost u opisivanju oblika krivulje protoka-vremena (Q-t) može biti ključna za analizu hidroloških procesa. Kontrolne točke omogućuju prilagodbu krivulje tako

da prolazi blizu zadanih podataka, čime se osigurava glatka i kontinuirana reprezentacija koja može biti korisna za daljnje analize, točnije funkcijске analize oblika krivulje hidrograma [2].

Za evaluaciju stabilnosti i kvalitete formiranih klastera korištena je analiza siluete. Koeficijent siluete mjeri koliko su objekti unutar istog klastera slični jedni drugima u usporedbi s objektima iz drugih klastera. Formalno, vrijednost siluete za podatkovnu točku je definirana kao:

$$s(i) = (b(i) - a(i)) / \max \{a(i), b(i)\}$$

gdje je $a(i)$ prosječna udaljenost između točke i te do svih ostalih objekata unutar istog klastera, a $b(i)$ najmanja prosječna udaljenost od točke do objekata iz drugih klastera. Ukupni koeficijent siluete je prosječna vrijednost $s(i)$ za sve objekte. Vrijednosti koeficijenta siluete kreću se od -1 do 1, gdje veće vrijednosti ukazuju na bolje definirane klastera. Visoke vrijednosti siluete (blizu 1) ukazuju na to da je podatkovna točka dobro smještena unutar svog klastera, dok niske vrijednosti (blizu -1) sugeriraju da bi podatkovna točka mogla biti bolje smještena u neki drugi klaster. [30]

Za optimizaciju parametara K-means algoritma, provedeno je pretraživanje mreže (engl. *grid search*) kako bi se pronašao optimalni broj klastera i odgovarajući rang Bspline krivulje. Testirani su različiti rangovi Bsplinea (4, 10, 15, 20, 25), kako bi se osiguralo da se podaci adekvatno modeliraju, dok je broj klastera varirao između 2 i 6. Ovaj pristup omogućio je sustavno ispitivanje kombinacija parametara kako bi se identificirala ona koja rezultira najkvalitetnijim klasterima. Dodatno je provjerena stabilnost indeksa siluete na podskupovima vremenskih nizova za oba sliva. Ovim postupkom se osigurava dosljednost i kvaliteta klasteriranja. Stabilni rezultati indeksa siluete na različitim podskupovima ukazuju na to da klasteri nisu slučajni, već da se njihova kvaliteta održava kroz različite vremenske nizove. To pruža dodatnu potvrdu pouzdanosti dobivenih grupacija i dokazuje robusnost klasteriranja prema promjenama u podacima.

Nakon klasteriranja provedena je analiza varijance (ANOVA) kako bi se provjerila statistička značajnost razlike parametara koji opisuju oblik projektnog hidrograma (protok, volumen, trajanje - rast i pad, oborine) po klasterima.

Dodatno je provedena i analiza količine oborina koja prethodi vršnom protoku vala. Uzete su oborine za vrijeme trajanja rasta hidrograma uz dodatak jednog dana prije početka hidrograma. Promatrane su dvije različite količine oborina: oborine s postaje s maksimalnim

vremenskim nizom podataka te oborina određena kao težinski prosjek (metoda Thiessenovih poligona) za najduže razdoblje s podacima iz svih postaja za koje su dostavljeni podaci. Provjerena je povezanost količine oborina s pripadnosti valova prethodno dobivenim klasterima.

U svrhu dodatne analize rezultata klasteriranja i boljeg razumijevanja faktora (parametri koji opisuju vodni val) koji utječu na grupiranje podataka, primijenjena je tehniku nadziranog učenja, konkretno stablo odluke. Na ovaj način, klasteri koji su formirani K-means algoritmom sada služe kao oznake (labele) odnosno klase za podatke, čime omogućavamo izgradnju prediktivnog modela koji može objasniti na koji način se specifične karakteristike hidrograma povezuju s dodijeljenim klasterima [31]. Korištenje stabla odluke pomaže nam u vizualizaciji i interpretaciji ključnih parametara koji određuju pripadnost određenom klasteru, što dodatno doprinosi razumijevanju hidroloških obrazaca i procesa specifičnih za odabране lokacije te može biti od velike koristi za buduće analize hidroloških ponašanja.

4 REZULTATI

4.1. Sliv rijeke Krapine

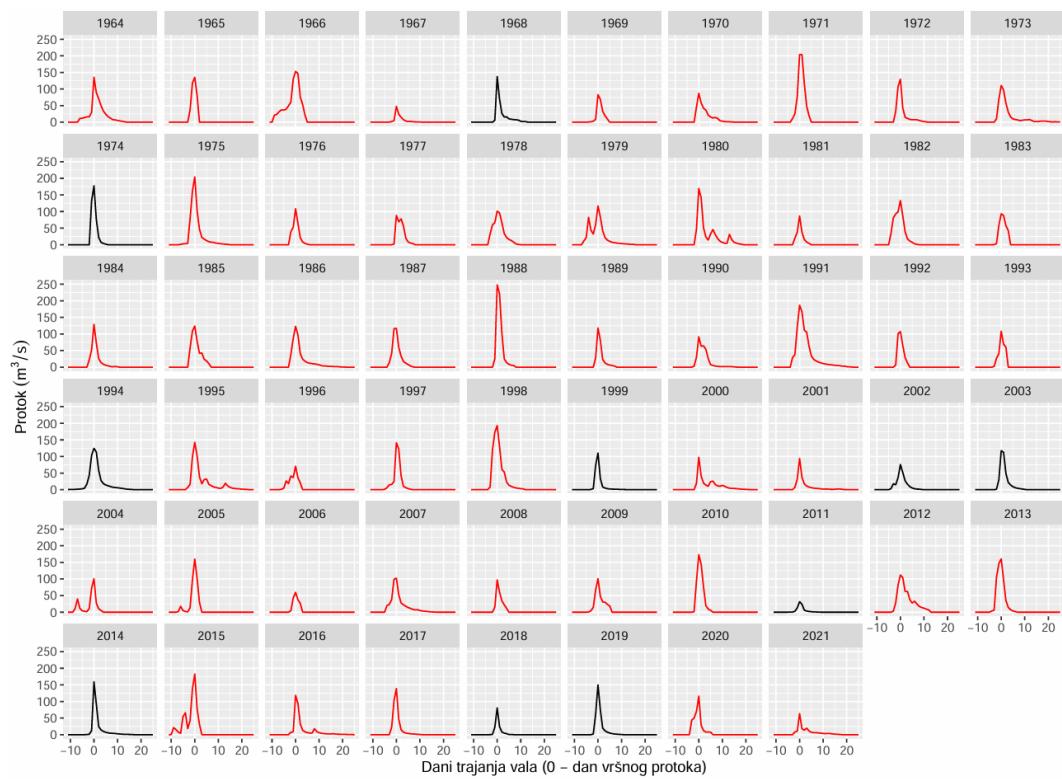
Za rijeku Krapinu definirano je 58 vodnih valova, čije su karakteristike dani u Tablica P1. Tablica 5 prikazuje statističke vrijednosti parametara za sve valove - prosječne, maksimalne i minimalne vrijednosti, standardnu devijaciju, medijan te prvi i treći kvartil.

Tablica 5: Statističke vrijednosti parametara vodnih valova na rijeci Krapini

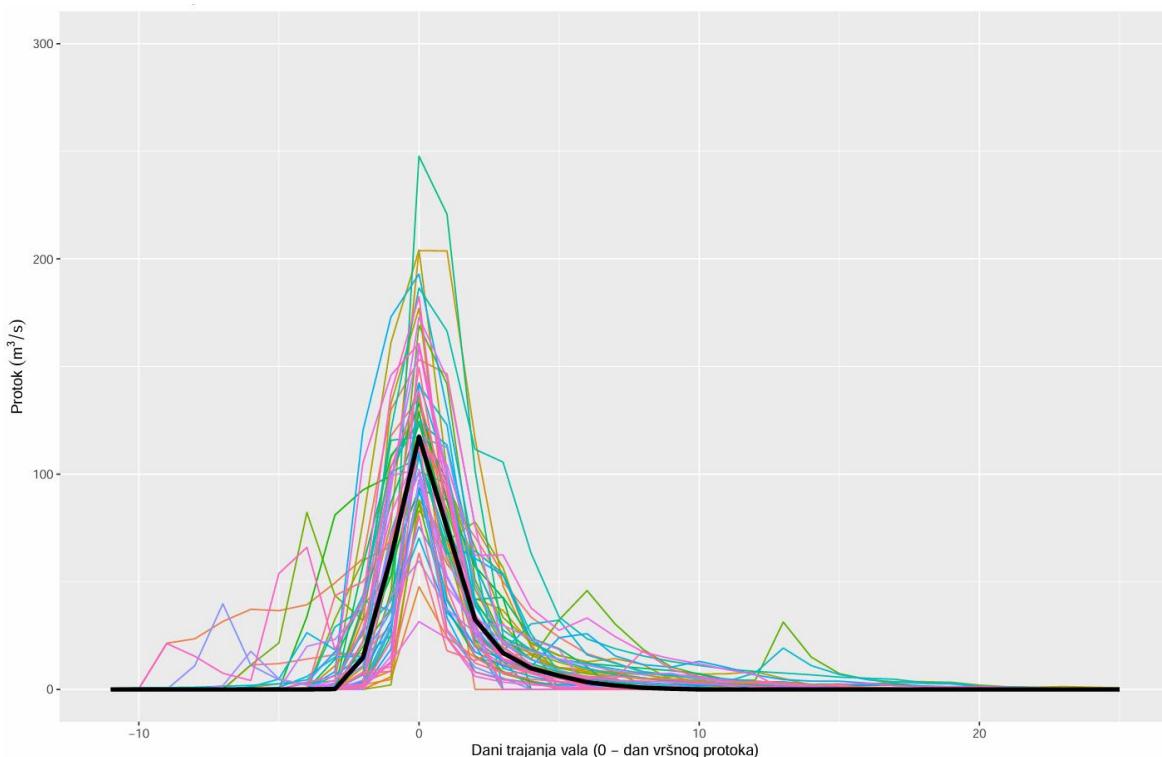
	Izmjereni protok [m ³ /s]	Direktno otjecanje [m ³ /s]	Trajanje rasta [dani]	Trajanje pada [dani]	Ukupno trajanje [dani]	Volumen vodnog vala [10 ⁶ m ³]
Prosjek	129.29	122.82	3.67	10.19	14.86	36.28
Min	33.00	31.51	1.00	1.00	5.00	7.47
Max	252.00	247.70	11.00	25.00	29.00	85.87
St. dev.	41.54	41.30	2.29	6.50	6.24	16.36
1. kvartil	103.68	96.09	2.00	5.00	10.00	25.79
Medijan	125.00	117.35	3.00	9.00	13.50	31.17
3. kvartil	154.50	144.23	4.00	15.00	20.00	48.69

Prosječni vršni protok maksimalnih godišnjih valova na rijeci Krapini iznosi 129.29 m³/s, uz prosječni volumen vodnog vala 36.28 milijuna kubičnih metara, a prosječno trajanje iznosi 14.86 dana. Prosječno trajanje grane rasta iznosi 3.67 dana, a grana pada traje u prosjeku 10.19 dana. Val s minimalnim vršnim protokom od 33 m³/s zabilježen je 2011. godine, a maksimalni protok je za val iz 1988. hidrološke godine - 252 m³/s. Val s maksimalnim trajanjem od 29 dana je iz 1994. hidrološke godine, a val iz 1965. godine najmanjeg je ukupnog trajanja od 5 dana. Trajanje grane rasta je od minimalno 1 dan (1970., 1980. i 2000. hidrološka godina) do maksimalno 11 dana (1994.), a grane pada od minimalno 1 dan (1965.) do maksimalno 25 dana (1973. i 2016. hidrološka godina). Val s minimalnim protokom iz 2011. godine ima i najmanji pripadni volumen (7.47 milijuna m³), dok je val s maksimalnim volumenom (85.87 milijuna m³) onaj iz 1991. godine.

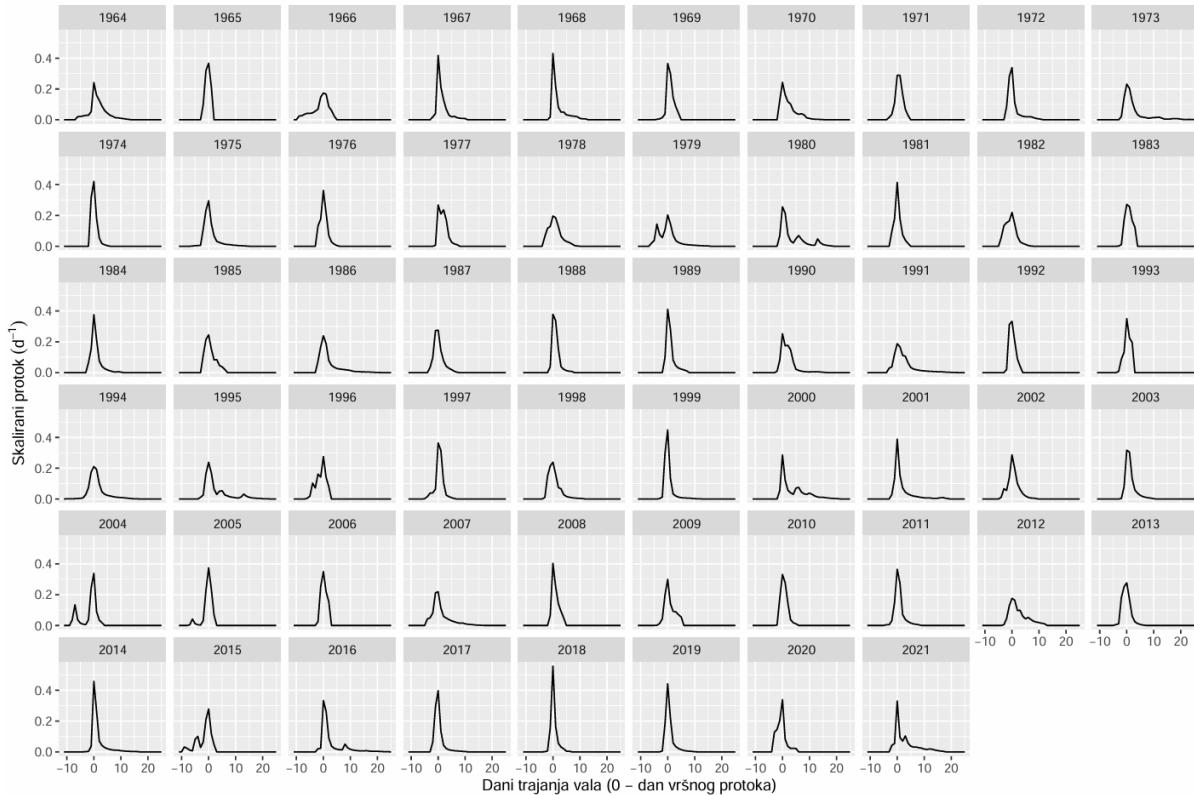
Slika 8 prikazuje sve definirane valove na rijeci Krapini prema ukupnom protoku. Na slici su različito prikazani vodni valovi kod kojih su detektirani te „odrezani“ sekundarni valovi (crveno) te valovi kod kojih nisu bili prisutni sekundarni zavisni valovi (crno). Na svega 10 valova nisu uočeni sekundarni, dok je kod preostalih 48 bio prisutan nezavisni sekundarni val.



Slika 8: Centrirani hidrogrami vodnih valova na rijeci Krapini, crvenom bojom prikazani su valovi kod kojih su odrezani nezavisni sekundarni valovi



Slika 9: Centrirani hidrogrami vodnih valova na rijeci Krapini, crnom bojom prikazan je medijan



Slika 10: Centrirani i normalizirani hidrogrami vodnih valova na rijeci Krapini

Slika 9 je objedinjeni prikaz svih hidrograma vodnih valova na rijeci Krapini centriranih na način da je dan vršnog protoka definiran kao nulti dan, a na Slika 10 vidljivi su pojedinačni vodni valovi u normaliziranom obliku.

Kroz analizu indeksa siluete za različite kombinacije ranga krivlje i broja klastera (Tablica 6), za sliv rijeke Krapine identificirana je najstabilnija karakteristika klastera s indeksom siluete oko 0.6 za Bspline rang 4 i broj klastera 2. Unatoč tome, odabrana su 4 klastera za konačnu analizu. Ova odluka donesena je na temelju vizualne procjene, koja je pokazala da se s 4 klastera mogu bolje uočiti i razumjeti razlike među različitim oblicima hidrograma. Iako je indeks siluete bio nešto niži za ovu konfiguraciju, prednosti u smislu interpretabilnosti i korisnosti rezultata za hidrološku analizu prevagnule su u korist korištenja 4 klastera. Napravljena je i provjera stabilnosti indeksa siluete na podskupovima podataka za rijeku Krapinu (1964-1992 i 1993-2022) te je uočeno da se rezultati otprilike poklapaju, u obje polovice najstabilnija su 2 klastera ranga 4, s vrijednostima koje se malo razlikuju (0.59 za prvu polovicu podataka i 0.63 za drugu polovicu podataka).

Tablica 6: Indeks siluete za različite kombinacije ranga krivulje i broja klastera, Kupljenovo

Rang Bspline krivulje	Broj centroida u K-means	Indeks siluete
4	2	0.60348
4	3	0.53968
4	4	0.56821
4	5	0.49652
4	6	0.50179
10	2	0.40955
10	3	0.40771
10	4	0.39629
10	5	0.36739
10	6	0.34405
15	2	0.41442
15	3	0.36085
15	4	0.33964
15	5	0.30460
15	6	0.30044
20	2	0.34103
20	3	0.30360
20	4	0.32839
20	5	0.32251
20	6	0.30187
25	2	0.55227
25	3	0.46341
25	4	0.44776
25	5	0.40021
25	6	0.42447

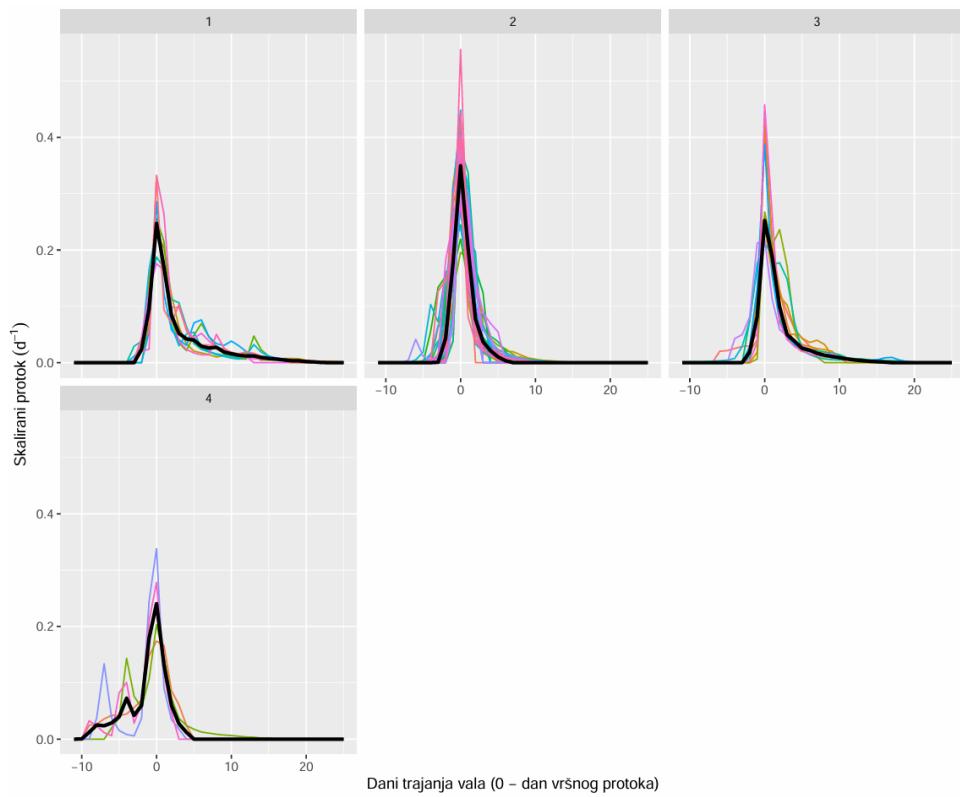
U nastavku su prikazani rezultati klasteriranja vodnih valova na rijeci Krapini u 4 klastera. Tablica 7 prikazuje koliko valova je svrstano u pojedini klaster, Slika 11 prikazuje normalizirane valove po klasterima sa crno iscrtanim centroidom za svaki klaster, Slika 12 valove prema protoku po klasterima, a pregled parametara valova po klasterima, kao i oborine za valove po klasterima, dani su u Tablica 8, odnosno Tablica 9.

Tablica 7: Broj valova po klasterima za hidrograme na rijeci Krapini

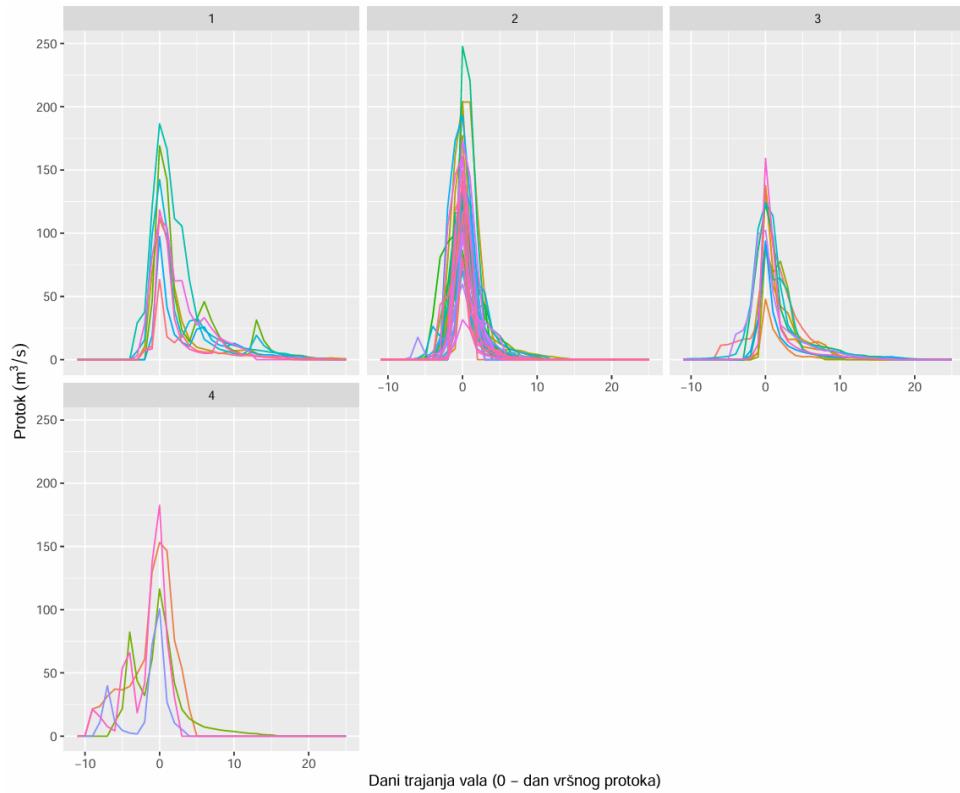
Klaster	Broj valova	Udio [%]
1	8	13.8
2	35	60.3
3	11	19.0
4	4	6.9

Najveći broj valova, 35 od 58, svrstano je u 2. klaster. Slijede 3. i 1. klaster s po 11 i 8 valova, dok je u 4. klaster raspoređeno najmanje klastera - 4.

U nastavku su dani kutijasti dijagrami (Slika 13 do Slika 18) koji prikazuju raspon parametara po klasterima. Iz navedenih dijagrama vidljivo je veliko preklapanje vrijednosti parametara za protok, međutim za trajanje rasta, pada i ukupno trajanje, te volumen i količinu oborina vidljive su razlike među klasterima. Navedena vizualna procjena potvrđena je i analizom varijance, čiji su rezultati pokazali da je razlika trajanja (ukupnog, grane rasta i pada) među klasterima statistički vrlo značajna (na razini 0.001), a umjerenou značajna je razlika ukupnog volumena vodnog vala i količine oborina po klasterima (na razini 0.05). Analiza varijance nije pokazala statistički značajnu razliku protoka po klasterima. Za svaki klaster prikazan je i hidrogram centroida svih vodnih valova u koji pripadaju tom klasteru u normaliziranom obliku (Slika 19 do Slika 22), a centroidi hidrograma po klasterima s originalnim protokom dani su u prilogu (Slika P1 do P4). Centroid predstavlja krivulju koja prolazi kroz medijane protoka ili normaliziranog protoka po y-osi za svaki dan na x-osi.



Slika 11: Normalizirani hidrogrami vodnih valova na rijeci Krapini po klasterima, crnom bojom prikazan centroid klastera



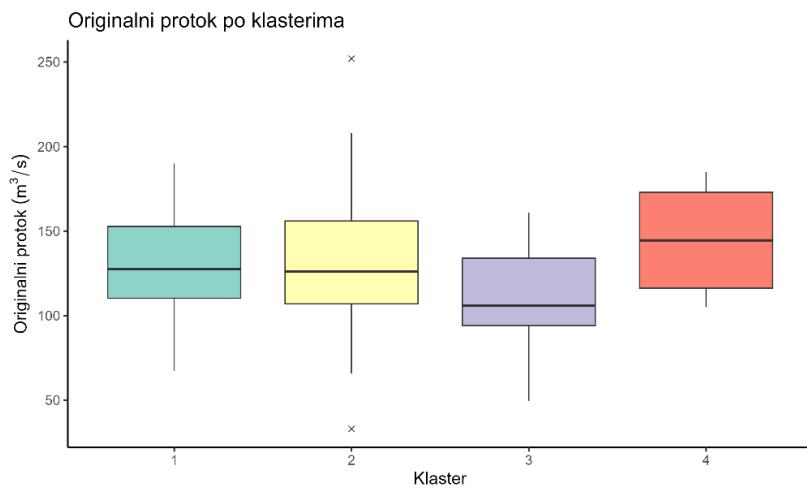
Slika 12: Hidrogrami vodnih valova na rijeci Krapini po klasterima

Tablica 8: Pregled statistike valova po klasterima za hidrograme na rijeci Krapini

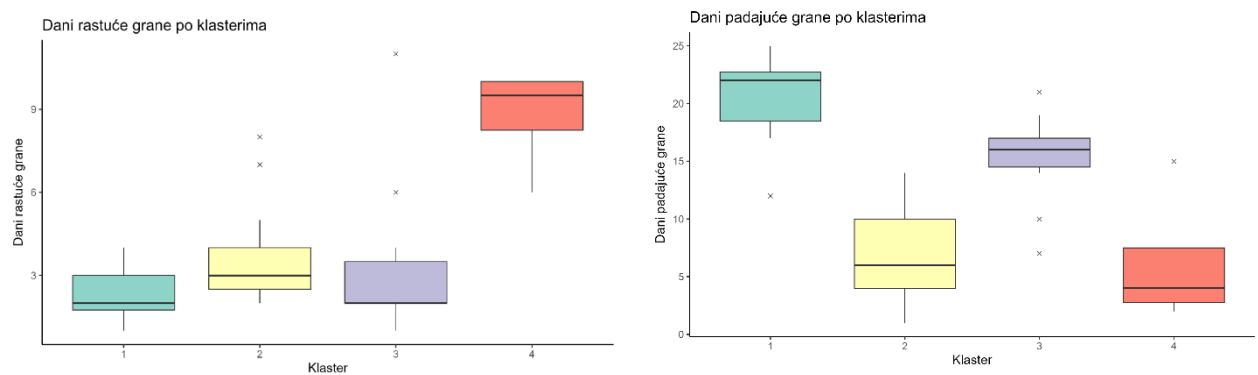
	Klaster	Izmjereni protok [m ³ /s]	Direktno otjecanje [m ³ /s]	Trajanje rasta [dani]	Trajanje pada [dani]	Ukupno trajanje [dani]	Volumen vodnog vala [10 ⁶ m ³]
Projek	Svi valovi	129.29	121.27	3.67	10.19	14.86	36.28
	1	131.04	124.96	2.25	20.50	23.75	45.95
	2	132.55	125.17	3.51	6.71	11.23	33.30
	3	112.04	108.19	3.36	15.18	19.55	33.04
	4	144.75	138.17	8.75	6.25	16.00	52.00
Max	1	190.00	186.50	4.00	25.00	28.00	85.87
	2	252.00	247.70	8.00	14.00	22.00	69.73
	3	161.00	159.15	11.00	21.00	29.00	50.96
	4	185.00	182.48	10.00	15.00	22.00	76.07
Min	1	67.30	63.28	1.00	12.00	17.00	16.57
	2	33.00	31.51	2.00	1.00	5.00	7.47
	3	49.60	47.71	1.00	7.00	10.00	9.87
	4	105.00	100.59	6.00	2.00	13.00	25.70
St. dev	1	39.00	39.68	1.04	4.38	4.03	21.45
	2	45.00	45.01	1.50	3.47	3.63	14.43
	3	31.35	31.16	2.87	3.89	5.16	12.22
	4	38.30	36.87	1.89	5.97	4.24	20.81

Tablica 9: Pregled oborina po klasterima za hidrograme na rijeci Krapini

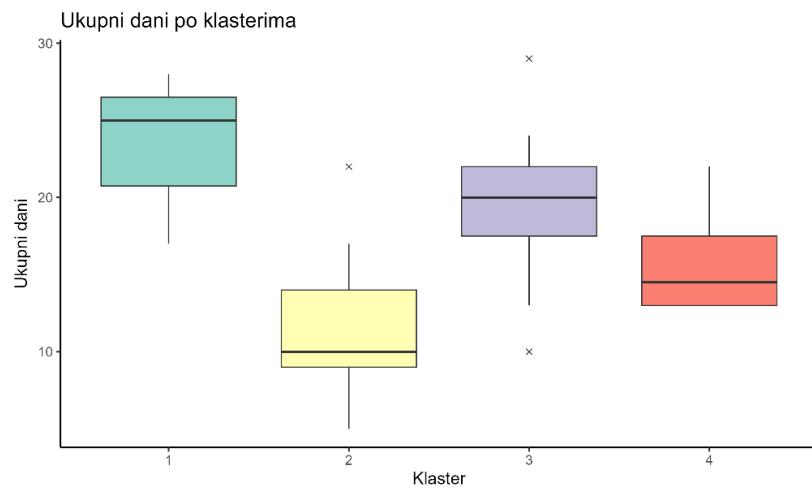
	Klaster	Oborine - postaja s najduljim nizom [mm]	Oborine - težinski prosjek svih postaja [mm]
Projek	1	23.33	22.55
	2	45.12	42.11
	3	26.51	35.34
	4	112.28	108.49
Max	1	46.70	66.68
	2	115.80	112.29
	3	94.40	70.47
	4	147.50	124.26
Min	1	1.50	1.82
	2	1.50	0.11
	3	4.30	6.48
	4	57.50	92.72
St. dev	1	18.37	27.92
	2	28.48	32.34
	3	27.54	27.01
	4	40.96	22.30



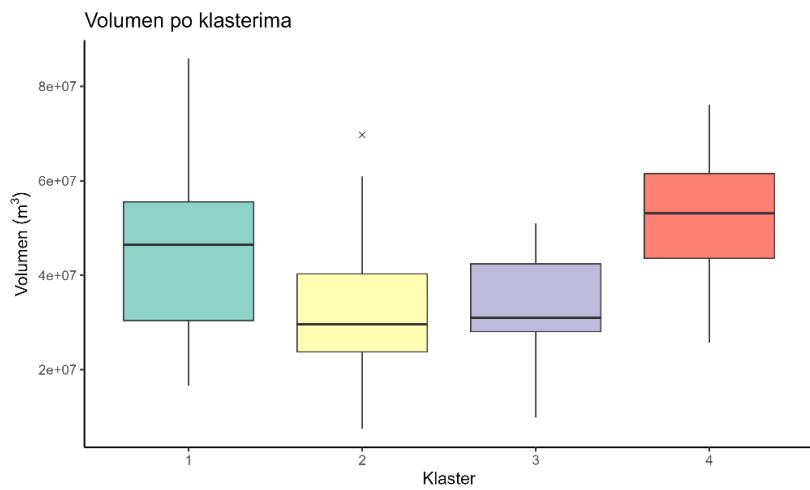
Slika 13: Vršni protok po klasterima za rijeku Krapinu



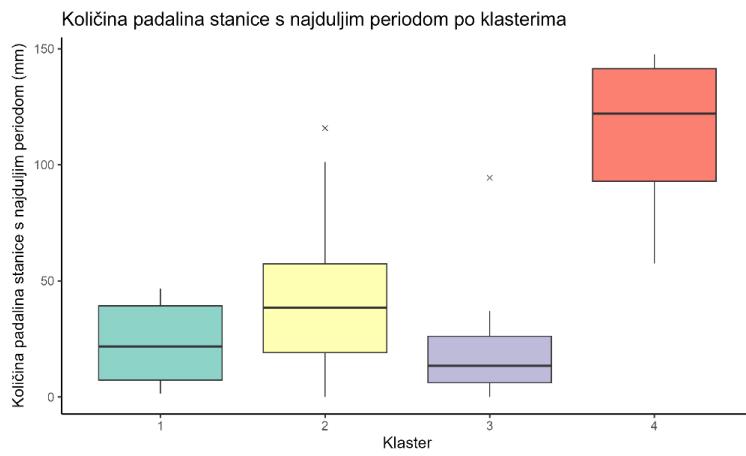
Slika 14: Trajanje rasta (lijevo) i pada (desno) vodnog vala po klasterima za rijeku Krapinu



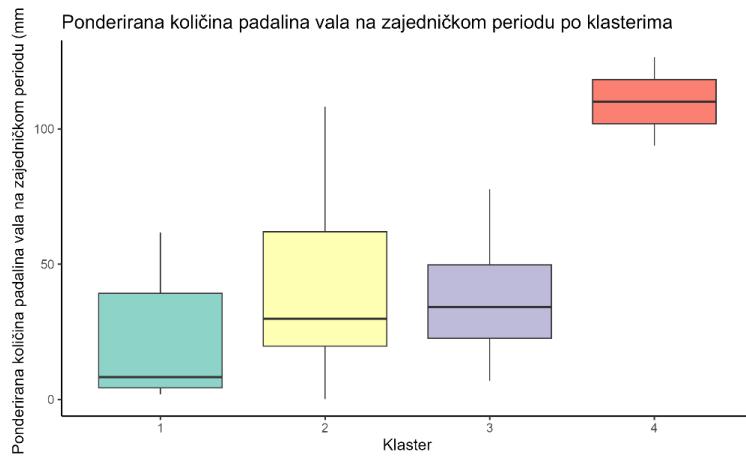
Slika 15: Ukupno trajanje vodnog vala po klasterima za rijeku Krapinu



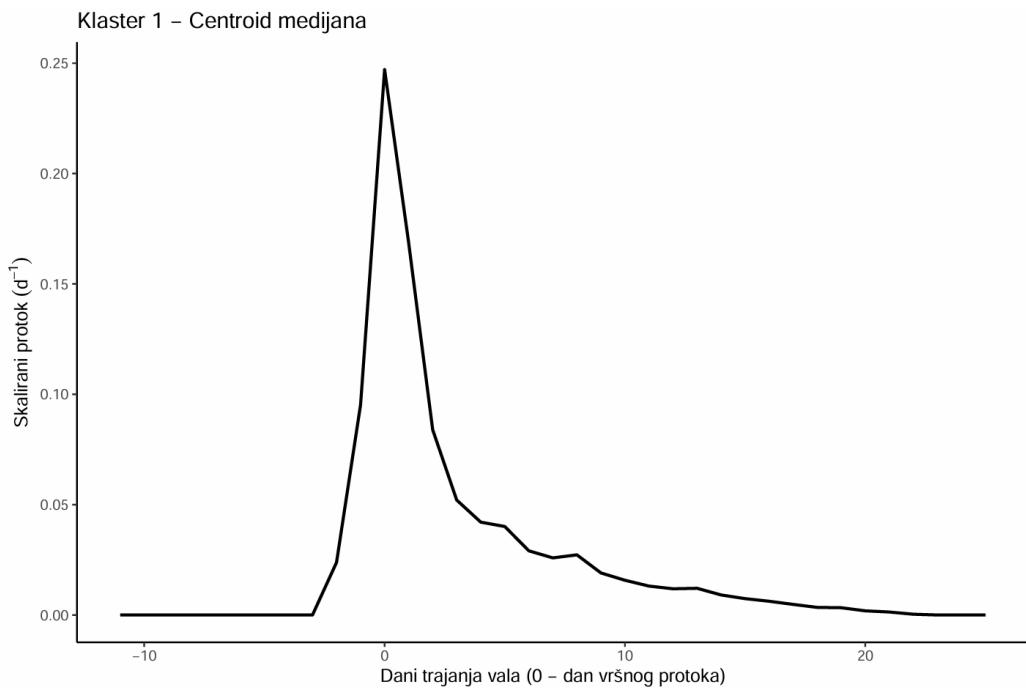
Slika 16: Volumen vodnih valova po klasterima za rijeku Krapinu



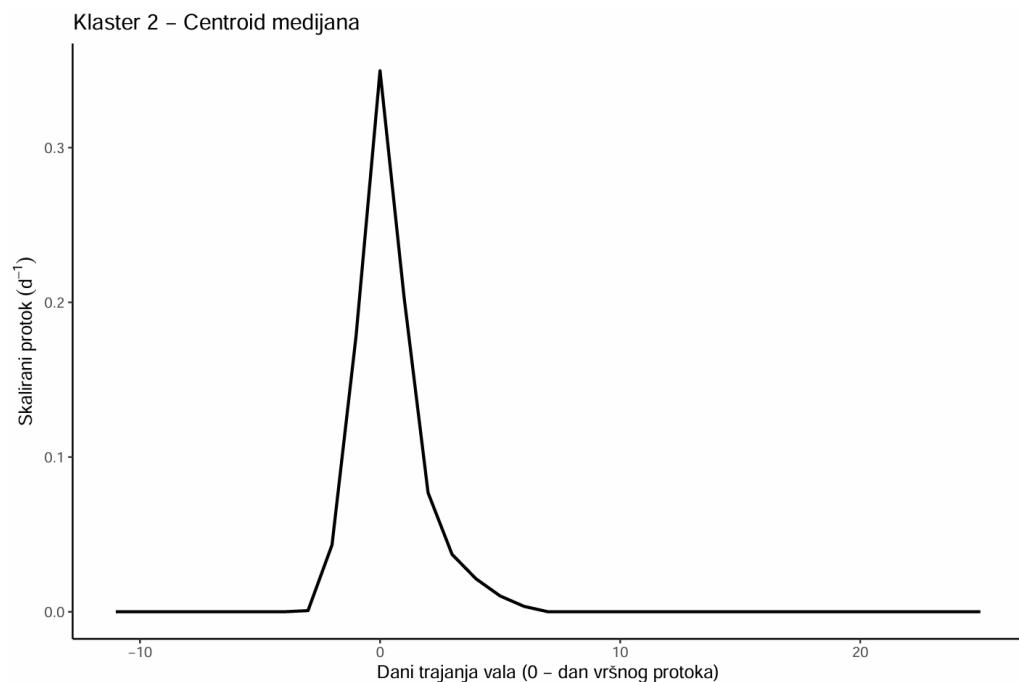
Slika 17: Oborine s kišomjerne stanice s najduljim periodom mjerenja po klasterima za rijeku Krapinu



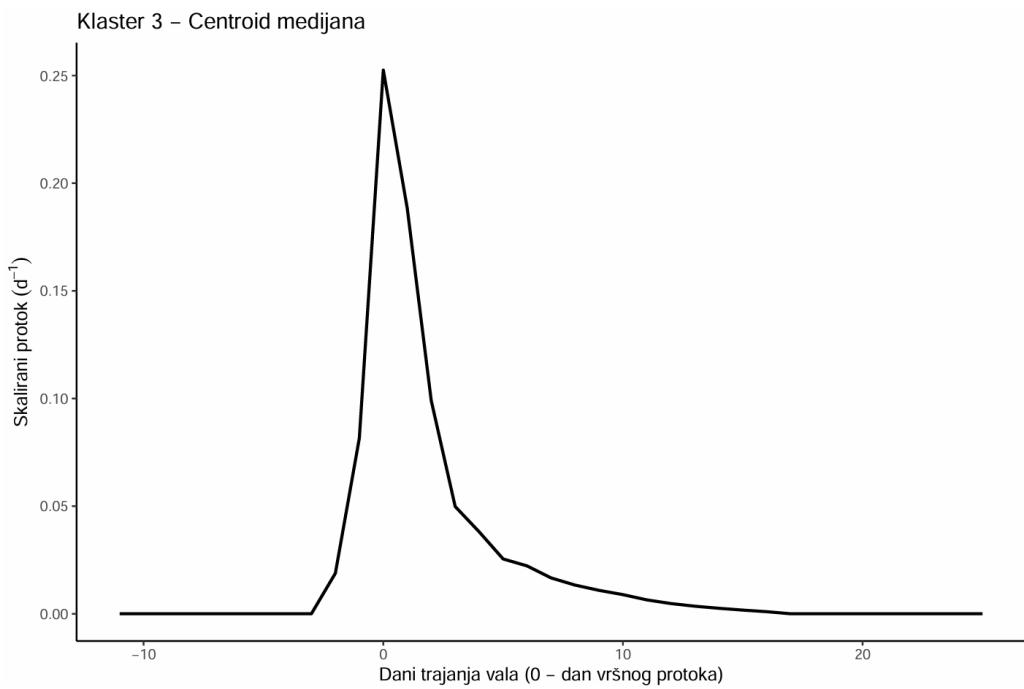
Slika 18: Težinski prosjek količine oborina svih kišomjernih stanica (metoda Thiessena) po klasterima za rijeku Krapinu



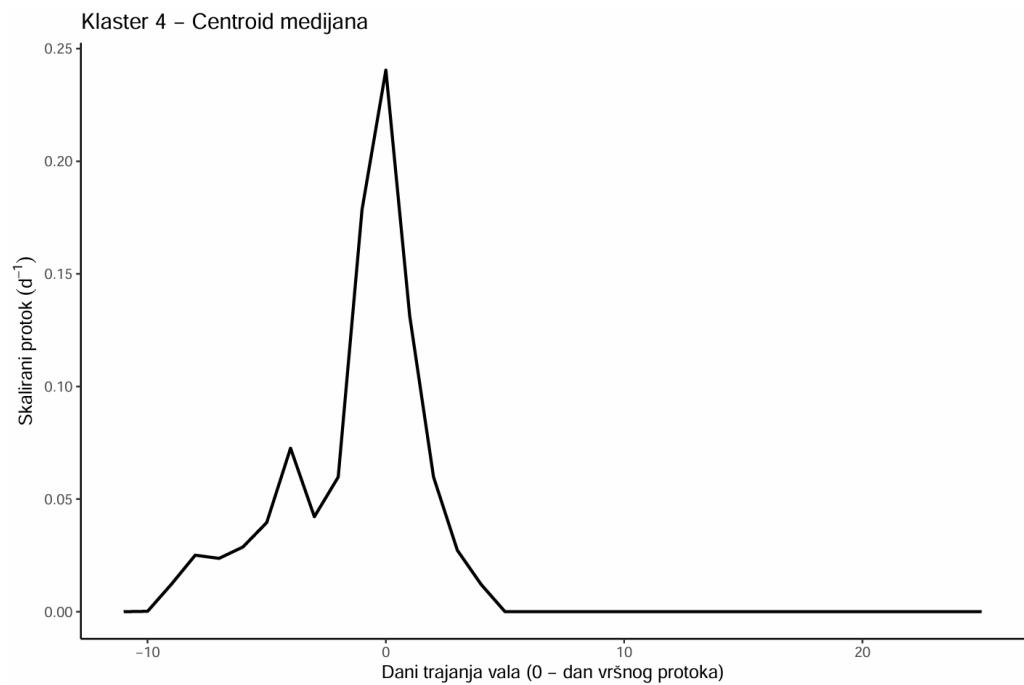
Slika 19: Normalizirani hidrogram centroida vodnog vala za 1. klaster, rijeka Krapina



Slika 20: Normalizirani hidrogram centroida vodnog vala za 2. klaster, rijeka Krapina

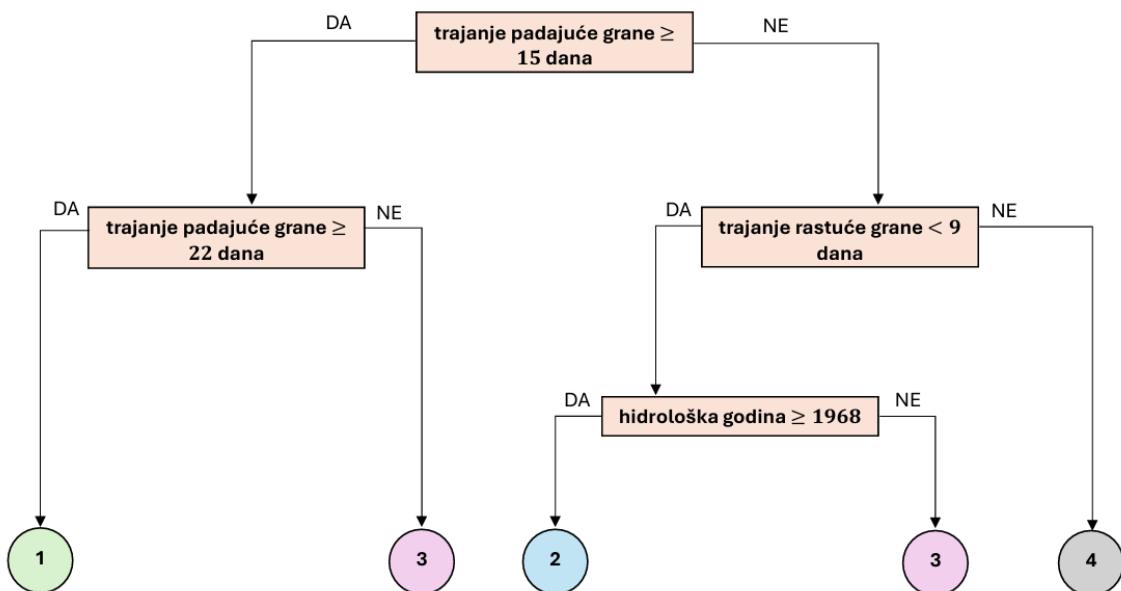


Slika 21: Normalizirani hidrogram centroida vodnog vala za 3. klaster, rijeka Krapina



Slika 22: Normalizirani hidrogram centroida vodnog vala za 4. klaster, rijeka Krapina

Za postaju Kupljenovo, nakon treniranja modela stabla odluke, dobivena je vizualizacija (Slika 23) koja jasno prikazuje kako se podaci granaju u različite klastere na temelju specifičnih kriterija. Prikazano stablo odluke pokazuje da su trajanje rastuće i padajuće grane vodnog vala te ukupno trajanje među najvažnijim parametrima za klasifikaciju. Na primjer, ako broj dana padajuće grane premašuje 15 dana, model sugerira daljnju podjelu na temelju drugih kriterija kao što je nova granica padajuće grane od 22 dana, što vodi do različitih klastera. S druge strane stabla, kriteriji za daljnju grupaciju su trajanje rastuće grane te hidrološka godina vala.



Slika 23: Prikaz stabla odlučivanja za Kupljenovo

4.2. Sliv rijeke Kupe

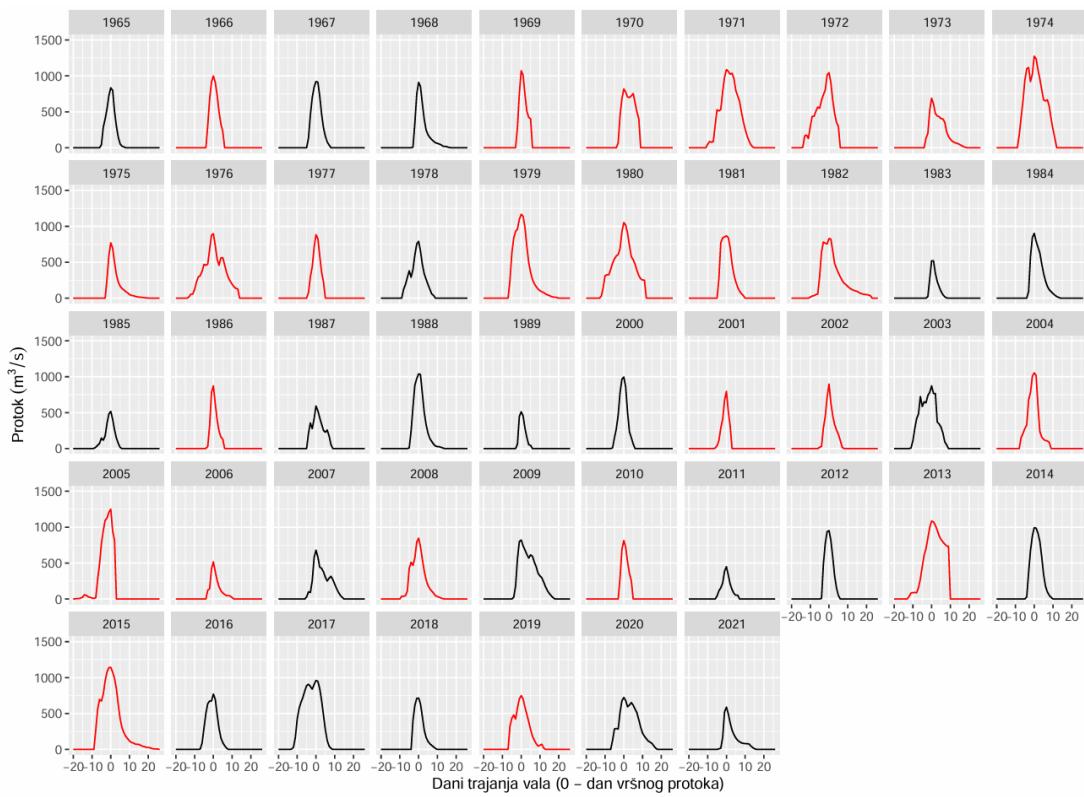
Za rijeku Kupu definirano je 47 vodnih valova, čije su karakteristike dani u tablici P2 u prilogu. Tablica 10 prikazuje statističke vrijednosti parametara za sve valove - prosječne, maksimalne i minimalne vrijednosti, standardnu devijaciju, medijan te prvi i treći kvartil.

Tablica 10: Statističke vrijednosti parametara vodnih valova na rijeci Kupi

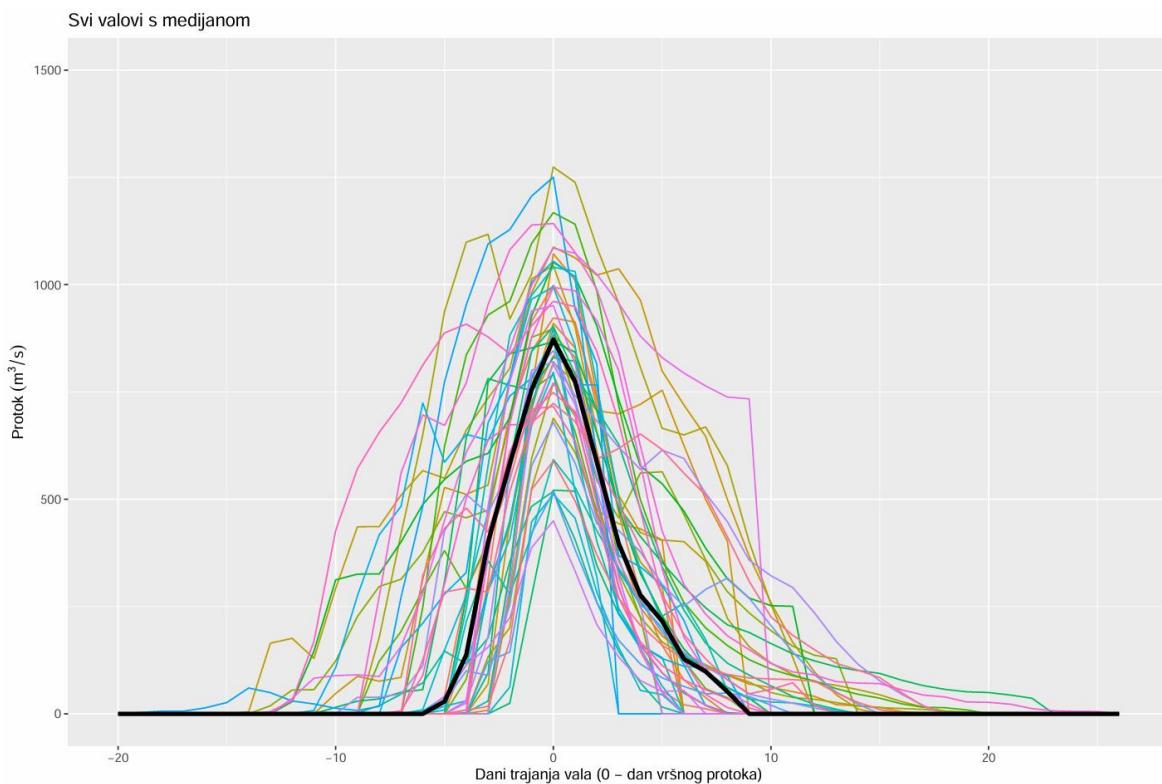
	Izmjereni protok [m ³ /s]	Direktno otjecanje [m ³ /s]	Trajanje rasta [dani]	Trajanje pada [dani]	Ukupno trajanje [dani]	Volumen vodnog vala [10 ⁶ m ³]
Prosjek	1015.68	862.62	7.06	10.72	18.79	578.14
Max	1585.00	1273.96	20	26	35	1300.95
Min	530.00	449.85	2	2	8	174.10
St. dev.	223.28	197.95	3.85	5.61	6.44	285.63
1. kvartil	882.00	748.78	5.00	6.00	14	369.23
Medijan	1001.00	872.42	6.00	9.00	19	493.12
3. kvartil	1192.00	998.83	9.00	14.00	24	781.62

Prosječni vršni protok vodnih valova rijeke Kupe je 1015.68 m³/s. Prosječni volumen vodnog vala iznosi 578.14 milijuna kubičnih metara, a prosječno trajanje 18.79 dana. Prosječno trajanje rasta iznosi 7.06 dana, a grana pada traje prosječnih 10.72 dana. Minimalni vršni protok od 530 m³/s zabilježen je 2011. godine, a maksimalni protok od 1585 m³/s zabilježen je 1974. hidrološke godine. Maksimalno trajanje vala je 35 dana iz 2015. hidrološke godine, a najmanjeg ukupnog trajanja je val iz od 8 dana. Trajanje grane rasta je u rasponu od minimalno 2 dana (1975. hidrološka godina) do maksimalno 20 dana (2005.), a grane pada od minimalno 2 dana (2001.) do maksimalno 26 dana (2015. hidrološka godina). Val s minimalnim pripadnim volumenom je iz 1989. godine (174.1 milijuna m³), dok je val s maksimalnim volumenom (1300.95 milijuna m³) ujedno i val s najvećim protokom iz 1974. hidrološke godine.

Slika 24 prikazuje sve definirane valove na rijeci Kupi prema ukupnom protoku. Na slici su različito prikazani vodni valovi kod kojih su detektirani te „odrezani“ sekundarni valovi (crveno) te valovi kod kojih nisu bili prisutni sekundarni zavisni valovi (crno). Na nešto manje od polovice valova, njih 22, nisu uočeni sekundarni valovi, dok je kod preostalih 25 bio prisutan nezavisni sekundarni val.



Slika 24: Centrirani hidrogrami vodnih valova na rijeci Kupi, crvenom bojom prikazani su valovi kod kojih su odrezani nezavisni sekundarni valovi



Slika 25: Centrirani hidrogrami vodnih valova na rijeci Kupi, crnom bojom prikazan je medijan



Slika 26: Centrirani i normalizirani hidrogrami vodnih valova na rijeci Kupi

Na temelju pretraživanja mreže (engl. *grid search*) za sliv rijeke Kupe određeno je da je najstabilnija kombinacija je Bspline rang 4 uz 6 klastera, indeks siluete iznosi 0.56 (Tablica 11). Navedena kombinacija je i na temelju vizualne procjene izabrana valova po klasterima izabrana kao odgovarajuća. Provedena je analiza stabilnosti indeksa siluete na podskupovima podataka s rijeke Kupe (1965-1989 i 2000-2021) te je uočeno da se rezultati ne poklapaju u potpunosti, u prvoj polovici najstabilnije je 6 klastera (indeks siluete 0.54), a u drugoj polovici najstabilnija su 2 klastera (0.65). Ipak, u oba slučaja vrijednosti indeksa su slične i prihvatljive za 2 do 6 klastera za rang 4 te je potrebna vizualna procjena. Treba istaknuti i nedostatak podatcima o dnevnim protocima od 1990. do 1999. godine uz mogućnost nehomogenosti podataka nakon ponovne uspostave mjerjenja.

Niže su dani rezultati klasteriranja vodnih valova na Kupi u 6 klastera. Tablica 12 prikazuje broj valova raspoređenih u pojedini klaster, Slika 27 prikazuje normalizirane valove po klasterima sa crno iscrtanim centroidom za klaster, a Slika 28 valove prema protoku po klasterima. Pregled parametara valova po klasterima dan je u Tablica 13, a oborine za valove po klasterima u Tablica 14.

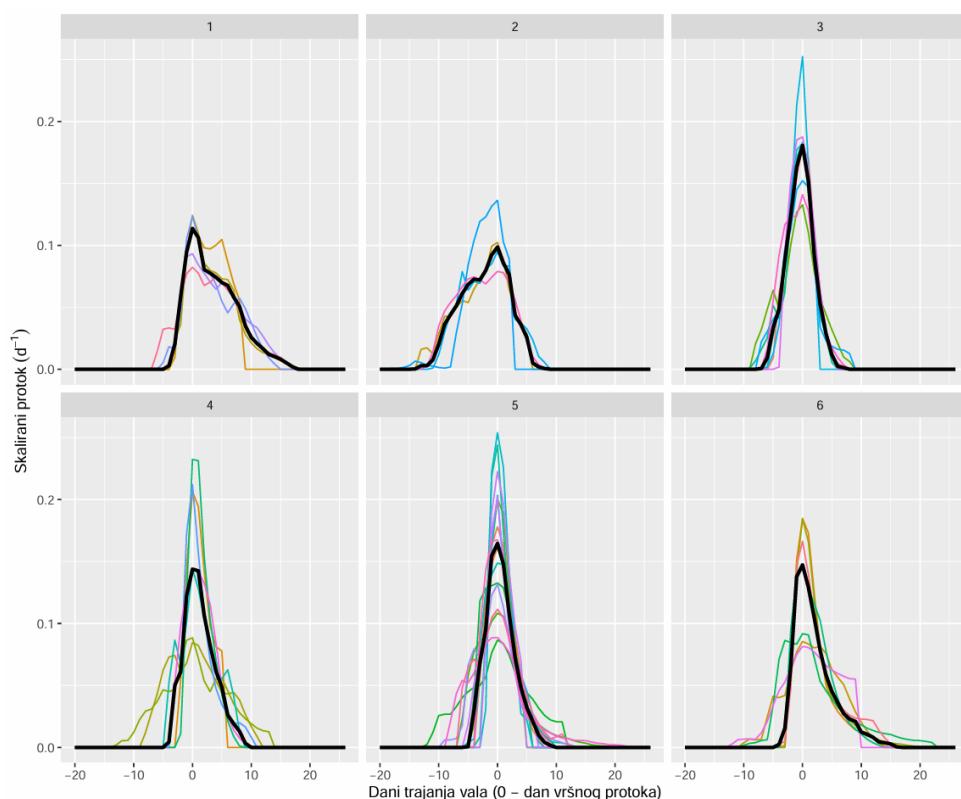
Tablica 11: Indeks siluete za različite kombinacije ranga krivulje i broja klastera, Farkašić

Rang Bspline krivulje	Broj centroida u K-means	Indeks siluete
4	2	0.53910
4	3	0.53445
4	4	0.55174
4	5	0.55772
4	6	0.56256
10	2	0.51663
10	3	0.39442
10	4	0.36514
10	5	0.38992
10	6	0.38280
15	2	0.43383
15	3	0.38245
15	4	0.31532
15	5	0.29546
15	6	0.29621
20	2	0.35705
20	3	0.31467
20	4	0.27262
20	5	0.25204
20	6	0.26909
25	2	0.30364
25	3	0.27464
25	4	0.24840
25	5	0.24295
25	6	0.24949

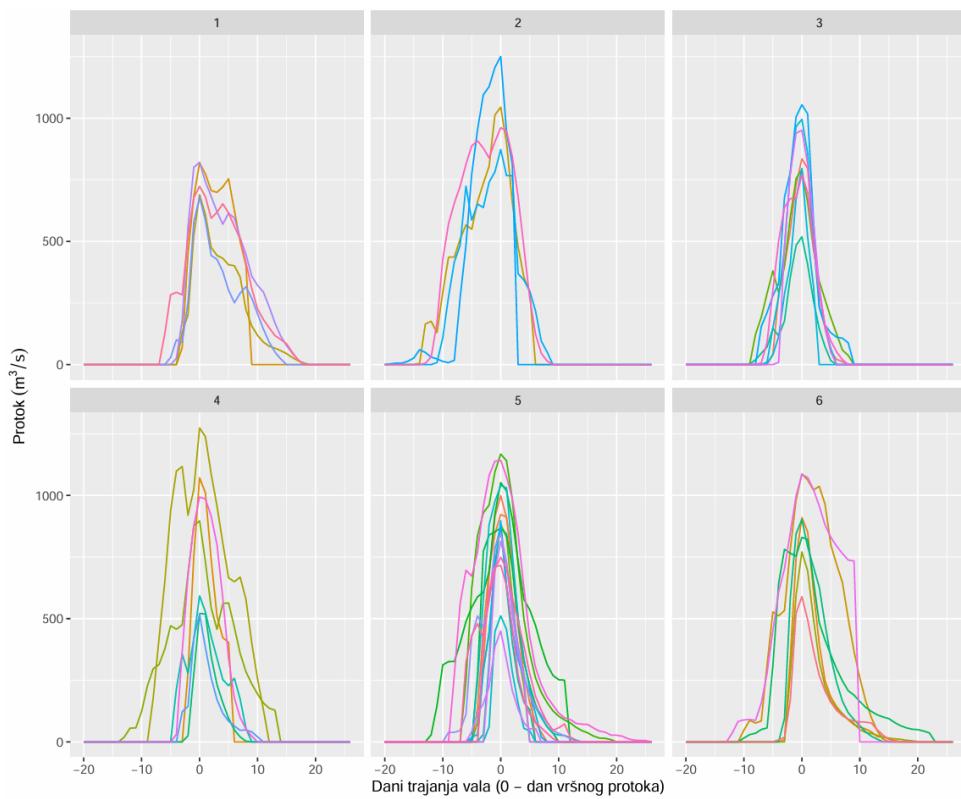
Najveći broj valova, otprilike trećina odnosno njih 16 od 47, svrstano je u 5. klaster. Slijedi 4. klaster s 8 valova, 4. i 6. s po 7 valova. Najmanje valova svrstano je u 1. klaster (5 valova) i 2. klaster (4 vala).

Tablica 12: Broj valova po klasterima za hidrograme na rijeci Kupi

Klaster	Broj valova	Udio [%]
1	5	10.6
2	4	8.5
3	8	17.0
4	7	14.9
5	16	34.0
6	7	14.9



Slika 27: Normalizirani hidrogrami vodnih valova na rijeci Kupi po klasterima, crnom bojom prikazan centroid klastera



Slika 28: Hidrogrami vodnih valova na rijeci Kupi po klasterima

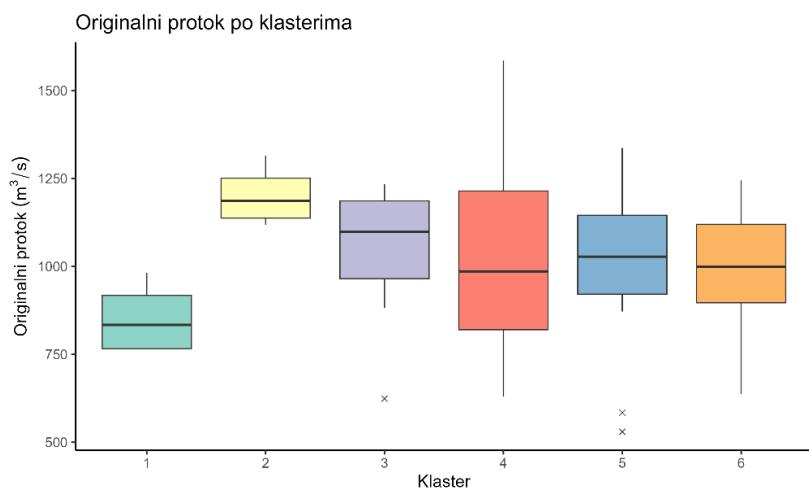
Tablica 13: Pregled statistike valova po klasterima za hidrograme na rijeci Kupi

	Klaster	Izmjereni protok [m ³ /s]	Direktno otjecanje [m ³ /s]	Trajanje rasta [dani]	Trajanje pada [dani]	Ukupno trajanje [dani]	Volumen vodnog vala [10 ⁶ m ³]
Prosjek	Svi valovi	1015.68	862.62	7.06	10.72	18.79	578.14
	1	853.00	746.11	5.00	15.60	21.60	619.42
	2	1202.00	1032.25	15.25	6.25	22.50	880.33
	3	1041.00	839.02	7.63	6.63	15.25	427.08
	4	1038.71	838.12	5.86	9.86	16.71	569.09
	5	1009.31	870.61	5.88	10.19	17.06	531.27
Max	1	982.00	821.00	7.00	19.00	26.00	760.87
	2	1315.00	1250.53	20.00	9.00	24.00	1052.09
	3	1234.00	1054.52	10.00	9.00	20.00	598.65
	4	1585.00	1273.96	14.00	13.00	28.00	1300.95
	5	1337.00	1167.94	13.00	26.00	35.00	1112.95
	6	1245.00	1087.66	13.00	22.00	34.00	1150.07
Min	1	765.00	679.36	3.00	8.00	13.00	475.60
	2	1119.00	872.57	12.00	2.00	21.00	792.76
	3	624.00	518.46	5.00	2.00	10.00	244.88
	4	630.00	516.74	3.00	5.00	9.00	193.73
	5	530.00	449.85	3.00	4.00	8.00	174.10
	6	638.00	589.12	2.00	9.00	19.00	306.06
St. dev.	1	95.58	68.99	1.58	4.51	5.03	141.31
	2	89.08	161.59	3.40	3.40	1.29	121.88
	3	208.95	166.95	1.77	2.20	3.20	118.87
	4	325.18	299.00	4.02	2.61	6.13	400.05
	5	220.03	198.79	2.47	6.10	7.58	281.70
	6	216.24	175.88	4.14	4.43	4.61	349.89

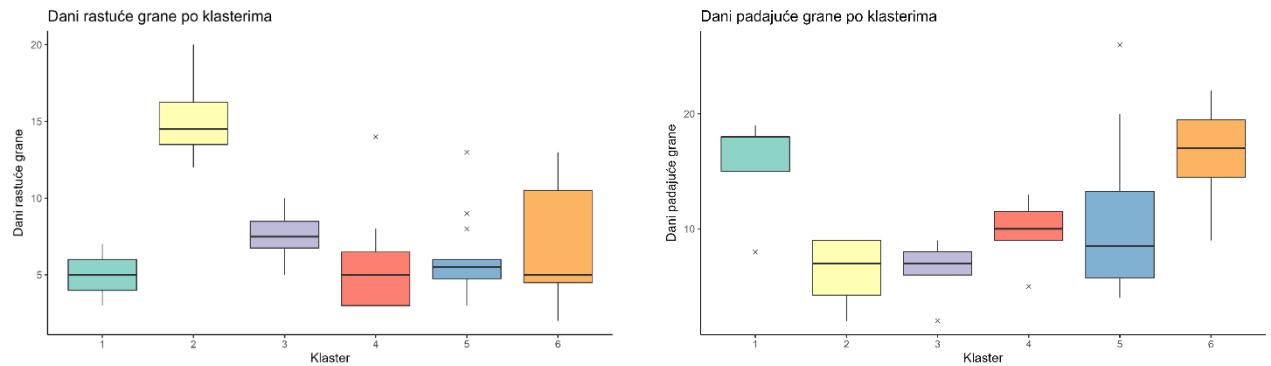
Tablica 14: Pregled oborina po klasterima za hidrograme na rijeci Kupi

	Klaster	Oborine - postaja s najduljim nizom [mm]	Oborine - težinski prosjek svih postaja [mm]
Prosjek	1	49.12	76.86
	2	118.48	79.10
	3	71.86	84.31
	4	92.73	139.55
	5	73.89	104.73
	6	72.01	159.79
Max	1	91.20	94.57
	2	149.90	79.10
	3	103.80	88.51
	4	164.60	139.55
	5	170.20	181.05
	6	121.00	159.79
Min	1	8.00	47.17
	2	55.70	79.10
	3	32.90	80.11
	4	32.30	139.55
	5	10.80	70.75
	6	23.80	159.79
St. dev.	1	33.04	25.87
	2	42.90	-
	3	27.84	5.94
	4	57.31	-
	5	33.69	39.71
	6	35.38	-

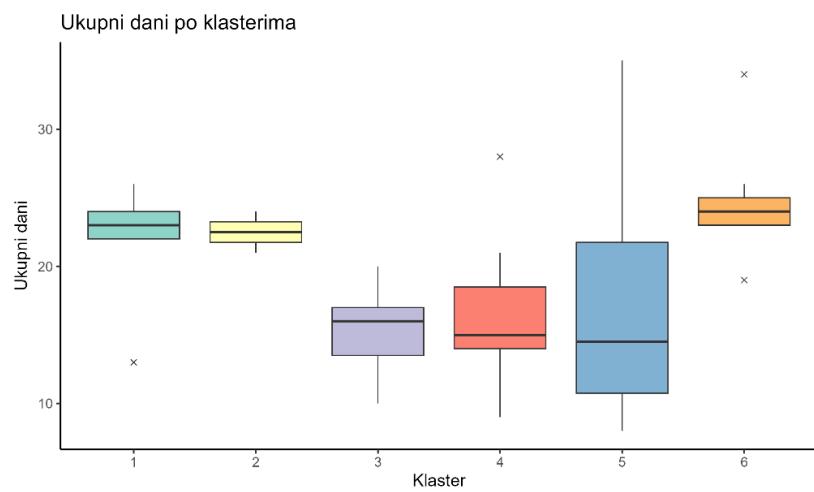
U nastavku su dani kutijasti dijagrami (Slika 29 do Slika 34) koji prikazuju raspone parametara po klasterima. Iz navedenih dijagrama vidljivo je veliko preklapanje vrijednosti parametara za protok, međutim za trajanje rasta, pada i ukupno trajanje, te volumen i količinu oborina vidljive su razlike među klasterima. Navedena vizualna procjena djelomično je potvrđena i analizom varijance, čiji su rezultati pokazali da je razlika trajanja (ukupnog, grane rasta i pada) među klasterima statistički vrlo značajna (na razini 0.001), no razlika volumena vodnih valova i količine oborina među klasterima nije statistički značajna. Analiza varijance nije pokazala statistički značajnu razliku protoka po klasterima. Za svaki klaster prikazan je i normalizirani hidrogram centroida svih vodnih valova koji pripadaju tom klasteru (Slika 35 do Slika 40), a centroidi hidrograma po klasterima s originalnim protokom dani su u prilogu (Slika P5 do P10).



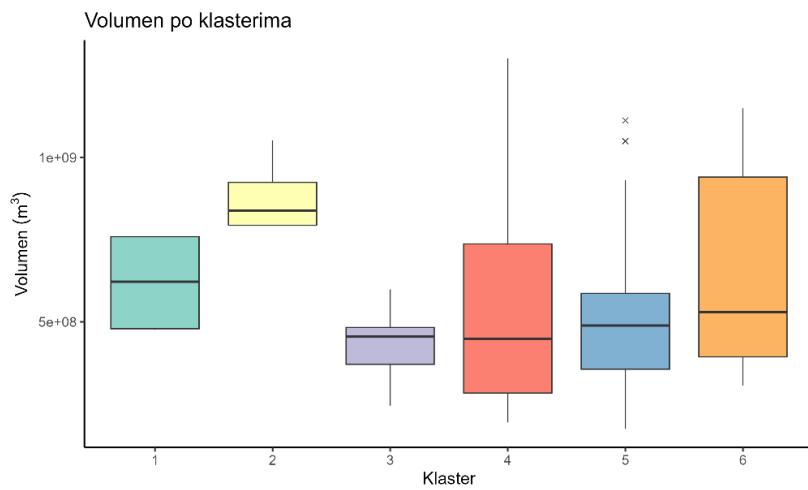
Slika 29: Vršni protok po klasterima za rijeku Kupu



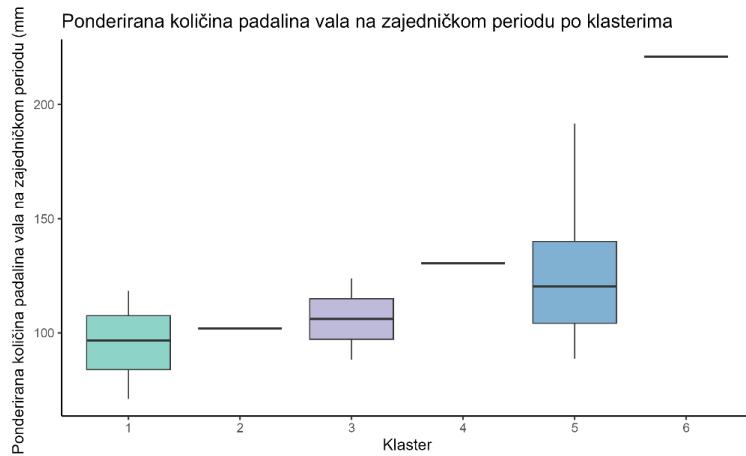
Slika 30: Trajanje grane rasta (lijevo) i pada (desno) vodnog vala po klasterima za rijeku Kupu



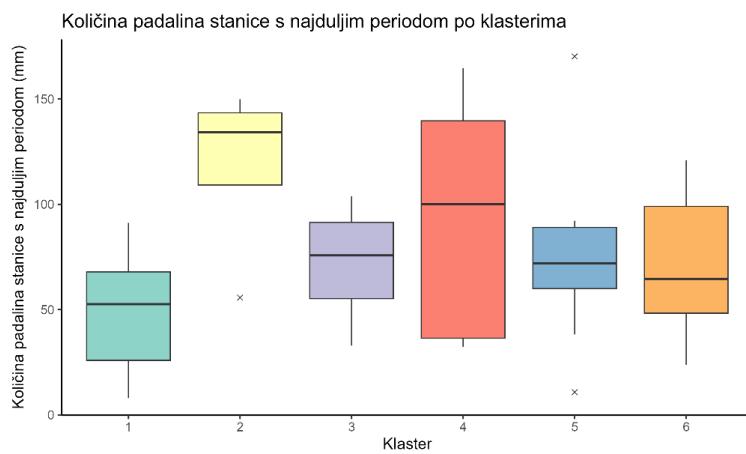
Slika 31: Ukupno trajanje vodnog vala po klasterima za rijeku Kupu



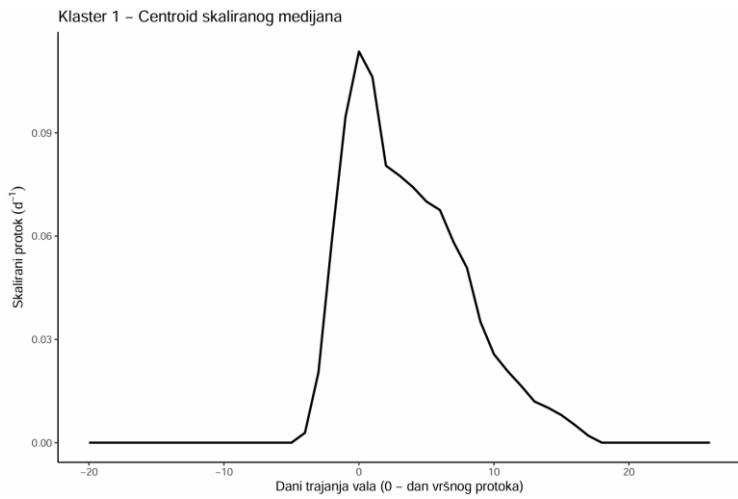
Slika 32: Volumen vodnih valova po klasterima za rijeku Kupu



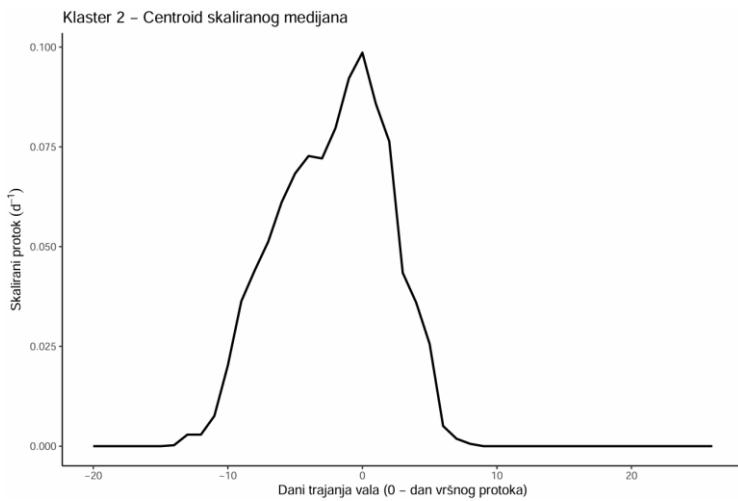
Slika 33: Oborine s kišomjerne stanice s najduljim periodom mjerena po klasterima za rijeku Kupu



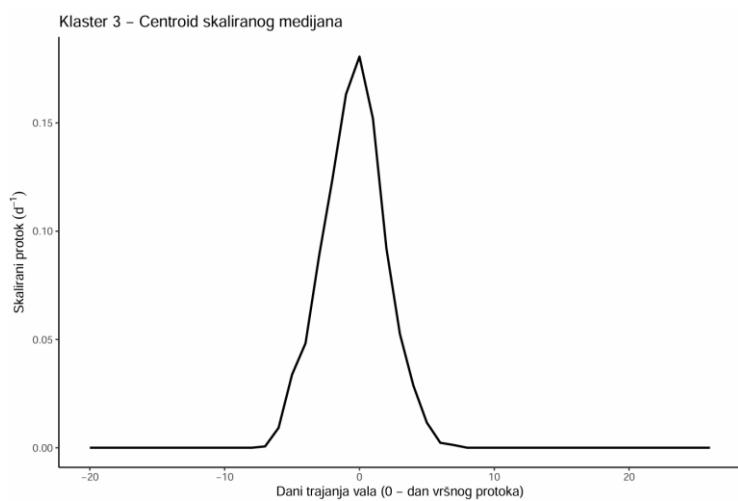
Slika 34: Težinski prosjek količine oborina svih kišomjernih stanica (metoda Thiessena) po klasterima za rijeku Kupu



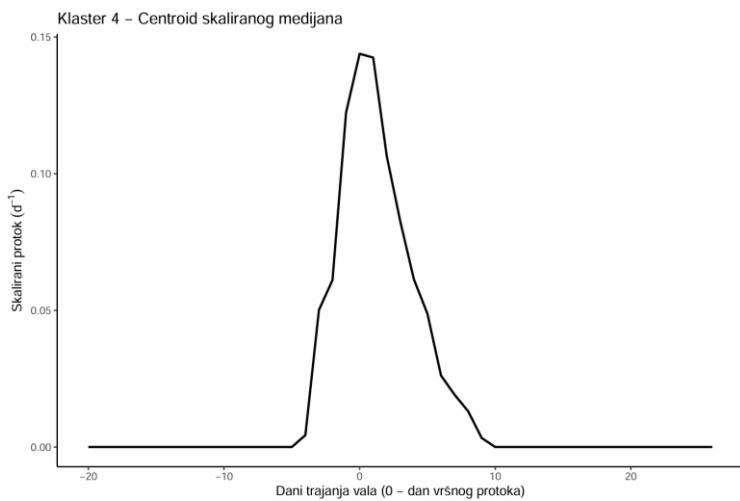
Slika 35: Hidrogram centroida vodnog vala za 1. klaster, rijeka Kupa



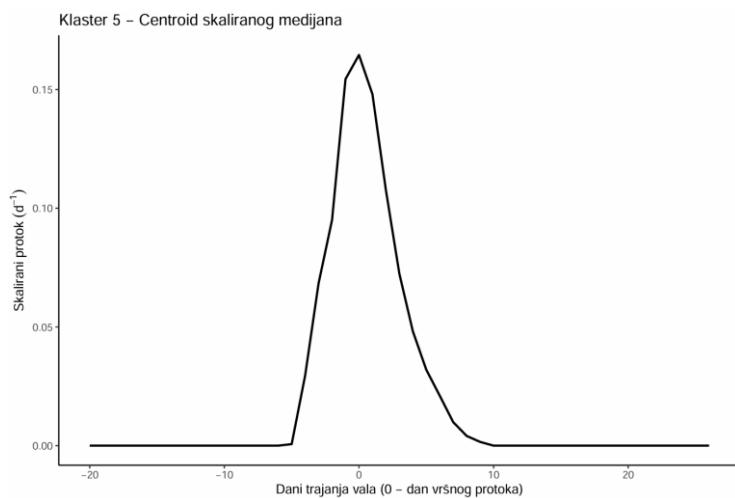
Slika 36: Hidrogram centroida vodnog vala za 2. klaster, rijeka Kupa



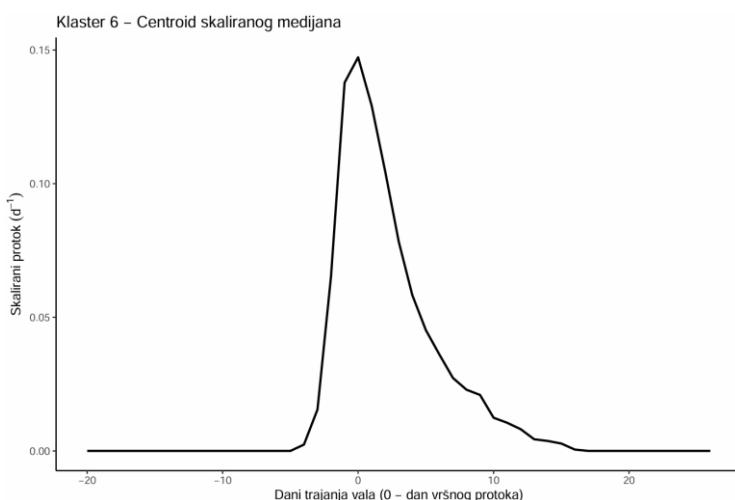
Slika 37: Hidrogram centroida vodnog vala za 3. klaster, rijeka Kupa



Slika 38: Hidrogram centroida vodnog vala za 4. klaster, rijeka Kupa

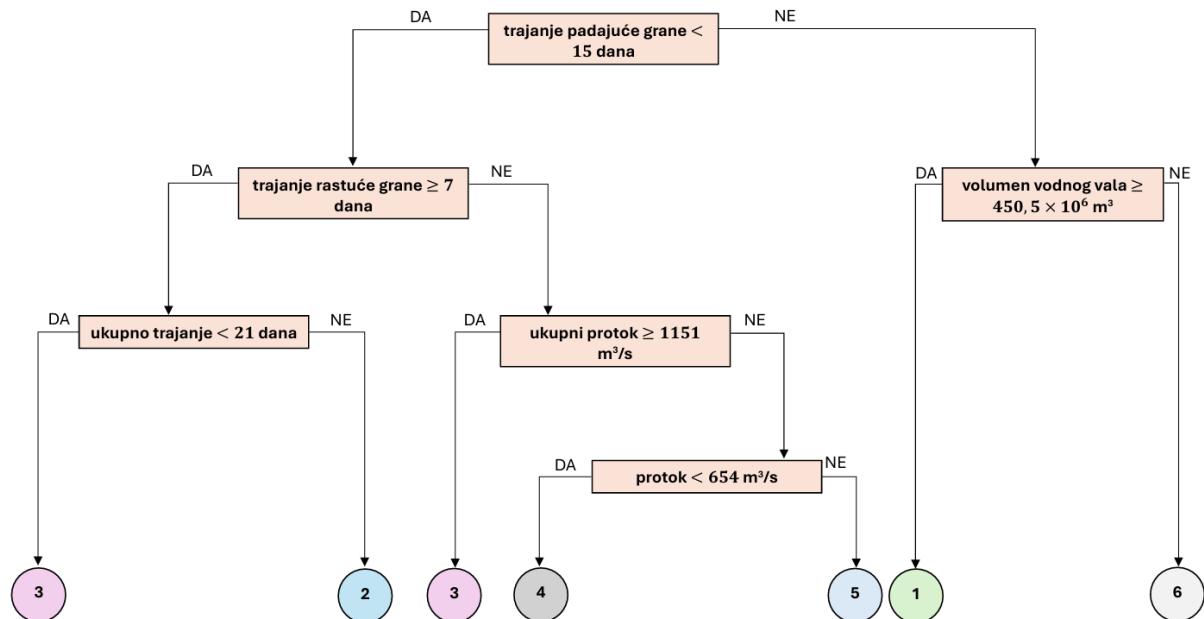


Slika 39: Hidrogram centroida vodnog vala za 5. klaster, rijeka Kupa



Slika 40: Hidrogram centroida vodnog vala za 6. klaster, rijeka Kupa

Za postaju Farkašić također je istrenirano i vizualizirano stablo odlučivanja (Slika 41). Uz značajke trajanja rastuće i padajuće grane te ukupnog trajanja uzima se u obzir i volumen te protok kao dodatne kriterije za kategorizaciju u klastere. To ima smisla iz razloga što postoji i veći broj različitih labela na koje se podaci dijele, pa mora postojati i veća razgranatost samog stabla uz više različitih kriterija za grupaciju.



Slika 41: Prikaz stabla odlučivanja za Farkašić

5 RASPRAVA

Slivovi rijeka Krapine i Kupe razlikuju se prema svojim morfološkim karakteristikama kao i po hidrološkim procesima koji se na njima odvijaju, stoga je prije usporedbe i analize rezultata klasteriranja potrebno usporediti promatrane slivove i vodne valove zabilježene na njima. Sliv Kupe do postaje Farkašić je 7.82 puta veće površine od sliva rijeke Krapine do postaje Kupljenovo (8992 km^2 prema 1150 km^2). U skladu s odnosom površina ponaša se i odnos prosječnog protoka definiranih vodnih valova - prosječni vršni protok na rijeci Kupi 7.86 puta je veći od prosječnog vršnog protoka na rijeci Krapini. Trajanje, kako ukupno tako i grane rasta i pada, nešto je dulje na rijeci Kupi nego na Krapini. Rast traje prosječno dvostruku dulje na Kupi, pad traje prosječno 1 dan duže, dok je ukupno trajanje vala na Kupi otprilike 4 dana duže nego na Krapini. Prosječni volumen vodnog vala na rijeci Kupi gotovo je 16 puta veći od volumena vodnih valova na Krapini. Navedene razlike, osim površine sliva, mogu proizlaziti i iz njihovog oblika, topografije i rasporeda riječne mreže. Sliv rijeke Kupe obilježava velika visinska razlika najviše i najniže točke - od iznad 1600 m n.m. u planinama Gorskog kotara do otprilike 90 m na izlazu iz sliva. Slive rijeke Krapine s druge strane ima manju razliku - najviši dijelovi iznad 1000 m n.m. nalaze se na Ivanšćici i Medvednici, izlaz sa sliva je na oko 130 m n.m., a slivom prevladava brežuljkasti teren nadmorske visine 150 do 350 m. Nadalje, slive rijeke Krapine obilježavaju brojni kraći pritoci, rječice i potoci dok su u slivu rijeke Kupe prisutni i značajniji pritoci poput Mrežnice, Dobre i Korane. Uzimajući u obzir navedene razlike, realno je očekivati veću raznolikost oblika hidrograma na rijeci Kupi, a u skladu s tim i njihovo svrstavanje u veći broj klastera.

S obzirom da je klasteriranje provođeno na temelju oblika, na normaliziranim hidrogramima, logično da nije uočeno grupiranje vodnih valova po protocima u klastere ni za Krapinu ni za Kupu, već su prosječne vrijednosti protoka svih klastera za pojedini sliv slične. S druge strane, kako su glavne značajke koje opisuju hidrogram trajanje rasta i pada te površina ispod krivulje/volumen vodnog vala, u skladu s očekivanjem je da su navedene značajke ispale statistički značajno različite po klasterima.

Klasteri hidrograma na rijeci Krapini mogu se opisati na sljedeći način: 1. klaster - kratko trajanje grane rasta, duga i nepravilnija grana pada, najdulje ukupno trajanje, veliki raspon volumena, manja količina oborina na slivu koja prethodi vršnom protoku; 2. klaster - pravilne i kratkotrajne grane rasta i pada uz najkraće ukupno trajanje, relativno manji volumen, srednja količina oborina; 3. klaster - kratko trajanje grane rasta, srednje trajanje grane pada

(koja je blaža nego u 2. klasteru), relativno manji volumen, srednja količina oborina; 4. klaster - duga i nepravilna grana rasta, brzi i nagli pad, veći volumen vodnog vala, najveće oborine. Treba uzeti u obzir veliku razliku broja valova po klasterima, od tek 4 vala za 4. klaster do 35 valova za 2. klaster. Prema pretraživanju mreže (engl. *grid search*) provedenom prije samog klasteriranja optimalan broj klastera za vodne valove definirane za rijeku Krapinu bio bi 2 klastera, no tako dobiveni klasteri manje bi se međusobno razlikovali po parametrima trajanja i volumena te se ne bi dobili značajno različiti oblici hidrograma između kojih bi se moglo birati pri kasnijoj izradi projektnog hidrograma. S obzirom na manju površinu sliva rijeke Krapine te njegove relativno uniformne morfološke i klimatske karakteristike, postavlja se pitanje potrebe i mogućnosti definiranja različitih oblika hidrograma. Radi se o slivu koji „brzo reagira“ na oborinske događaje te bi se većina valova mogla aproksimirati jednim oblikom hidrograma, no klasteriranjem u 4 klastera ipak su uzete u obzir manje razlike među valovima koje se prema potrebi mogu uzeti u obzir pri definiranju projektnog hidrograma, odnosno projektiranju.

Za razliku od rijeke Krapine, razlike klastera za hidrograme na rijeci Kupi nisu statistički značajne za parametre volumena vodnog vala i količine oborina. Prema tome se klasteri na rijeci Kupi mogu opisati na sljedeći način promatrujući trajanje: 1. klaster - kratko trajanje grane rasta, dugo trajanje padajuće grane, dugo ukupno trajanje; 2. klaster - dugo trajanje rastuće grane, kratki pad, dugo ukupno trajanje; 3. klaster - srednje trajanje rasta, kratki pad, kratko ukupno trajanje; 4. klaster - kratki rast, srednje dugo trajanje padajuće grane, kratko ukupno trajanje; 5. klaster - kratki rast, srednje kratak pad, srednje ukupno trajanje; 6. klaster - kratki do srednji rast, najduže trajanje grane pada, najduže ukupno trajanje. I ovdje je, iako u manjoj mjeri, prisutna razlika u broju valova po klasterima, od 4 vala za 2. klaster do 16 valova za 5. klaster. Upravo je kod 5. klastera s najvećim brojem valova najveće rasipanje vrijednosti promatranih parametara. Podaci o oborinama ovdje su manje uporabivi jer za najduže razdoblje za koje su dostupni podaci svih promatranih kišomjernih postaja nema valova koji pripadaju 3 od 6 klastera. Podaci s meteorološke postaje s najdužim trajanjem mjerenja (Karlovac), ne mogu dovoljno dobro reprezentirati neujednačenost oborine pale na cijeli sliv s obzirom na već ranije opisane značajke sliva rijeke Kupe.

6 ZAKLJUČCI

Provedeno je klasteriranje vodnih valova na rijeci Krapini i Kupi po obliku hidrograma korištenjem metoda strojnog učenja. Funkcijsko klasteriranje hidroloških podataka, kao što je ovdje primijenjeno, predstavlja relativno efikasan alat za analizu složenih hidroloških procesa. Kombinacija Bspline krivulja za aproksimaciju podataka i K-means algoritma za klasteriranje omogućuje preciznu i fleksibilnu segmentaciju hidrograma prema njihovim oblikovnim karakteristikama. Međutim, postoji prostor za poboljšanje i daljnji razvoj ove metodologije, posebice primjenom tehnika dubokog učenja kao što su rekurentne neuronske mreže (RNN), koje su dizajnirane za rad s vremenskim nizovima i sekvencijskim podacima, ili razvojem hibridnih modela kombiniranjem metoda nadziranog i nenadziranog strojnog učenja kako bi se u potpunosti iskoristili dostupni podaci te povećala interpretabilnost rezultata. Također, poboljšanje predstavlja i detaljnije pretraživanje mreže uz veći broj kombinacija parametra te provjeru stabilnosti klastera pomoću nekoliko različitih mjera uz već korišten koeficijent siluete.

Za svaki sliv na temelju dnevnih podataka o protocima definirani su vodni valovi metodom godišnjih maksimuma. Za sliv rijeke Krapine dobivena su 4 klastera koja se međusobno statistički razlikuju po parametrima trajanja rastuće i padajuće grane, ukupnog trajanja, volumena vodnog vala te količine oborine koja prethodi danu vršnog protoka vodnog vala. Za sliv rijeke Kupe definirano je 6 klastera između kojih postoji statistički značajna razlika u trajanju rastuće i padajuće grane te ukupnog trajanja. Dobiveni set klastera za Krapinu i Kupu predstavljaju dobru polazišnu točku i dovoljno raznolik izbor za eventualno definiranje sintetičkog projektnog hidrograma za promatrane slivove. Takav pristup predstavlja iskorak u odnosu na još uvijek prevladavajuću praksu u hidrotehnici, gdje se u projektiranju najčešće koriste samo jednoznačno definirane vrijednosti maksimalnog protoka i volumena vodnog vala, čime se zanemaruje različitost velikovodnih događaja na slivu. Glavna prednost korištenja metodologije iz ovog rada kao koraka prema definiranju projektnog hidrograma je raznolikost oblika hidrograma koji predstavljaju različite uvjete na slivu i varijabilnost stvarnih događaja. Osim toga, postupak je lagano prilagoditi za druge slivove te se rezultati mogu dobiti relativno brzo. Međutim, kvaliteta ulaznih podataka i njihova dostupnost značajno utječe na pouzdanost rezultata klasteriranja. Radi se o metodi koja zahtjeva stručno znanje, kako pri izboru broja klastera tako i u potencijalnom naknadnom izboru odgovarajućeg oblika hidrograma. Izbor previše klastera može rezultirati oblicima koji su temeljeni na premalom broju vodnih valova, a premalo klastera može uzrokovati gubitak raznolikosti između oblika.

Daljnji koraci kojima bi se došlo do sintetičkih projektnih hidrograma iz ovdje predstavljenih oblika hidrograma po klasterima su određivanje maksimalnog protoka korištenjem određenih vjerojatnosnih raspodjela te skaliranje oblika hidrograma na tu vrijednost protoka. Pri definiranju vodnih valova umjesto metode godišnjih maksimuma vrijedilo bi isprobati metodu premašenja praga jer su za sušne godine godišnjim maksimumima dobiveni relativno mali vodni valovi, a zanemareni su neki veći vodni valovi u godinama s više velikovodnih događaja. Osim toga, metodologiju bi vrijedilo primijeniti na dužem nizu podataka ili na nizu s više definiranih valova kako bi se dobio veći broj podataka, odnosno vodnih valova po klasterima. Metodologija iz ovog rada mogla bi se primijeniti i na nizu drugih slivova na prostoru Republike Hrvatske, a u slučaju neizučenih slivova, odnosno rijeka na kojima nisu provođena mjerenja, za regionalnu analizu koriste se hidrogrami dobiveni za slivove s dostupnim mjeranjima. Proširenje ovog rada na ostale slivove u RH i definiranje više tipova sintetičkih projektnih hidrograma za svaki sliv stvorilo bi bazu oblika projektnih hidrograma koji bi mogli pomoći u projektiranju hidrotehničkih građevina različitih namjena, a dajući projektantu fleksibilnost izbora prema relevantnim parametrima. Moguća je i nadogradnja metodologije korištenjem podataka o oborinama dobivenih na temelju daljinskih mjeranja (radari, sateliti), povezivanjem oblika s podacima o štetama uzrokovanim pojedinim valovima te uzimanjem u obzir nesigurnosti podataka.

7 ZAHVALE

Zahvaljujemo kolegici Martini Lacko, mag. ing. aedif., na savjetima prilikom izrade rada te Državnom hidrometeorološkom zavodu na ustupljenim podacima u zatraženom obimu i formatu.

8 POPIS LITERATURE

- [1] Gądek W, Baziak B, Tokarczyk T, Szalińska W. A novel method of design flood hydrographs estimation for flood hazard mapping. *Water*. 2022;14:1856. Dostupno na: <https://doi.org/10.3390/w14121856>
- [2] Brunner M.I., Viviroli D., Furrer R., Seibert J., Favre A.-C. Identification of flood reactivity regions via the functional clustering of hydrographs. *Water Resources Research*. 2018;54:1852-1867. Dostupno na: <https://doi.org/10.1002/2017WR021650>
- [3] Brunner M.I., Viviroli D., Sikorska A.E., Vannier O., Favre A.C., Seibert J. Flood type specific construction of synthetic design hydrographs. *Water Resources Research*. 2017;53(2):1390-1406. doi:10.1002/2016WR019535
- [4] Žugaj R. *Hidrologija*. Zagreb: Rudarsko-geološko-naftni fakultet; 2000.
- [5] Bezak N., Lebar K., Mikoš M., Rusjan S., Šraj M. Razvoj metodologije za izračun visokovodnih valov na podlagi ekstremnih padavinskih dogodkov. *Končno poročilo*. Ljubljana: Univerza v Ljubljani, Fakulteta za gradbeništvo in geodezijo; 2023.
- [6] Wilson D., Fleig A.K., Lawrence D., Hisdal H., Pettersson L.-E., Holmqvist E. A review of NVE's flood frequency estimation procedures. *Report no. 9 – 2011*. Oslo: Norwegian Water Resources and Energy Directorate; 2011.
- [7] Fischer S., Schumann A.H. Generation of type-specific synthetic design flood hydrographs. *Hydrological Sciences Journal*. 2023;68(7):982-997. doi:10.1080/02626667.2023.2195560
- [8] Whipple A.A., Viers J.H., Dahlke H.E. Flood regime typology for floodplain ecosystem management as applied to the unregulated Cosumnes River of California, United States. *Ecohydrology*. 2017;e1817. doi:10.1002/eco.1817
- [9] Lacko M., Potočki K., Pintar D., Humski L., Bojanjac D. The applicability of functional clustering in analyzing historical floods of the Sava River in Zagreb. U: Lončarić S.,

Šmuc T., urednici. *Abstract Book, Sixth International Workshop on Data Science*. Zagreb: Centre of Research Excellence for Data Science and Cooperative Systems, Research Unit for Data Science; 2021. p. 67-69.

- [10] Vodoprivredno-projektni biro d.d. Sustav obrane od poplava Srednjeg Posavlja: Aktualizacija rješenja, SAŽETAK - verzija 1. Zagreb: Vodoprivredno-projektni biro d.d.; 2012.
- [11] Bekić D., Kadić V., Ivezić V., Kerin I. Flood hazard maps for the Krapina River Basin. U: Proceedings of the 6th Croatian Water Conference with International Participation; Svibanj 2015; Opatija; 2015.
- [12] Biondić B., Biondić R., Kapelj S. Karst groundwater protection of the Kupa River catchment area and sustainable development. *Environmental Geology*. 2006; doi:10.1007/s00254-006-0178-4
- [13] Bombač M., Duh L., Filipan S., Mlačnik J., Petan S., Polić S., Potočnik B., Prešeren T., Rački V., Šlehta T., Vošnjak S., Vučak Ž., Žagar Kopitar M. Projekt FRISCO1 prekogranično usklađeno slovensko-hrvatsko smanjenje rizika od poplava – negrađevinske mjere. Ljubljana: Ministarstvo okoliša i prostornog planiranja Republike Slovenije; 2019.
- [14] Wickham H. *Package ‘tidyverse’*. Dostupno na: <https://tidyverse.tidyverse.org>, <https://github.com/tidyverse/tidyverse>. [Pristupljeno: 28. svibnja 2024.].
- [15] Spinu V., Grolemund G., Wickham H., Vaughan D., Lyttle I., Costigan I., Law J., Mitarotonda D., Larmarange J., Boiser J., Lee C.H. *Package ‘lubridate’*. Dostupno na: <https://lubridate.tidyverse.org>, <https://github.com/tidyverse/lubridate>. [Pristupljeno: 28. svibnja 2024.].
- [16] Ramsay J., Hooker G., Graves S. *Package ‘fda’*. Dostupno na: <http://www.functionaldata.org>. [Pristupljeno: 28. svibnja 2024.].
- [17] Ooms J. *Package ‘writexl’*. Dostupno na: <https://docs.ropensci.org/writexl/>, <https://github.com/ropensci/writexl>, <https://libxlsxwriter.github.io>. [Pristupljeno: 28. svibnja 2024.].
- [18] Zeileis A., Grothendieck G., Ryan J.A., Ulrich J.M., Andrews F. *Package ‘zoo’*. Dostupno na: <https://zoo.R-Forge.R-project.org/>. [Pristupljeno: 28. svibnja 2024.].

- [19] Happ-Kurz C. *Package 'funData'*. Dostupno na: <https://github.com/ClaraHapp/funData>. [Pristupljeno: 28. svibnja 2024.].
- [20] Maechler M., Rousseeuw P., Struyf A., Hubert M., Hornik K., Studer M., Roudier P., Gonzalez J., Kozlowski K., Schubert E., Murphy K. *Package 'cluster'*. Dostupno na: <https://svn.r-project.org/R-packages/trunk/cluster/>. [Pristupljeno: 28. svibnja 2024.].
- [21] Kassambara A., Mundt F. *Package 'factoextra'*. Dostupno na: <http://www.sthda.com/english/rpkgs/factoextra>. [Pristupljeno: 28. svibnja 2024.].
- [22] Therneau T., Atkinson B., Ripley B. *Package 'rpart'*. Dostupno na: <https://github.com/bethatkinson/rpart>; <https://cran.r-project.org/package=rpart>. [Pristupljeno: 28. svibnja 2024.].
- [23] Gauster T., Laaha G., Koffler D. *Package 'lfstat'*. Dostupno na: <https://cran.r-project.org/web/packages/lfstat/index.html>. [Pristupljeno: 28. svibnja 2024.].
- [24] Chow V.T. *Applied hydrology*. McGraw-Hill series in water resources and environmental engineering. New York: McGraw-Hill; 1988.
- [25] Yue S., Ouarda T.B.M.J., Bobée B., Legendre P., Bruneau P. Approach for describing statistical properties of flood hydrograph. *Journal of Hydrologic Engineering*. 2002;7(2):147-157. doi:10.1061/(ASCE)1084-0699(2002)7:2(147)
- [26] Robson A.J., Reed D.W. Statistical procedures for flood frequency estimation. In: *Flood Estimation Handbook*, vol. 3. Wallingford: Centre for Ecology & Hydrology; 1999.
- [27] Bezak N., Brilly M., Šraj M. Comparison between the peaks-over-threshold method and the annual maximum method for flood frequency analysis. *Hydrological Sciences Journal*. 2014;59(5):959-977. doi:10.1080/02626667.2013.831174
- [28] Sinaga K.P., Yang M.-S. Unsupervised K-Means clustering algorithm. *IEEE Access*. 2020;8:67655-67666. doi:10.1109/ACCESS.2020.2988796
- [29] Dehariya V.K., Shrivastava S.K., Jain R.C. Clustering of image data set using K-means and fuzzy K-means algorithms. U: Proceedings of the 2010 International Conference on Computational Intelligence and Communication Networks; Nov 26–28, 2010; Bhopal, Indija; 2010. p. 386-391. doi:10.1109/CICN.2010.80

- [30] Řezanková H. Different approaches to the silhouette coefficient calculation in cluster evaluation. U: Proceedings of the 21st International Scientific Conference AMSE Applications of Mathematics and Statistics in Economics; 29. kolovoza–2. rujna 2018; Kutná Hora, Češka; 2018.
- [31] Jain AK, Murty MN, Flynn PJ. Data Clustering: A Review. *ACM Computing Surveys*. 1999;31(3).

9 SAŽETAK

Autori rada: Lovro Štefan, Luka Đud

Naslov rada: Klasifikacija hidrograma prema obliku na slivu rijeke Kupe i Krapine pomoću algoritma strojnog učenja

Sažetak: Projektni hidrogram, koji objedinjuje podatke o vršnom protoku, trajanju i volumenu vodnog vala, ključan je ulazni podatak u projektiranju različitih hidrotehničkih objekata i sustava zaštite od poplava. U mnogim projektnim zahtjevima, osim vršne vrijednosti projektnog hidrograma, važan je i njegov oblik. Ovaj rad primjenjuje metodu klasifikacije hidrograma vodnih valova na temelju oblika za slivove rijeka Krapine i Kupe. Korištena je metoda strojnog učenja i funkcijeske analize. K-means klasteriranjem valova reprezentiranih kontinuiranim Bspline funkcijama definirana su 4 klastera za rijeku Krapinu i 6 za rijeku Kupu. Uočene su i opisane razlike parametara vodnih valova, a u manjem opsegu ispitana je i povezanost oborina s dobivenim klasterima. Klasteri mogu poslužiti kao baza oblika za izbor pri kasnijem definiranju projektnog hidrograma. Provedena metodologija može se ponoviti za slivove sličnih karakteristika ili primijeniti prilagođene dobivene oblike na neizučenim slivovima. Metoda je osjetljiva na ulazne podatke i zavisi o subjektivnom odabiru broja klastera, no bolje predstavlja raznolikost poplavnih događaja na slivu od tradicionalnih metoda koje se koriste u praksi.

Ključne riječi: poplavni valovi, projektni hidrogram, klasteriranje, K-means, funkcijeska analiza

10 SUMMARY

Authors: Lovro Štefan, Luka Đud

Title: Shape-Based Hydrograph Classification for the Kupa and Krapina Catchments Using a Machine Learning Algorithm

Summary: The design hydrograph, which integrates data on peak flow, duration, and volume of flood waves, is a crucial input for designing various hydraulic structures and flood protection systems. In many design requirements, the shape of the design hydrograph, in addition to its peak value, is also important. This paper employs a hydrograph classification method based on the shape of flood waves for the Krapina and Kupa river catchments. Machine learning methods and functional analysis were used. K-means clustering of waves, represented by continuous Bspline functions, identified 4 clusters for the Krapina River and 6 for the Kupa River. Differences in flood wave parameters were observed and described, and the relationship between precipitation and the identified clusters was also examined to a lesser extent. The clusters can serve as a basis for shape selection in the subsequent definition of the design hydrograph. The applied methodology can be replicated for catchments with similar characteristics or adapted for unstudied catchments using the derived shapes. While the method is sensitive to input data and depends on the subjective choice of the number of clusters, it provides a more nuanced representation of flood event diversity in the catchment compared to traditional methods used in practice.

Key words: flood waves, design hydrograph, clustering, K-means, functional analysis

Lovro Štefan rođen je 3. travnja 2001. godine u Varaždinu, gdje je završio prirodoslovno-matematički smjer Prve gimnazije Varaždin te srednje glazbeno obrazovanje na Glazbenoj školi u Varaždinu. Student je prve godine diplomskog studija Građevinarstvo, smjer Hidrotehnika na Građevinskom fakultetu Sveučilišta u Zagrebu na kojem je prijediplomski studij završio s najvećom pohvalom (*summa cum laude*). Plaketom je nagrađen za uspjeh na prvoj i drugoj godini studija.

Luka Đud rođen je 29. kolovoza 2001. godine u Varaždinu, gdje je i završio prirodoslovno-matematički smjer Prve gimnazije Varaždin. Student je prve godine diplomskog studija Fakulteta elektrotehnike i računarstva u Zagrebu, smjer Računarska znanost, uz završeni prijediplomski studij smjera Računarstvo.

PRILOG

Tablica P1: Vodni valovi na postaji Kupljenovo, rijeka Krapina

Redni broj	Hidrološka godina	Datum vršnog protoka	Izmjereni protok [m ³ /s]	Direktno otjecanje [m ³ /s]	Trajanje rasta [dani]	Trajanje pada [dani]	Ukupno trajanje [dani]	Volumen vodnog vala [10 ⁶ m ³]
1	1964	25.10.1964.	143	135.31	6	14	21	48.40
2	1965	9.12.1965.	153	135.32	3	1	5	31.88
3	1966	26.11.1966.	169	153.24	10	5	16	76.07
4	1967	16.6.1968.	49.6	47.71	2	10	13	9.87
5	1968	19.12.1968.	141	137.63	2	15	18	27.63
6	1969	6.1.1970.	89.8	83.10	4	4	9	19.67
7	1970	25.1.1971.	92.6	87.08	1	15	17	31.01
8	1971	14.7.1972.	206	203.81	4	4	9	60.94
9	1972	25.11.1972.	135	129.48	2	12	15	33.04
10	1973	6.5.1974.	112	110.72	2	25	28	41.29
11	1974	22.10.1974.	193	177.03	2	6	9	36.38
12	1975	29.4.1976.	208	204.11	7	14	22	59.89
13	1976	11.12.1976.	119	108.18	2	6	9	25.83
14	1977	14.11.1977.	93.3	88.16	2	7	10	28.48
15	1978	30.1.1979.	109	101.16	3	10	14	44.62
16	1979	20.11.1979.	120	116.37	6	15	22	49.55
17	1980	13.10.1980.	173	169.09	1	19	21	57.12
18	1981	13.12.1981.	96.8	86.45	2	5	8	18.04
19	1982	25.12.1982.	152	132.88	4	8	13	52.33
20	1983	27.2.1984.	99.7	93.11	3	3	7	29.64
21	1984	8.5.1985.	141	128.77	2	11	14	29.58
22	1985	17.6.1986.	131	123.91	3	6	10	43.69
23	1986	21.2.1987.	127	123.00	2	21	24	44.45
24	1987	16.11.1987.	124	117.26	3	8	12	36.66
25	1988	4.7.1989.	252	247.70	2	7	10	56.51
26	1989	29.9.1989.	121	117.87	2	7	10	24.76
27	1990	2.11.1990.	94.9	91.59	2	16	19	31.33
28	1991	22.11.1991.	190	186.50	3	22	26	85.87
29	1992	31.10.1992.	110	107.54	3	3	7	27.90

Tablica P1 (nastavak): Vodni valovi na postaji Kupljenovo, rijeka Krapina

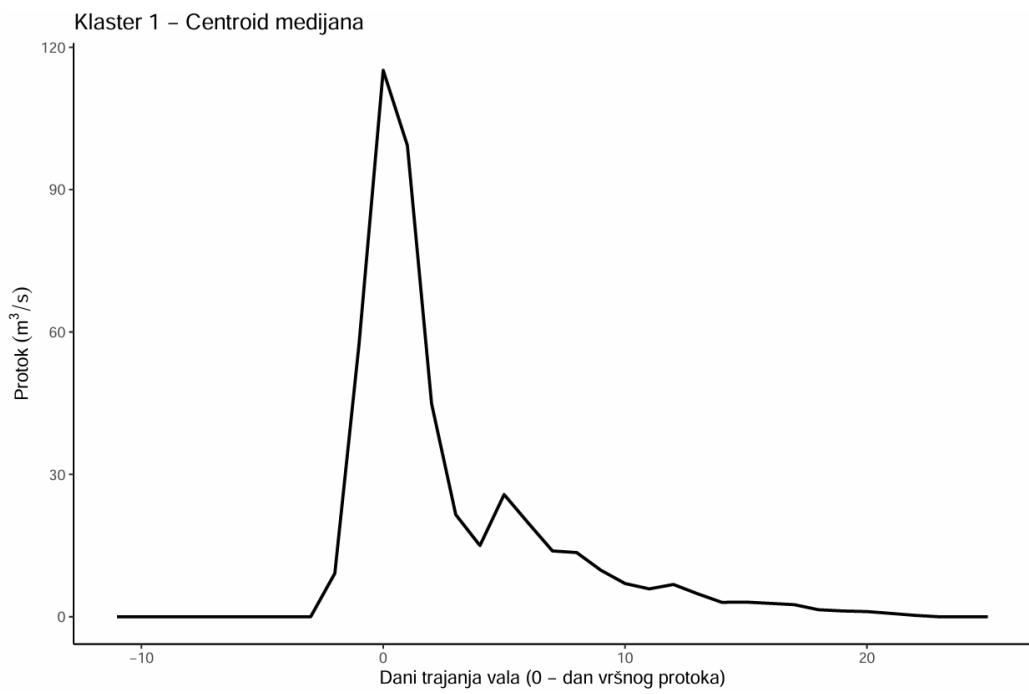
Redni broj	Hidrološka godina	Datum vršnog protoka	Izmjereni protok [m ³ /s]	Direktno otjecanje [m ³ /s]	Trajanje rasta [dani]	Trajanje pada [dani]	Ukupno trajanje [dani]	Volumen vodnog vala [10 ⁶ m ³]
30	1993	12.12.1993.	126	108.11	4	2	7	26.67
31	1994	1.1.1995.	127	124.49	11	17	29	50.96
32	1995	16.9.1995.	146	142.43	3	22	26	51.60
33	1996	23.11.1996.	74	70.31	7	2	10	22.05
34	1997	4.12.1997.	148	141.26	5	7	13	33.59
35	1998	15.9.1998.	197	192.99	4	11	16	69.73
36	1999	5.10.1999.	113	110.15	3	13	17	21.23
37	2000	12.4.2001.	105	97.36	1	22	24	29.42
38	2001	28.4.2002.	97	93.76	2	19	22	20.86
39	2002	6.12.2002.	81.4	75.64	5	10	16	22.79
40	2003	24.3.2004.	124	117.44	3	11	15	31.97
41	2004	18.10.2004.	105	100.59	9	3	13	25.70
42	2005	31.5.2006.	163	159.85	8	2	11	36.92
43	2006	21.3.2007.	65.7	59.70	3	2	6	14.75
44	2007	25.3.2008.	106	102.24	4	17	22	40.41
45	2008	22.1.2009.	107	96.86	2	4	7	20.78
46	2009	21.2.2010.	107	100.70	4	5	10	29.15
47	2010	19.9.2010.	175	172.90	3	5	9	45.11
48	2011	17.12.2011.	33	31.51	5	8	14	7.47
49	2012	31.3.2013.	133	111.96	4	12	17	55.01
50	2013	12.11.2013.	168	160.68	4	7	12	50.16
51	2014	23.5.2015.	161	159.15	3	16	20	30.04
52	2015	16.10.2015.	185	182.48	10	2	13	56.70
53	2016	12.11.2016.	122	118.35	2	25	28	30.75
54	2017	20.9.2017.	142	138.90	2	10	13	30.27
55	2018	30.5.2019.	88.9	80.51	3	6	10	12.51
56	2019	23.12.2019.	159	149.62	3	10	14	29.28
57	2020	20.5.2021.	127	115.97	4	5	10	29.63
58	2021	3.12.2021.	67.3	63.28	2	17	20	16.57

Tablica P2: Vodni valovi na postaji Farkašić, rijeka Kupa

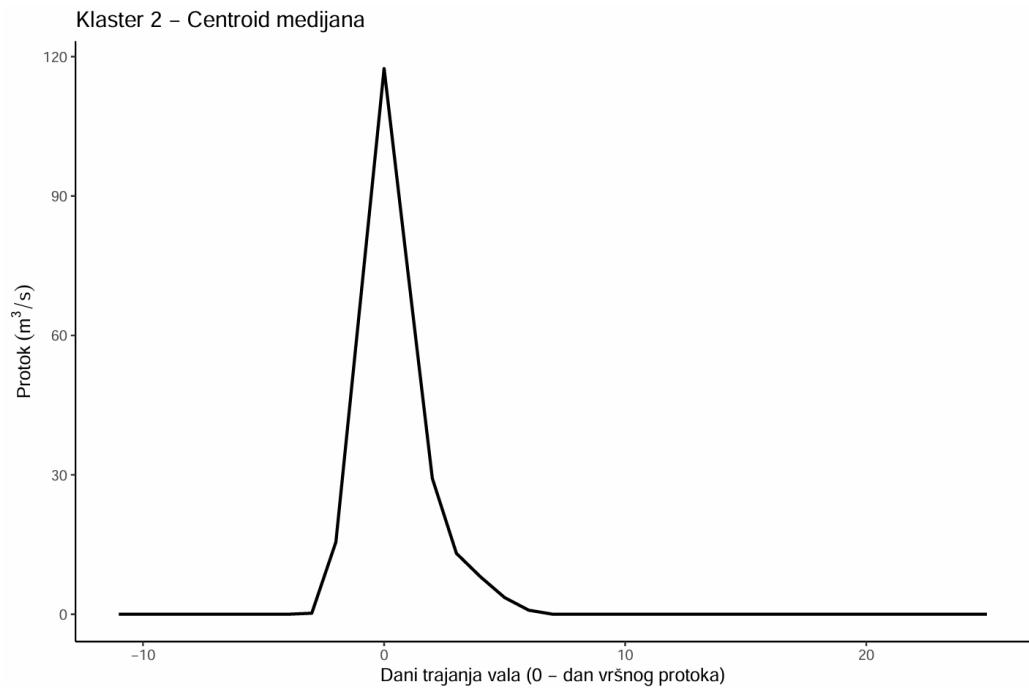
Redni broj	Hidrološka godina	Datum vršnog protoka	Izmjereni protok [m ³ /s]	Direktno otjecanje [m ³ /s]	Trajanje rasta [dani]	Trajanje pada [dani]	Ukupno trajanje [dani]	Volumen vodnog vala [10 ⁶ m ³]
1	1965	13.12.1965.	1166	835.14	6	8	15	403.36
2	1966	6.12.1966.	1337	998.83	4	5	10	484.59
3	1967	28.12.1967.	1054	922.37	5	8	14	493.12
4	1968	21.12.1968.	975	909.85	5	18	24	425.42
5	1969	8.1.1970.	1220	1071.91	3	5	9	448.75
6	1970	24.1.1971.	918	818.08	4	8	13	622.12
7	1971	20.4.1972.	1245	1087.66	10	15	26	1099.50
8	1972	26.11.1972.	1119	1044.64	15	5	21	882.11
9	1973	3.5.1974.	765	688.97	3	19	23	479.30
10	1974	8.10.1974.	1585	1273.96	8	12	21	1300.95
11	1975	19.10.1975.	819	770.57	2	21	24	361.30
12	1976	12.4.1977.	986	897.02	14	13	28	876.22
13	1977	18.11.1977.	982	882.76	5	4	10	380.92
14	1978	1.2.1979.	993	790.75	10	9	20	514.87
15	1979	21.11.1979.	1272	1167.94	6	20	27	931.48
16	1980	15.3.1981.	1174	1052.15	13	11	25	1049.92
17	1981	3.1.1982.	1119	867.24	5	9	15	564.72
18	1982	31.3.1983.	1001	830.05	11	22	34	781.62
19	1983	15.4.1984.	888	521.10	3	9	13	193.73
20	1984	26.1.1985.	999	902.51	4	14	19	529.44
21	1985	8.6.1986.	624	518.46	10	6	17	244.88
22	1986	9.5.1987.	978	872.42	5	5	11	308.93
23	1987	16.11.1987.	753	592.46	5	9	15	355.76
24	1988	20.5.1989.	1136	1040.75	6	14	21	604.23
25	1989	1.10.1989.	584	511.68	3	6	10	174.10

Tablica P2 (nastavak): Vodni valovi na postaji Farkašić, rijeka Kupa

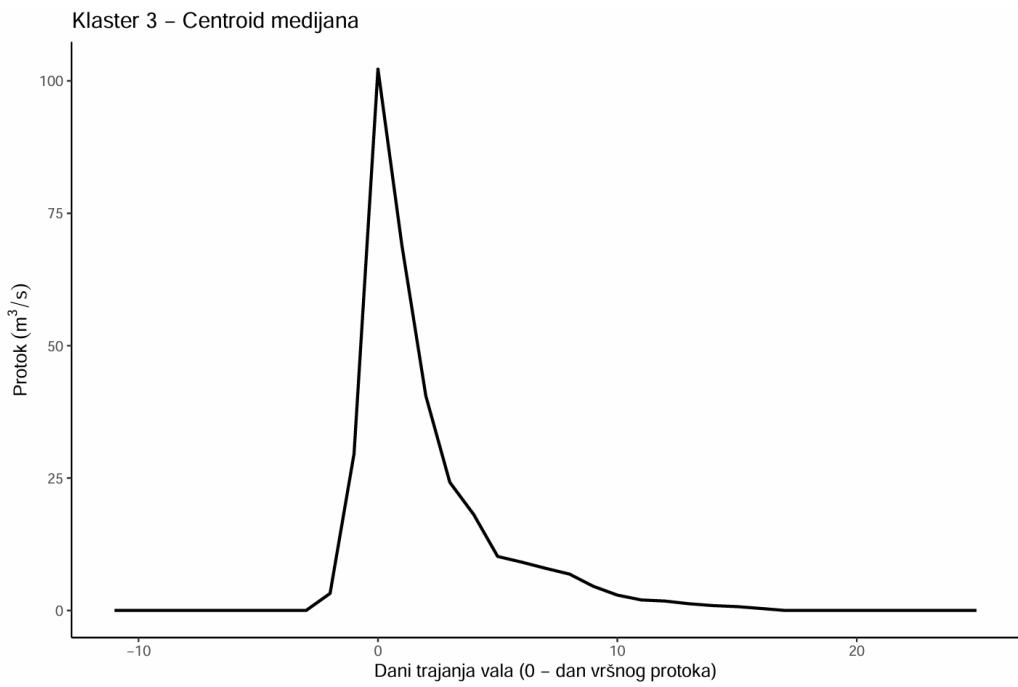
Redni broj	Hidrološka godina	Datum vršnog protoka	Izmjereni protok [m ³ /s]	Direktno otjecanje [m ³ /s]	Trajanje rasta [dani]	Trajanje pada [dani]	Ukupno trajanje [dani]	Volumen vodnog vala [10 ⁶ m ³]
26	2000	8.3.2001.	1224	995.29	7	6	14	471.55
27	2001	16.4.2002.	1031	796.24	7	2	10	272.66
28	2002	9.12.2002.	1096	897.11	6	7	14	381.13
29	2003	16.4.2004.	1144	872.57	12	9	22	794.36
30	2004	15.4.2005.	1234	1054.52	8	8	17	598.65
31	2005	4.12.2005.	1315	1250.53	20	2	23	792.76
32	2006	15.2.2007.	630	516.74	3	11	15	210.52
33	2007	25.10.2007.	766	679.36	6	15	22	475.60
34	2008	5.4.2009.	932	846.24	9	14	24	556.13
35	2009	23.2.2010.	982	821.00	5	18	24	760.87
36	2010	21.9.2010.	889	813.17	3	4	8	315.63
37	2011	19.12.2011.	530	449.85	6	7	14	192.55
38	2012	25.1.2013.	1174	951.31	5	6	12	438.26
39	2013	14.2.2014.	1239	1085.00	13	9	23	1150.07
40	2014	26.10.2014.	1209	993.67	5	10	16	597.72
41	2015	19.10.2015.	1192	1142.55	8	26	35	1112.95
42	2016	8.2.2017.	882	770.45	8	8	17	472.45
43	2017	19.3.2018.	1230	961.26	14	9	24	1052.09
44	2018	17.5.2019.	1002	716.00	4	10	15	369.23
45	2019	19.11.2019.	872	748.78	6	13	20	580.64
46	2020	1.1.2021.	834	723.15	7	18	26	759.22
47	2021	3.4.2022.	638	589.12	5	17	23	306.06



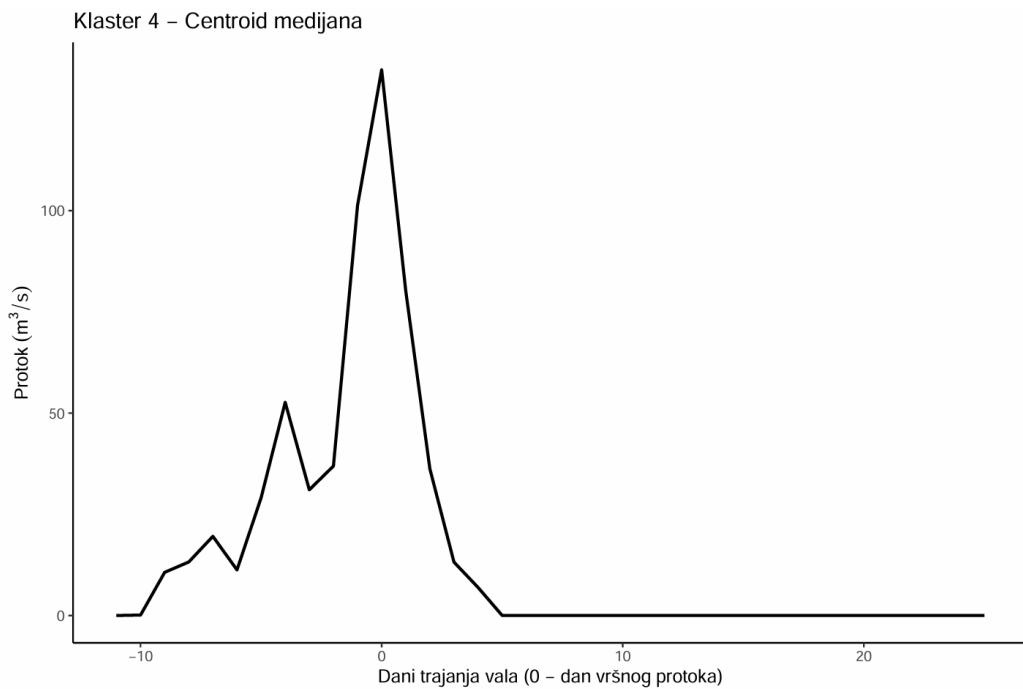
Slika P1: Hidrogram centroida vodnog vala za 1. klaster, rijeka Krapina



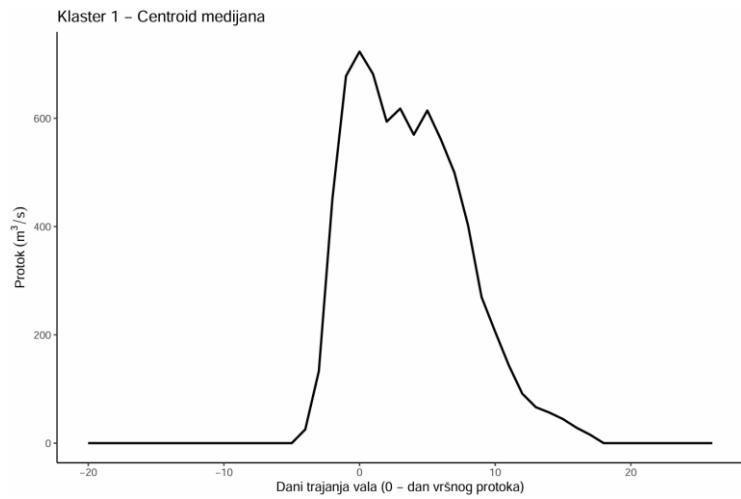
Slika P2: Hidrogram centroida vodnog vala za 2. klaster, rijeka Krapina



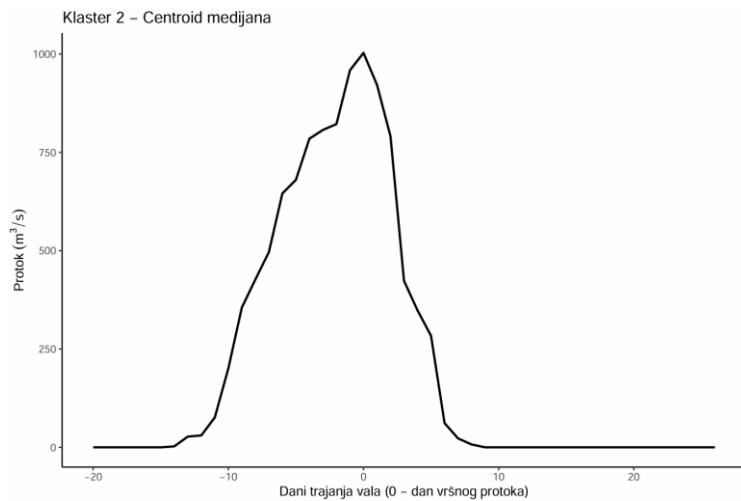
Slika P3: Hidrogram centroida vodnog vala za 3. klaster, rijeka Krapina



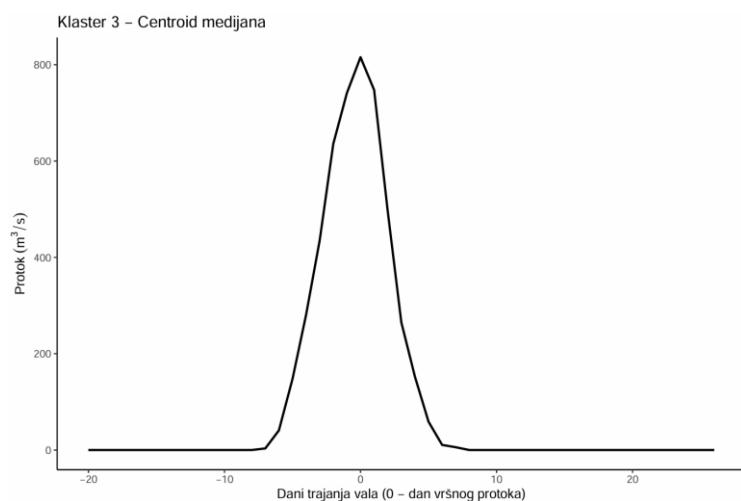
Slika P4: Hidrogram centroida vodnog vala za 4. klaster, rijeka Krapina



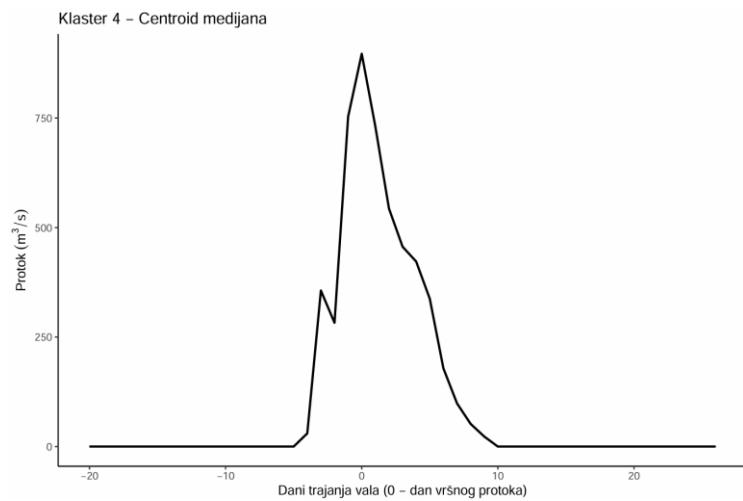
Slika P5: Hidrogram centroida vodnog vala za 1. klaster, rijeka Kupa



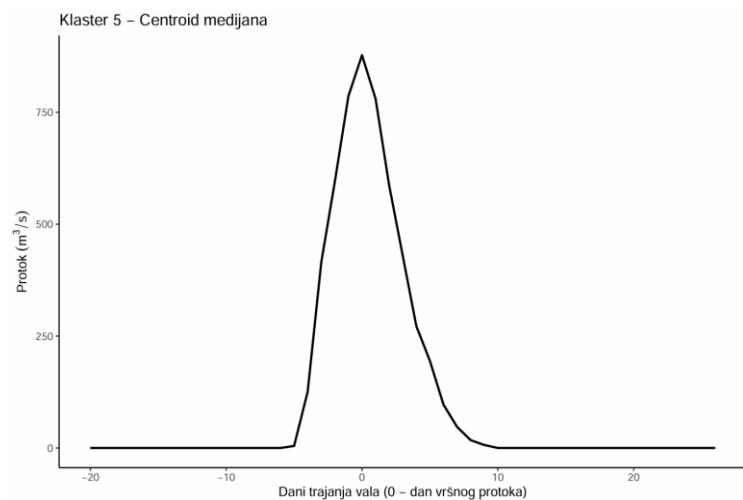
Slika P6: Hidrogram centroida vodnog vala za 2. klaster, rijeka Kupa



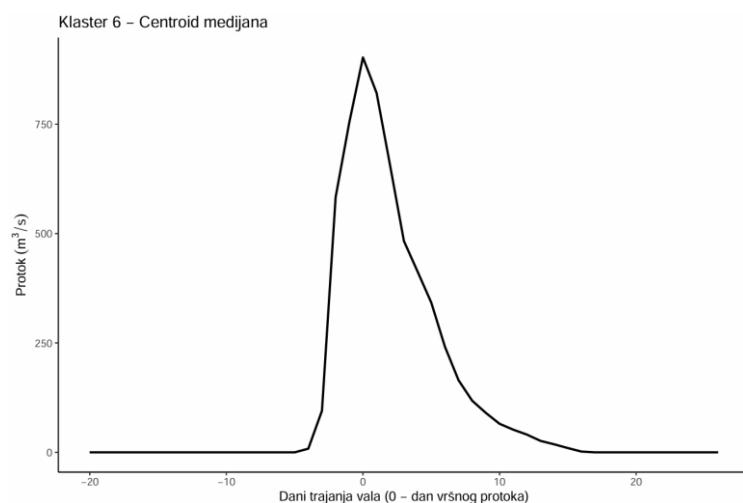
Slika P7: Hidrogram centroida vodnog vala za 3. klaster, rijeka Kupa



Slika P8: Hidrogram centroida vodnog vala za 4. klaster, rijeka Kupa



Slika P9: Hidrogram centroida vodnog vala za 5. klaster, rijeka Kupa



Slika P10: Hidrogram centroida vodnog vala za 6. klaster, rijeka Kupa