

Modeliranje sintetičkih nizova podataka mjesečnih količina oborina

Igor Ružić, Josip Rubinić, Nevenka Ožanić

Ključne riječi

sintetički nizovi podataka, mjesečne količine oborina, empirijsko-stohastički model, vremenske serije, registrirani podaci, generirani podaci

Key words

synthetic data series, monthly precipitation rate, empirical-stochastic model, time series, registered data, generated data

Mots clés

série des données synthétiques, précipitations mensuelles, modèle empirique-stochastique, séries temporelles, données enregistrées, données générées

Ключевые слова

синтетические ряды данных, месячные количества осадков, эмпирическо-стохастическая модель, временные ряды, зарегистрированные данные, генерированные данные

Schlüsselworte

synthetische Datenfolgen, monatliche Niederschlagsmengen, empirisch-stochastisches Modell, Zeitserien, registrierte Daten, generierte Daten

I. Ružić, J. Rubinić, N. Ožanić

Pregledni rad

Modeliranje sintetičkih nizova podataka mjesečnih količina oborina

Razrađen je novi pristup modeliranju (empirijsko-stohastički) sintetičkih nizova podataka o mjesečnim količinama oborina. Za analizu i verifikaciju modela rabljeni su podaci s više različitih oborinskih postaja. Dobiveni rezultati uspoređeni su s originalnim, promotrenim vremenskim serijama uz pomoć niza parametarskih testova osjetljivih na pojedine značajke distribucije registriranih i generiranih podataka. Rezultati su pokazali prihvatljivost predloženog modelskog pristupa.

I. Ružić, J. Rubinić, N. Ožanić

Subject review

Modelling synthetic data series for monthly precipitation

A novel (empirical-stochastic) approach to the modelling of synthetic series of monthly precipitation data has been developed. Data collected from several precipitation stations were used in the analysis and verification of the new model. The results were compared with the originally measured time series by means of a number of parametric tests sensitive to individual distribution-related properties of the registered and generated data. The results have shown that the proposed model may be considered acceptable.

I. Ružić, J. Rubinić, N. Ožanić

Ouvrage de synthèse

Modélage des séries des données synthétiques relatives aux précipitations mensuelles

Une nouvelle approche empirique et stochastique du modélage des séries synthétique des données sur les précipitations mensuelles a été développée. Les données recueillies sur plusieurs stations de mesure des précipitations ont été utilisées dans l'analyse et la vérification du nouveau modèle. Les résultats ont été comparés avec les séries de temps originelles à l'aide d'un nombre des essais paramétriques susceptibles aux propriétés individuelles de distribution des données enregistrées et générées. Les résultats ont montré que le modèle proposé peut être considéré acceptable.

И. Рујић, Ђ. Рубинић, Н. Ожанић

Обзорная работа

Моделирование синтетических рядов данных месячных количеств осадков

В работе разрабатывается новый подход к моделированию (эмпирическо-стохастический) синтетических рядов данных о месячных количествах осадков. Для анализа и верификации модели использовались данные с большего числа разных метеорологических станций. Полученные результаты сравнивались с оригинальными, наблюдаемых во временных сериях при помощи ряда параметрических тестов, чувствительных на отдельные характерные свойства, дистрибуции зарегистрированных и генерированных данных. Результаты показали приемлемость предложенного модельного подхода.

I. Ružić, J. Rubinić, N. Ožanić

Übersichtsarbeit

Modellieren synthetischer Datenfolgen monatlicher Niederschlagsmengen

Erarbeitet ist ein neuer Zutritt zum Modellieren (empirisch-stochastisch) synthetischer Datenfolgen über monatliche Niederschlagsmengen. Zur Analyse und Verifikation gebrauchte man Daten von mehreren verschiedenen Niederschlagsstationen. Die erhaltenen Ergebnisse verglich man mit originalen, beobachteten Zeitserien mit Hilfe einer Folge von Parametertesten die gegen einzelne Merkmale der Distribution registrierter und generierter Daten empfindlich sind. Die Ergebnisse zeigten die Annehmbarkeit des vorgeschlagenen Modellzutritts.

Autori: **Igor Ružić**, dipl. ing. građ.; mr. sc. **Josip Rubinić**, dipl. ing. građ.; prof. dr. sc. **Nevenka Ožanić**, dipl. ing. građ., Sveučilište u Rijeci, Građevinski fakultet, Rijeka

1 Uvod

U radu je predstavljen jedan od mogućih modela formiranja sintetičkih nizova meteoroloških podataka. U navedenom se slučaju radi o generiranju mjesečnih količina oborina koje su uz protoke najčešće primijenjen parameter u hidrološkim analizama. Pritom je iskorišten stohastički pristup modeliranju koji se često rabi za simulaciju vremenskih serija u hidrologiji. Većina stohastičkih modela razvijenih za simulaciju vremenskih serija su autoregresivni (AR) i autoregresivni pomičnih prosjeka (ARMA) [9]. Determinističko modeliranje oborinskih procesa veoma je složeno i upitne pouzdanosti zbog složene prirode procesa koje je vrlo teško parametarski opisati. Vremenski i prostorni raspored oborina ovisi o velikom broju čimbenika (geografski položaj, udaljenost od mora, konfiguracija tla, utjecaj šuma, velikih gradova...) i njihovih međudjelovanja. Zbog toga se pri modeliranju serija oborinskih podataka u hidrologiji gotovo isključivo primjenjuju stohastički modeli. Pri tome je česta primjena i sintetički generiranih nizova – uglavnom pri simulacijama oborinskih procesa u nekim sustavima ili pri nadopunjavanju nedostajućih podataka nekoga vremenskog niza.

Oborine su osnovni pokretač poplava, suša, klizišta, puzanja krša, puzanja tla ... Njihove analize i modeliranje karakteristični su problemi primijenjene hidrometeorologije. Oborine karakterizira velika vremenska i prostorna varijabilnost, stoga stohastičko modeliranje oborina nije jednostavan zadatak [5].

Cox i Isham [4] definiraju tri osnovna tipa modeliranja oborinskih podataka: empirijsko statistički modeli, modeli dinamičke meteorologije i polustohastički modeli. Podjela se zasniva na udjelu stvarnih podataka pri modeliranju. Model prikazan u ovome radu je empirijsko-stohastički, a modeliranje sintetičkih podataka zasniva se na stvarnim podacima.

Promjene količina i rasporeda oborina kao posljedica klimatskih promjena mogu se simulirati sintetički modeliranim oborinama, pod uvjetom da se uzme u obzir model trenda izmjerenih oborina. Hidrološki sustavi simuliraju se modelima otjecanja, koji često sadrže ili preuzimaju podatke stohastičkih oborinskih modela. Tako se može simulirati ponašanje vodoopskrbnih, hidroenergetskih, melioracijskih i sličnih sustava u različitim hidrološkim prilikama. Pri tome se rabi različita razina vremenske diskretizacije – godišnja (vrlo rijetko – npr. za analize vjerojatnosti koincidencije uzastopne pojave više sušnih godina), mjesečna (npr. za dimenzioniranje akumulacija) ili pak dnevna (npr. pri zadacima vezanim uz obranu od poplave).

U svijetu je vrlo široka primjena takvih modela u hidrologiji. Nasuprot tome, u Hrvatskoj su i dalje vrlo usam-

ljeni slučajevi primjene generiranih sintetičkih vremenskih serija u hidrologiji, iako prvi radovi iz toga područja potječu još iz razdoblja potkraj šezdesetih [2]. Rubinić i Margeta [7] primjenjuju generirane sintetičke vremenske serije o dotocima pri analizi akumulacije Boljunčica u Istri, Brezak i Spajić [3] za analizu malih voda rijeke Drave, a Margeta i Fistanić [6] za generiranje pojava mutnoća u vodama krškoga izvora Jadro.

U ovome radu razvijen je stohastički model generiranja sintetičkih nizova mjesečnih oborina. Modeliranje oborina zasniva se na nizu izmjerenih podataka koji su ujedno i ulazni parametri modela. Upotrijebljen je generatorom slučajnih brojeva normalne distribucije. U modelu su moguće i simulacije budućih događaja budući da uzima u obzir trendove izmjerenih serija podataka.

Model je razvijen u edukativne svrhe provođenja vježbi iz predmeta Hidrologija na dodiplomskom studiju građevinskog fakulteta. Model omogućava zadavanje različitih sintetičkih nizova oborinskih podataka realnih značajki velikom broju studenata. No, model ima i mnogo širu praktičnu primjenu u hidrologiji realnih sustava – pri analizama različitih scenarija pojava različitih nizova kišnih događaja.

2 Osnovne metodološke postavke provedenih analiza

2.1 Osnovni prikaz modela

Shema razvijenoga modela prikazana je na slici 1. Sintetički podaci dobiveni su s pomoću generatora slučajnih brojeva normalne distribucije programskog paketa MATLAB 6.5.

Ulazni su parametri za model srednja vrijednost (SV_{mj}) i standardno odstupanje (STD_{mj}) izmjerenih podataka određenog mjeseca (korak b u shemi modela na slici 1.). Ulazne srednje vrijednosti mijenjaju se u ovisnosti o trendu izmjerenih podataka (jednadžba 1 – korak c).

$$P_{mj_ulaz}(GOD) = TREND(SV_{mj}, GOD) \quad (1)$$

P_{mj_ulaz} - trendizirana srednja vrijednost oborina nekog mjeseca – ulaz u model $TREND(SV_{mj}, GOD)$ – jednadžba trenda određenog mjeseca

SV_{mj} - srednja oborina određenog mjeseca

GOD - godina.

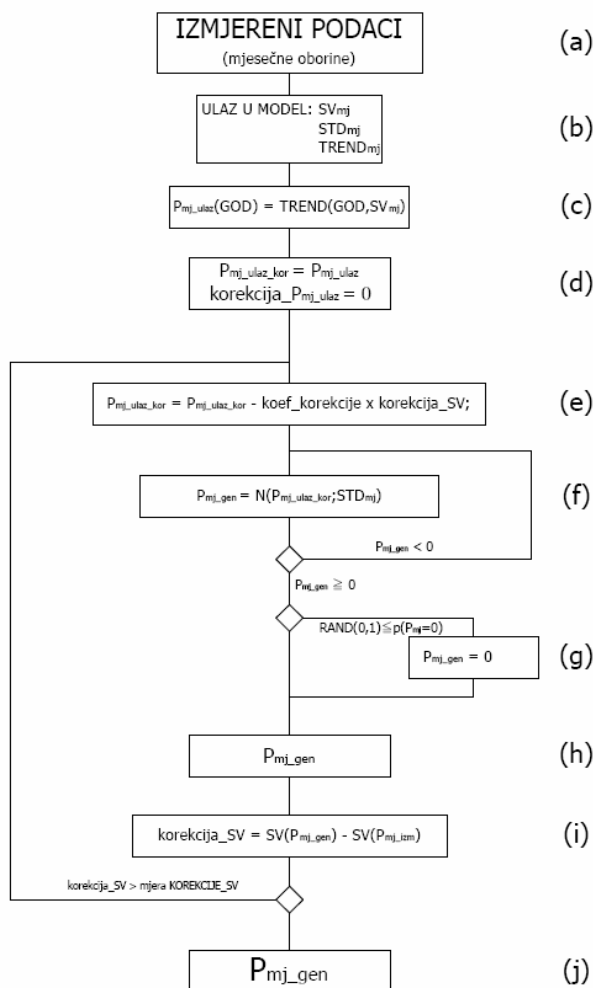
Srednje vrijednosti u ovisnosti o trendu i standardno odstupanje nekog mjeseca ulaze u model, odnosno generator slučajnih brojeva normalne distribucije (jednadžba 2):

$$P_{mj_gen}(GOD) = N(P_{mj_ulaz}(GOD), STD_{mj}) \quad (2)$$

P_{mj_gen} - generirana mjesečna oborina

$N(P_{mj_ulaz}(GOD), STD_{mj})$ - generator slučajnih brojeva normalne distribucije

STD_{mj} - standardno odstupanje određenog mjeseca.



Slika 1. Shema razvijenog modela sintetičkoga generiranja mjesečnih količina oborina

Ova distribucija daje i negativne vrijednosti. Problem „negativnih“ oborina riješen je upotrebom *while* petlje (jednadžba 3 – korak f). Generirane se vrijednosti manje od „0“ odbacuju, normalna se distribucija izvršava unutar petlje dokle god postoje „negativne“ oborine.

While ($P_{mj_gen} < 0$) (3)

$P_{mj_gen}(GOD) = N(P_{sij_ulaz}(GOD), STD_{mj})$

Problem pojave mjeseca bez oborina u modelu je dosta složen budući da se radi o veoma rijetkim događajima. Opisanom se distribucijom uglavnom ne dobije realno stanje. Taj je problem riješen uvođenjem dodatnog uvjeta da u nekom mjesecu nisu pale oborine, uporabom generatora slučajnih brojeva uniformne distribucije i vjerojatnosti pojave mjeseca bez oborina (korak g).

Tako generirane oborine daju relativno dobre rezultate statističkih testova, no zbog upotrebe spomenute *while* petlje, odnosno opisane distribucije dolazi do povećanja srednjih vrijednosti sintetičkih podataka.

Stoga model sam provodi popravak ulaznih podataka na osnovi razlike izmjerenih i generiranih podataka (korak i). Popravak se provodi dok je razlika izmjerenih i generiranih srednjih vrijednosti veća od zadane u modelu.

2.2 Statistička obrada izmjerenih i generiranih podataka

Statistička obrada izmjerenih i generiranih podataka napravljena je uporabom u nastavku navedenih statističkih parametara [8].

Centar, aritmetička sredina :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

gdje je n broj članova niza, a x_i je i-ti član niza.

Standardno odstupanje:

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$

Trend je usmjeravanje (padajuće ili rastuće) u vremenskim nizovima. Jednadžba linearnog trenda računa se metodom najmanjih kvadrata odstupanja:

$$x - \bar{x} = \frac{\sum (\Delta x_i \cdot \Delta y_i)}{\sum (\Delta y_i)^2} (y - \bar{y})$$

x_i, y_i - koordinate i-te točke

\bar{x}, \bar{y} - srednja vrijednost koordinata.

Učestalost neke vrijednosti slučajne varijable je broj koji pokazuje koliko puta se ta vrijednost varijable pojavila unutar nekoga skupa. Ako se broj pojavljivanja neke učestalosti podijeli s ukupnim brojem pojava dobije se relativna učestalost.

2.3 Distribucije slučajnih brojeva

U ovome radu upotrijebljeni su generatori slučajnih brojeva uporabom računalnog programa MATLAB.

Uniformna distribucija vjerojatnosti diskretne slučajne varijable jest

$$p(x) = \frac{1}{n} \quad x = 1, 2, 3, \dots, n \quad n \in \mathbb{N}$$

Za modeliranje mjesečnih oborina uporabljena je Gaussova krivulja raspodjele s određenim modifikacijama zbog specifičnosti učestalosti mjesečnih količina oborina.

Gaussova (normalna) krivulja raspodjele

Gaussova (normalna) krivulja raspodjele je simetrična, dvoparametarska raspodjela neprekinute slučajne varijable zvonolikog oblika. Većina pojava u hidrologiji nije simetrična, stoga se Gaussova raspodjela najčešće rabi za usporedbu drugih raspodjela. Po Gaussovu zakonu raspodjela vjerojatnosti jest:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}$$

\bar{x} - prosjek članova niza

σ - standardno odstupanje.

Eksponecijalna distribucija

Eksponecijalna krivulja raspodjele je nesimetrična raspodjela neprekinute slučajne varijable

$$p(x) = \int_0^x \frac{1}{\mu} e^{-\frac{t}{\mu}} dt = 1 - e^{-\frac{x}{\mu}}$$

μ - prosjek članova niza.

2.4 Provjeravanje generiranih podataka

Provjera pripadnosti istomu skupu sintetički generiranih i izmjerenih podataka provedeno je sljedećim parametarskim testovima sadržanim unutar programskog paketa MATLAB:

Ttest2 provjerava razlike srednjih vrijednosti dvaju neovisnih uzoraka, odnosno njihovu pripadnost jednakoj distribuciji.

Wilcoxon rank sum test - provjerava hipotezu da dva neovisna uzorka potječu iz iste distribucije jednakih medijana.

Ansari-Bradley test jednakih disperzija provjerava hipotezu da dva neovisna uzorka potječu iz iste distribucije jednakih medijana i oblika no različitih disperzija (odstupanja).

Kstest2 je *Kolmogorov-Smirnov test* pripadnosti istoj distribuciji dvaju skupova podataka X_1 i X_2 duljina n_1 i n_2 .

Ranksum test testira hipotezu pripadnosti dvaju neovisnih uzoraka distribucijama jednakih medijana.

Kod svih je testova upotrijebljena standardna razina značajnosti od 5 % odnosno interval pouzdanosti 95 %.

3 Razvoj modela

Kao što je već spomenuto u uvodu, ovaj je model empirijsko-stohastički. Osnovni ulazni parametri modela dobiveni su obradom izmjerenih podataka.

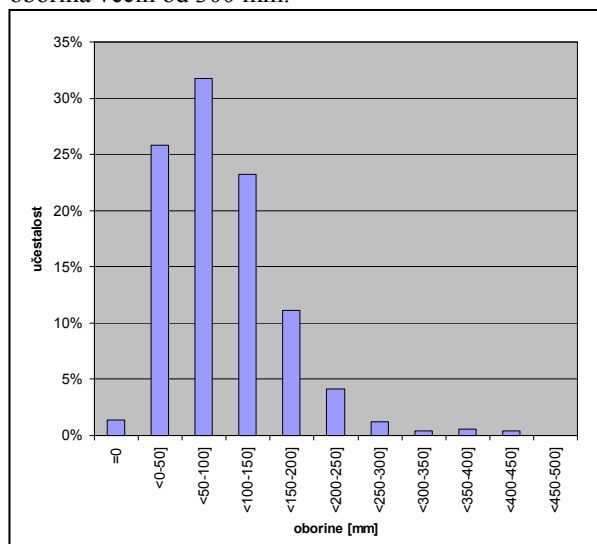
3.1 Oborinska osnova kao ulaz u stohastički model

Pri razvoju modela uporabljeni su i modelski analizirani mjesečni podaci o oborinama s više meteoroloških i kišomjernih postaja na području sjeveroistočne obale Istre: Čepić, Labin, Rakonek, Mošćenička Draga, Letaj (brana) i Barban. Obrađeno je razdoblje od 1961. do 2002. godine,

Tablica 1. Osnovni statistički parametri mjesečnih oborina, Čepić (1961.-2002.)

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	GOD
S.V.	95	81	84	86	73	89	67	88	114	123	149	100	1148
MAX	356	186	206	179	195	213	179	294	427	393	432	261	1657
MIN	0	4	0	19	5	10	0	0	4	0	11	22	801
STD	75	57	56	40	43	43	41	62	84	95	86	56	199
Trend	-1,13	-1,19	-0,46	-0,15	-0,45	0,13	-0,48	-0,64	-0,24	0,58	0,45	-0,14	-3,73

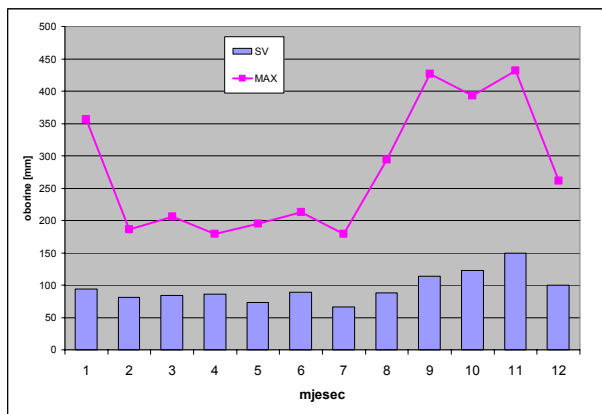
što je za domaće prilike relativno dugačak niz. Primarna obrada oborinskih podataka i nadopunjavanje eventualnih nedostajajućih pojedinih vrijednosti provedena je u studiji Građevinskog fakulteta u Rijeci (2005.). Za ilustraciju primijenjenoga metodološkog postupka odabrani su podaci s meteorološke postaje Čepić. Prikazani su rezultati provedene statističke obrade promatranih vrijednosti, kao i njihova primjena u različitim fazama razvoja modela. Tablica 1. prikazuje osnovne statističke parametre promatranih mjesečnih oborina meteorološke stanice Čepić. Na slici 2. prikazan je histogram svih mjesečnih oborina promatranoga razdoblja postaje Čepić. Iz tog histograma vidimo da su najčešće mjesečne oborine između 50 i 100 mm. Rijetka je pojava mjesečnih količina oborina većih od 300 mm.



Slika 2. Histogram mjesečnih oborina registriranih na postaji Čepić (1961.-2002.)

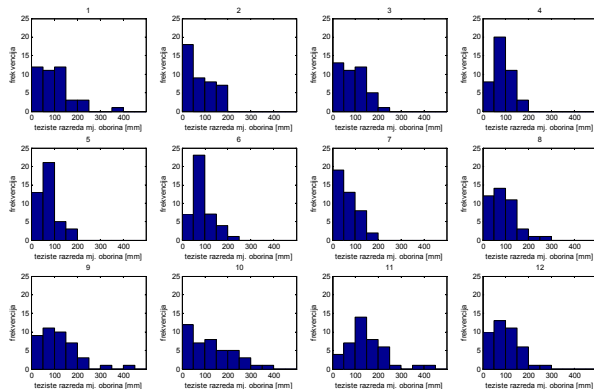
Kada se analizira svaki mjesec posebno, mjesečne količine oborina znatno variraju (slike 3. i 4.). Slika 3. prikazuje godišnji hod srednjih i maksimalnih oborina. Tablica 1. prikazuje osnovne statističke parametre svakog

mjeseca posebno. Iz navedene slike 3. i tablice 1. vidi se da ni jedan mjesec nema približno jednake statističke parametre. Najsušni mjesec je srpanj s prosječno 67 mm oborina, a u tom je mjesecu zabilježen i najmanji mjesečni maksimum od 179 mm oborina. Najkišoviti mjesec je studeni s prosječno 149 mm oborina, a drugi po redu najkišoviti mjesec je listopad sa 123 mm oborina. Nabrojene karakteristike odgovaraju maritimnom tipu srednjega godišnjega hoda oborina [2].



Slika 3. Godišnji hod srednjih i maksimalnih mjesečnih oborina registriranih na postaji Čepić (1961-2002)

Slika 4. prikazuje histograme oborina svakog mjeseca. Sam izgled histograma pokazuje različitost distribucija oborina tijekom svakog mjeseca iz čega proizlazi potreba generiranja sintetičkih oborina svakog mjeseca posebno.



Slika 4. Histogrami mjesečnih oborina registriranih na postaji Čepić (1961.-2002.)

3.2 Eksponecijalna raspodjela

Početni pristup generiranju sintetičkog niza bio je uporabom generatora slučajnih brojeva ekspancijalne distribucije. Ekspancijalna distribucija nije simetrična, pa stoga njezina krivulja razdiobe nalikuje razdiobi mjesečnih oborina. Pri generiranju ekspancijalnom distribucijom, kao ulazni parametar upotrijebljena je izmjeren srednja mjesečna oborina. Svaki je mjesec generiran posebno (4). Generirana je serija vremenskih nizova mje-

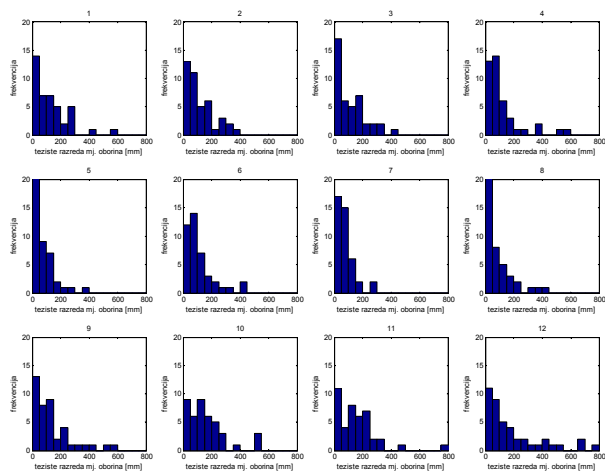
sečnih oborina, pri čemu je usvojena identična duljina kao i kod ulazne – registrirane serije (42 godine). Godišnja količina oborina rezultirajućih sintetičkih vremenskih serija dobivena je kao zbroj pojedinih mjesečnih generiranih vrijednosti.

$$P_{mj_gen}(GOD) = EXP(SV(P_{mj})) \quad (4)$$

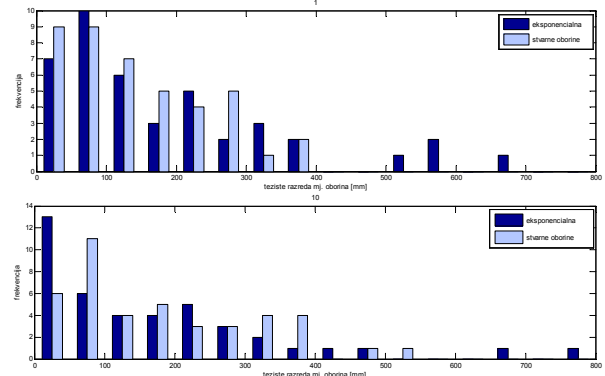
$EXP(SV(P_{mj}))$ - generator slučajnih brojeva ekspancijalne distribucije.

Rezultati u obliku mjesečnih histograma prikazani su na slici 5. Usporedbom s histogramom na slici 4. vidi se da distribucije nisu znatno različite za male vrijednosti, no bitna su odstupanja kod većih vrijednosti. To se vidi na slici 6. koja, primjera radi, prikazuje usporedni histogram učestalosti izmjerenih i generiranih podataka za siječanj i listopad.

Zbog već očitog znatnog odstupanja većih vrijednosti, odbačeno je generiranje sintetičkih oborina ekspancijalnom distribucijom. Iz tih razloga nisu provedene ni statističke analize vezane uz ocjenu dobivenih rezultata njezinom uporabom.



Slika 5. Histogrami mjesečnih oborina generiranih ekspancijalnom distribucijom za postaju Čepić

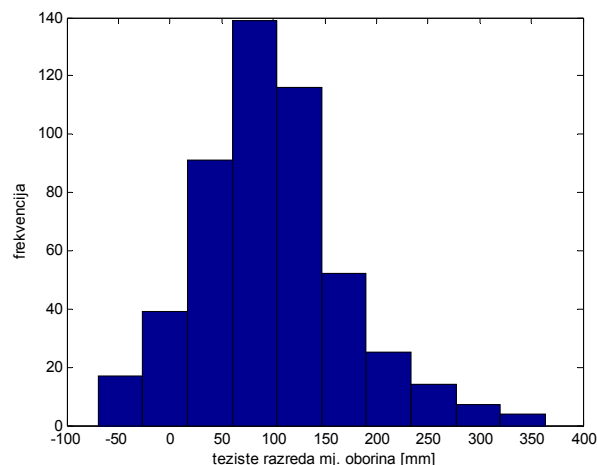


Slika 6. Usporedni histogrami izmjerenih i generiranih siječanjskih i listopadskih oborina za postaju Čepić

3.3 Normalna raspodjela

Sljedeći je korak bio uporaba normalne distribucije s određenim ograničenjima. Ulazni parametri generatora slučajnih brojeva normalne distribucije su srednja mjesečna oborina i standardno odstupanje tog mjeseca (2).

Histogram svih sintetičkih podataka identičnih po dužini trajanja i generiran na osnovi podataka postaje Čepić prikazan je na slici 7.



Slika 7. Histogram učestalosti svih generiranih normalnom distribucijom mjesečnih oborina, sintetički niz generiran na osnovi podataka postaje Čepić

Iz navedenog se prikaza vidi da analizirana raspodjela daje i negativne vrijednosti. Problem „negativnih“ oborina riješen je, kao što je već rečeno (poglavlje 2.1), upotrebom *while* petlje. Generirane vrijednosti manje od „0“ odbacuju se, a petlja provjerava podatke dok god postoje „negativne“ oborine ((3) - shema modela f).

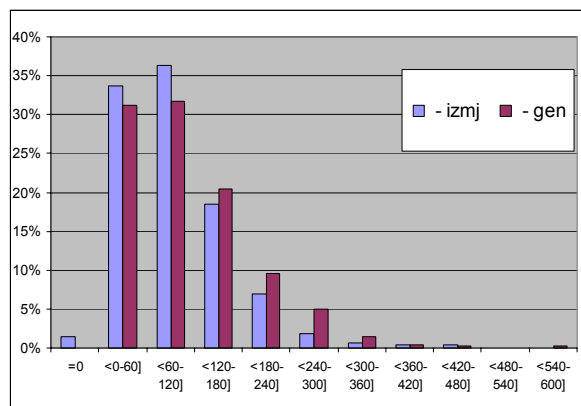
Provedeni su statistički parametarski testovi koji potvrđuju pripadnost istom skupu generiranih i izmjerenih podataka. Tablica 2. prikazuje rezultate parametarskih ispitivanja 100 generiranih serija. Mjesečne vrijednosti za svaki mjesec zadovoljavaju, no godišnje i sve mjesečne vrijednosti ne zadovoljavaju. To je posljedica pomaka distribucije odnosno povećanja srednje vrijednosti zbog prihvaćanja samo pozitivnih generiranih vrijednosti.

Tablica 2. Rezultati parametarskih ispitivanja modela na uzorku od 100 eksperimenata, Čepić (1961.-2002.)

Broj eksperimenata	ODBACIVANJE H_0 [%] (sig=0,05)												GOD	SVI
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII		
Ttest2	5%	6%	5%	0%	3%	0%	2%	3%	4%	5%	2%	1%	98%	97%
Wilcoxon rank sum	4%	4%	1%	0%	2%	0%	2%	4%	5%	8%	2%	1%	97%	83%
Ansari-Bradley	0%	2%	1%	11%	13%	24%	5%	4%	3%	0%	23%	5%	2%	51%
Vartest2	5%	1%	2%	22%	9%	19%	5%	6%	7%	1%	17%	13%	21%	100%
Kolmogorov-Smirnov	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Srednja vrijednost:	3%	3%	2%	7%	5%	9%	3%	3%	4%	3%	9%	4%	44%	66%

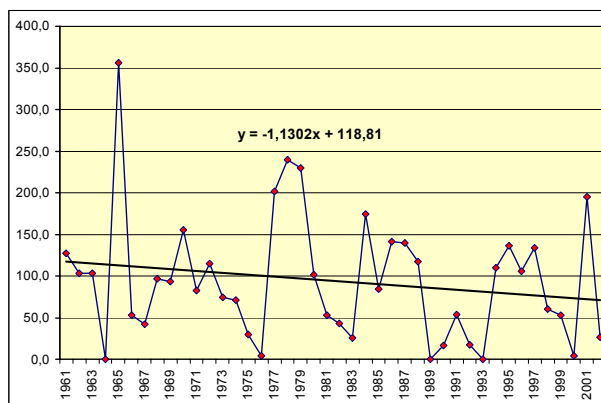
Napomena: GOD – niz godišnjih oborina, dužina niza 42 godine
 SVI – niz svih mjesečnih podataka, dužina niza: 12 x 42 = 504.

Slika 8. prikazuje usporedno histograme jedne generirane i izmjerene serije podataka. Vidi se razlika u učestalostima najvećih i najmanjih vrijednosti.



Slika 8. Histogrami učestalosti mjesečnih oborina, izmjerenih i generiranih normalnom distribucijom s uporabom *while* petlje postaje Čepić

Iduće poboljšanje modela dobiveno je upotrebom srednjih vrijednosti izmjerenog niza u ovisnosti o trendu kao ulaznih parametara. Te srednje vrijednosti dobivene su kao funkcija jednadžbe trenda.



Slika 9. Siječanjске oborine i njihov trend za postaju Čepić (1961-2002)

Ulazni parametri za model su srednja vrijednost (SV_{mj}) i standardno odstupanje (STD_{mj}) izmjerenih podataka za određeni mjesec. Ulazne se srednje vrijednosti mijenjaju ovisno o trendu izmjerenih podataka.

Npr. za mjesec siječanj (Čepić 1961.-2002.) srednja mjesečna oborina je 94,5 mm. Jednadžba trenda siječanjских oborina promatranog razdoblja prikazana je na (5) i slici 9.

$$P_{sij_ulaz}(GOD) = 118,8 - 1,13 \times (1961 - GOD + 1) = 2334 - 1,13 \times GOD \quad (5)$$

To je ujedno i jednadžba ulaznih srednjih vrijednosti.

Dodavanjem ulaznih srednjih vrijednosti u funkciji trenda nije došlo do promjena u rezultatima modela, no dobivena je mogućnost nadopuna nedostajućih i simulacija budućih serija podataka.

Model je poboljšán uvođenjem uvjeta da u nekom mjesecu nije bilo oborina. Broj mjeseci bez oborina teško je statistički korektno interpretirati. Radi se o rijetkim događajima. Primjera radi, za promatrano razdoblje u Mošćeničkoj Dragi nije zabilježen niti jedan mjesec bez oborina, dok je u radu detaljnije analiziranoj postaji Čepić zabilježeno 7 takvih slučajeva (slika 8.). Zbog toga se modeliranju pristupilo na osnovi tromjesečnih vjerojatnosti pojave mjeseca bez oborina.

Za promatrano razdoblje na postaji Čepić u prvom tromjesečju (siječanj – ožujak) zabilježena su ukupno četiri mjeseca kada su oborine bile jednake nuli, što znači da je vjerojatnost pojave mjeseca bez oborina u prvom tromjesečju (6):

$$p(P_{mj} = 0)_{I-III} = 4/42 = 0,095 = 9,5\%$$

Vjerojatnost za pojedini mjesec je trećina kvartalne vjerojatnosti (jednadžba 7):

$$p(P_{mj} = 0)_{mj} = P(P_{mj} = 0)_{kvartal}/3 \quad (7)$$

U modelu vrijednost generirane oborine poprima vrijednost nula ako je rezultat generatora slučajnih brojeva uniformne distribucije manji od vjerojatnosti pojave da u nekom mjesecu nema oborina ((8) - korak g):

$$if(RAND(0,1) < p(P_{mj} = 0)_{mj}) \quad (8)$$

$$P_{mj_gen}(GOD) = 0 \text{ i}$$

$RAND(0,1)$ - generator slučajnih brojeva uniformne distribucije

$(0,1)$ - granice uniformne distribucije.

Rezultati ispitivanja (tablica 3.) prikazuju znatno poboljšanje modela, no lošiji rezultati t-testa dviju varijabli govore nam da srednje vrijednosti generiranog niza imaju prevelika odstupanja od izmjerenog. Prevelika odstupanja medijana pokazuju rezultati *Wilcoxon rank sum* testa.

Odstupanje srednjih vrijednosti izmjerenog i generiranog niza anulirano je autokorekcijom modela. Korekcija ulazne srednje vrijednosti zasniva se na razlici izmjerene i generirane prosječne oborine nekog mjeseca ((9) - korak i):

$$\text{korekcija_SV} = SV(P_{mj_gen}) - SV(P_{mj_izm}) \quad (9)$$

$SV(P_{mj_gen})$ – srednja vrijednost generiranih oborina nekog mjeseca

$SV(P_{mj_izm})$ – srednja vrijednost izmjerenih oborina nekog mjeseca

Na osnovi korekcije srednje vrijednosti korigiraju se ulazne vrijednosti modela, po formuli (jednadžba 10 - korak e):

$$P_{mj_ulaz_kor} = P_{mj_ulaz_kor}(\text{prijašnji korak}) - \text{koef_kor} \times \text{kor_SV} \quad (10)$$

$P_{mj_ulaz_kor}$ - ulazna vrijednost za model – srednja vrijednost nekog mjeseca

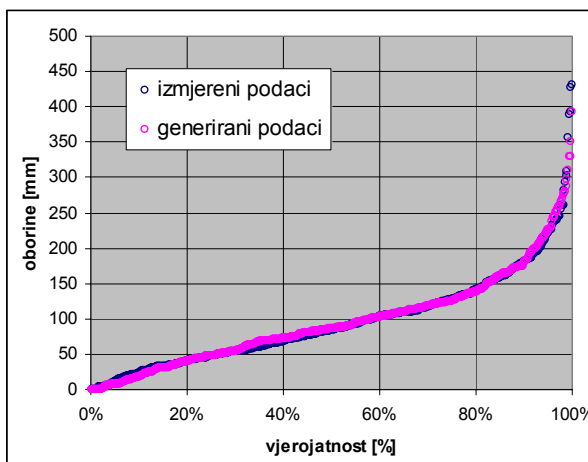
$P_{mj_ulaz_kor}(\text{prijašnji korak})$ - ulazna vrijednost u model iz predhodnog korake petlje koef_kor – koeficijent korekcije srednje vrijednosti.

Koeficijent korekcije srednje vrijednosti zadaje se u parametrima modela, a uveden je zbog modela odnosno manjeg broja koraka petlji autokorekcije.

Tablica 3. Rezultati ispitivanja poboljšanog modela na uzorku od 100 eksperimenata, Čepić

Broj eksperimenata 100	ODBACIVANJE H_0 [%] (sig=0,05)												GOD	SVI
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII		
Ttest2	6%	2%	3%	0%	1%	0%	1%	1%	1%	5%	1%	1%	49%	32%
Wilcoxon rank sum	9%	2%	4%	1%	4%	1%	2%	7%	3%	10%	7%	2%	56%	79%
Ansari-Bradley	3%	26%	7%	1%	0%	1%	2%	1%	0%	27%	2%	2%	1%	16%
Vartest2	7%	0%	1%	1%	1%	1%	8%	4%	8%	5%	2%	5%	1%	37%
Kolmogorov-Smirnov	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Srednja vrijednost:	5%	6%	3%	1%	1%	1%	3%	3%	2%	9%	2%	2%	21%	33%

Spomenuta petlja provjerava zadatke dok razlika srednjih mjesečnih vrijednosti izmjerenog i generiranog niza nije manja od 'mjere autokorekcije srednje vrijednosti' koja se zadaje kao parametar modela. Strožim uvjetom – manjom vrijednošću korekcije srednje vrijednosti dobivamo bolje rezultate. Ako se zada prestrogi uvjet, dolazi do tzv. „pucanja“ modela, odnosno nemogućnosti njegove realizacije.



Slika 11. Empirijska razdioba generiranih i mjenjenih podataka, Čepić

Tablica 4. Rezultati statističkih testova finalnog modela na uzorku od 100 eksperimenata, Čepić (Mjera_KOREKCIJE_SV = 10)

Broj eksperimenata 100	ODBACIVANJE H_0 [%] (sig=0,05)													
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	GOD	SVI
Ttest2	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Wilcoxon rank sum	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%
Ansari-Bradley	0%	16%	1%	1%	1%	0%	1%	0%	0%	0%	0%	0%	1%	0%
Vartest2	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	3%	0%
Kolmogorov-Smirnov	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Srednja vrijednost:	0%	3%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%

Kao primjer prikazani su rezultati ispitivanja generiranih mjesečnih količina oborina za postaju Čepić. Vremenski niz identičan je mjerenom nizu.

Mjera autokorekcije srednje vrijednosti zadana je 10 a koeficijent korekcije 0,25 (jedadžba 10 - korak e). Tablica 4. prikazuje rezultate provedenih statističkih analiza za 100 eksperimenata. U većini slučajeva nije došlo do odbacivanja nulte hipoteze da izmjereni i generirani niz pripadaju istom skupu podataka. Slika 11. prikazuje empirijsku razdiobu svih izmjerenih (504) i svih generiranih (504) mjesečnih oborina za jedan eksperiment. Iz dijagrama se vidi pripadnost istoj distribuciji.

Bolji rezultati su dobiveni strožim izborom mjere autokorekcije srednje vrijednosti, što je u ovom slučaju zadana vrijednost 5. Rezultati su prikazani u tablici 5. iz koje je vidljivo još bolje prihvaćanje nulte hipoteze.

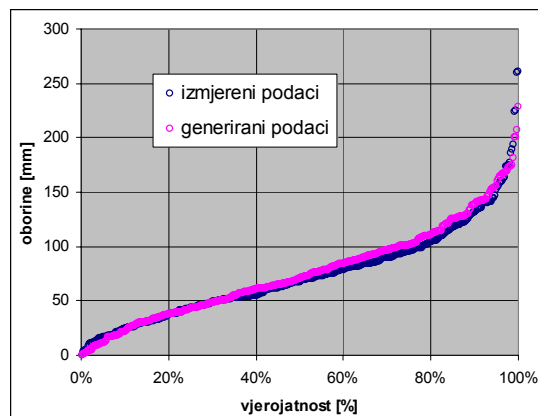
Tablica 5. Rezultati ispitivanju konačnog modela na uzorku od 100 eksperimenata, Čepić (Mjera_KOREKCIJE_SV = 5)

Broj eksperimenata 100	ODBACIVANJE H_0 [%] (sig=0,05)													
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	GOD	SVI
Ttest2	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Wilcoxon rank sum	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Ansari-Bradley	0%	2%	0%	0%	0%	0%	1%	0%	1%	1%	2%	0%	2%	1%
Vartest2	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	3%	0%
Kolmogorov-Smirnov	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Srednja vrijednost:	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%

Tablica 6. Osnovna statistička obrada izmjerenih i generiranih podataka, srednja vrijednost statističkih parametara 100 generiranih serija Čepić

		I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	GOD
		IZMJERENO	S.V.	95	81	84	86	73	89	67	88	114	123	149
	MAX	356	186	206	179	195	213	179	294	427	393	432	261	1679
	MIN	0	4	0	19	5	10	0	0	4	0	11	22	703
	StDev	75,4	56,6	55,7	40,3	42,7	43,2	41,4	62,1	84,5	95,3	85,7	56,1	203
	MEDIAN	88,5	78,8	75,6	84,0	66,5	81,5	52,9	77,0	101,8	105,5	137,2	87,6	1209
	MODE	0,0	36,0	69,0	64,0	44,0	55,0	47,0	86,0	33,0	33,0	10,5	73,0	703
GENERIRANO	S.V.	95	82	85	87	74	89	67	89	115	123	150	100	1155
	MAX	298	213	214	176	175	188	165	242	338	383	349	227	1669
	MIN	0	0	0	11	5	10	2	1	2	2	7	4	693
	StDev	74,5	55,5	54,3	39,8	41,1	42,7	40,1	60,7	84,0	94,4	85,3	55,4	221
	MEDIAN	81,7	76,3	81,2	85,9	71,0	87,2	64,0	80,8	100,5	105,4	143,8	96,8	1152
	MODE	0,3	0,4	0,4	10,6	5,0	10,1	2,1	1,1	2,3	2,3	7,0	4,2	693

Tablica 6. sadrži usporedni prikaz osnovnih statističkih parametara izmjerenih i generiranih podataka. Kod generiranih podataka prikazana je srednja vrijednost parametara 100 generiranih serija.



Slika 12. Histogram učestalosti mjesečnih oborina, izmjerenih i generiranih normalnom distribucijom, Čepić

Tablica 6. sadrži usporedni prikaz osnovnih statističkih parametara izmjerenih i generiranih podataka. Kod generiranih podataka prikazana je srednja vrijednost parametara 100 generiranih serija.

Slika 12. prikazuje histogram učestalosti izmjerenih i generiranih podataka. Generirani podaci su rezultat eksperimenta na indentičnoj vremenskoj seriji izmjerenom nizu.

Tablica 7. Rezultati statističkih testova finalnog modela na uzorku od 100 eksperimenata, Labin, (Mjera_KOREKCIJE_SV=5, Mjera_KOREKCIJE_STD=5)

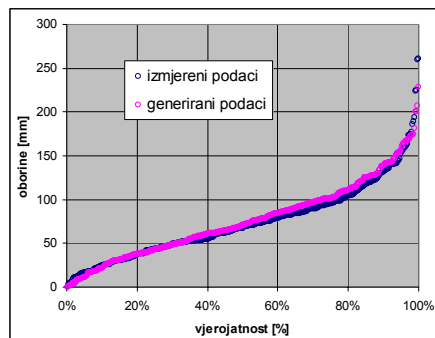
Broj eksperimenata 100	ODBACIVANJE H_0 [%] (sig=0,05)													
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	GOD	SVI
Ttest2	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Wilcoxon rank sum	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Ansari-Bradley	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	4%	0%
Vartest2	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%
Kolmogorov-Smirnov	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Srednja vrijednost:	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%

Na osnovi iznesenog može se zaključiti da je, na analizi ranom primjeru podataka o mjesečnim oborinama s postaje Čepić, provedeno uspješno modeliranje – generiranje sintetičkih serija podataka o mjesečnim količinama oborina uporabom prezentiranoga metodološkoga pristupa. Njegova je verifikacija provedena u nastavku – na primjeru podataka oborinskog režima s nekoliko drugih meteoroloških postaja.

4 Verifikacija modela

Model je ispitan na podacima za meteorološke postaje Labin i Zagreb.

Kod postaje Labin dobiveni su slični rezultati kao i za postaju Čepić, što je i logično s obzirom na njihovu relativnu blizinu od 13 km. Rezultati su prikazani u tablici 7.



Slika 13. Empirijska razdioba generiranih i mjerenih podataka

LITERATURA

- [1] Bonacci, O.: *Prilog proučavanju modela za stvaranje sintetičkih hidroloških nizova*, Vodoprivreda 1/2 (1969), 76-83.
- [2] Bonacci, O.: *Oborine, glavna ulazna veličina u hidrološki ciklus*, GEING, 1., Split, 1994.
- [3] Brezak, S.; Spajić, M.: *Modeliranje malih voda rijeke Drave*. U: Zborniku radova 3. Hrvatska konferencija o vodama (2003), Osijek 28-31.3.2003., Hrvatske vode, 725-730.
- [4] Cox, D. R.; Isham, V.: *Stochastic models of precipitation*, *Statistics for the environment 2*, Water issues, Wiley, New York, 1994.

Model je provjeren i na nešto kraćem nizu (1961.-1990.), meteorološke postaje Zagreb – Grič. Rezultati su prikazani u tablici 8. iz njih se može zaključiti da je model primjenjiv i za područje Zagreba. Slika 13. prikazuje empirijske vjerojatnosti

pojavljivanja izmjerenih podataka i jedne generirane serije podataka.

Tablica 8. Rezultati statističkih testova finalnog modela na uzorku od 100 eksperimenata, Zagreb–Grič, (Mjera_KOREKCIJE_SV=10)

Broj eksperimenata 100	ODBACIVANJE H_0 [%] (sig=0,05)													
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	GOD	SVI
Ttest2	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Wilcoxon rank sum	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%
Ansari-Bradley	0%	7%	0%	3%	0%	0%	0%	1%	0%	0%	0%	0%	1%	0%
Vartest2	0%	2%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%
Kolmogorov-Smirnov	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Srednja vrijednost:	0%	2%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%

5 Zaključak

U radu prikazani model generiranja vremenskih nizova mjesečnih količina oborina pokazao je da je razvijenim postupkom moguće formirati sintetičke nizove podataka koji se nalaze unutar statistički prihvatljivih granica odstupanja. Model uzima u obzir trendove u hodu oborina, kao i mogućnost da se pojave i mjeseci bez oborina.

Provedeni postupak generiranja sastoji se od sljedećih osnovnih koraka:

- zadavanje i obrada izmjerenog oborinskog niza podataka
- generiranje sintetičkih nizova podataka
- korekcija ulaznih parametara
- verifikacija dobivenih rezultata

Model je razvijen u programskom okruženju MATLAB, koraci se mogu vrlo jednostavno programski povezati čime je pojednostavnjen put do rješenja postavljenog zadatka modeliranja sintetičkih vremenskih serija podataka o mjesečnim količinama oborina.

- [5] De Michele, C., P. Bernardara: *Spectral analysis and modeling of space-time rainfall fields*, Atmospheric Research 77, 124-136, 2005
- [6] Margeta, J.; Fističić, I.: *Water quality modeling of Jadro spring*, Water Science & Technology (2004), Vol 50, No 11, 59-66.
- [7] Rubinić, J.; Margeta, J.: *Dimenzioniranje akumulacija primjenom generiranih protoka*, Građevinar 53 (2001) 1, 17-33.
- [8] Šošić, I.: *Primijenjena statistika*, Školska knjiga, 1., Zagreb, 2004.
- [9] Theyer, M. A.: *Modelling long-term persistence in hydrological time series*, The University of Newcastle, 2000