

Cvetko Andreeski¹ UDC 338.48: :[616.98:578.834(100)“2020/2021”

Originalan naučni rad

Primljen: 2. 1. 2022.

Prihvaćen: 11. 3. 2022.

VREMENSKA SERIJA NOĆENJA S UKLJUČENIM STRUKTURNIM PROMENAMA TOKOM PANDEMIJE KOVID-19: IZBOR OPTIMALNOG MODELA

APSTRAKT: Turizam je izložen raznim rizicima kao što su prirodne katastrofe, sve vrste kriza, negativna propaganda, pandemije i sl. Svi ovi rizici utiču ne samo na razvoj turizma već i na vremenske serije koje su predmet analize. Za donošenje odluka o radu i daljem razvoju turističkog sektora, potrebno je izvršiti tačnu analizu postojećeg stanja, ali i predvideti buduće vrednosti serije. Zbog navedenih rizika, ovaj složeni problem zahteva dubinsku analizu, izbor odgovarajućeg modela i testiranje rezultata i predviđanja modela. U radu se analiziraju vremenske serije u oblasti turizma za period pre i tokom pandemije, analiziraju se strukturne promene, kao i mogući modeli koji se ubuduće mogu koristiti za modeliranje takvih serija.

KLJUČNE REČI: vremenske serije, strukturne promene, analiza, modeli, turizam

1. Uvod

Turizam je bio jedna od najteže pogođenih privrednih delatnosti tokom pandemije KOVID-19. Po proglašenju pandemije u Evropi u martu 2020. godine, ugostiteljski i smeštajni objekti su zatvoreni, što

¹ Full Professor, Faculty of Tourism and Hospitality Ohrid, email: cvetko.andreeski@uklo.edu.mk

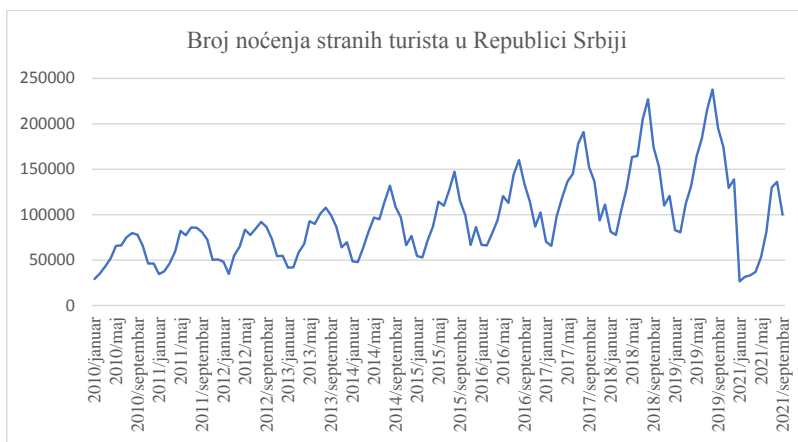
je onemogućilo redovno funkcionisanje ovog sektora. Prema podacima Svetske banke, 2019. godine, godinu dana pre izbijanja pandemije, izvozni efekat turizma u Srbiji iznosio 7,65 odsto BDP-a. U 2020. godini, međutim, ukupan broj noćenja iznosio je samo 24% od noćenja u 2019. godini, a prihodi su pali za više od 147 miliona dolara.

Značajne i dugotrajne strukturne promene u razvoju turizma tokom pandemije dovele su sektor na ivicu održivosti; štaviše, tokom pandemije istraživači nisu bili u mogućnosti da modeliraju analize vremenskih serija za razvoj turizma i predvide buduće vrednosti (O'Hare & Li, 2015; Asghar & Amena, 2012). Do danas je obavljeno mnogo istraživanja sa ciljem da se napravi validan model za identifikaciju i predviđanje budućih vremenskih serija podataka u oblasti razvoja turizma (Andreeski & Mechkaroska, 2020; Petrevska, 2017; Baldigara & Mamula, 2015), ali se čini da ovi modeli nisu ni od kakve koristi kada uzmemo u obzir ogromne strukturne promene kao što je pandemija KOVID-19. Pored linearnih modela za identifikaciju i predviđanje, koriste se složeniji nelinearni modeli, kao što su modeli zasnovani na veštačkim neuronskim mrežama, koji mogu da obuhvataju i promene nivoa i varijanse serije (Andreeski i Petrevska, 2021; Shi, 2019).). Kao što navode But i Pik (Boot & Pick, 2019), predviđanje može biti nešto tačnije ako se koristi uzorak nakon prekida, a ne puni uzorak. Međutim, ovaj rad će razmatrati samo kraće prekide, a ne trajne prekide. U nadi da će dobiti tačnije rezultate, neki istraživači su koristili tehnike dubokog učenja i analize velikih podataka. Ipak, svi ovi modeli se zasnivaju na pretpostavci da će buduće vrednosti biti zasnovane na trendovima i varijansama iz prošlosti, tj. da će budućí trendovi odražavati one iz prošlosti (Kaushnik et al., 2019; Jian et al., 2021). Međutim, čak ni ovi modeli ne mogu dati validne rezultate ako nije dostupna dovoljno duga serija podataka od početka pandemije. Neke studije modeliranja i predviđanja vremenskih serija turizma tokom pandemije ispituju trenutnu situaciju i porede je sa prognoziranim vrednostima serije razvoja turizma pre pandemije, sa ciljem da se izračunaju gubici nastali usled pandemije u ovom sektoru (Andreeski & Petrevska, 2021; Šenkova et al., 2021). Nedavno su u nekim publikacijama objavljeni predloženi modeli za identifikaciju i predviđanje budućih vrednosti serije razvoja turizma (Provenzano & Volo, 2021; Šenkova et al., 2021).

Prošlo je skoro dve godine od izbijanja pandemije, što je dovoljno vremena da se pokuša modeliranje analize vremenske serije, kako bi se mogli predvideti budući podaci. Prema preporukama za modelovanje serije dovoljnih podataka, potrebno je da prođe najmanje četiri godine, pod uslovom da su dostupni mesečni podaci. U ovom radu je predložen jedan takav model, zasnovan na linearnom SARIMA modelu, tokom kojeg se vrši određena predobrada vremenske serije. Ovaj model je izabran jer je relativno jednostavan i pruža mogućnost da se njegove performanse uporede sa performansama sličnih modela koji se koriste u studijama o modeliranju vremenskih serija u oblasti turizma. Model je kreiran na osnovu podataka o noćenjima stranih turista u Republici Srbiji deset godina pre pandemije i godinu dana nakon nje.

2. Podaci i modeliranje

Na Slici 1 prikazan je broj noćenja stranih turista u Republici Srbiji 2010–2019, kao i statistički podaci za period januar–septembar 2021. Podaci prikazani na Slici 1 preuzeti su iz Republičkog zavoda za statistiku.



Slika 1. Broj noćenja stranih turista u Republici Srbiji

Na osnovu prikazanih podataka može se uočiti nekoliko karakteristika ove serije:

- Serija ima definisan trend u periodu od 2010. do 2019. godine, a zatim se vidi značajan pad u 2021. godini. Ipak, oblik serije je i dalje konstantan.

- Serija ima promenu varijanse, odnosno izraženu heteroskedastičnost.

- Sezonska karakteristika serije očigledna je u čitavom analiziranom periodu.

- Verovatna je strukturna promena u 2021. godini zbog promene trenda.

U seriji su namerno izostavljeni podaci za 2020. godinu, jer je ne treba uzimati kao referencu za broj dolazaka turista, pošto bi to onemogućilo modeliranje serije i predviđanje budućih podataka.

Za modeliranje serije korišćen je model SARIMA. U delu predobrabe serije urađen je logaritam originalne serije, a nakon toga i diferencijacija serije kako bi se dobila stacionarna serija koja se može modelovati. Test jediničnog korena izvršen je za ovu modifikovanu seriju da bi se identifikovala stacionarnost. Tabela 1 prikazuje rezultat ovog testa.

Tabela 1. Test jediničnog korena modifikovane serije

Nulta hipoteza: NOĆENJA_STRANI_DLOG ima jedinični koren

Egzogeni: konstantni

Dužina kašnjenja: 10 (Automatski – bazirano na SIC-u, maksimalno kašnjenje = 12)

		T-statistika	Verovatnoća
Statistika proširenog Diki-Fulerovog testa		-19,04046	0.0000
Kritične vrednosti:	1% nivo	-3,491928	
	5% nivo	-2,888411	
	10% nivo	-2,581176	

*MacKinnon (1996) jednostrane p-vrednosti.

Tabela 1 pokazuje da je vrednost proširenog Diki-Fulerovog testa znatno niža od kritičnih vrednosti za stacionarni niz, a verovatnoća odbacivanja hipoteze da je serija stacionarna manja od 1%.

Pre nego što pređemo na modeliranje serije, potrebno je identifikovati strukturnu promenu koja će verovatno nastupiti 2021. godine. U tu svrhu napravljen je test loma korena jedinice u nizu. Rezultati ispitivanja su dati u Tabeli 2.

Tabela 2. Test loma korena jedinice modifikovane serije

Nulta hipoteza: NOĆENJA_STRANI_DLOG ima jedinični koren

Specifikacija trenda: samo presretanje

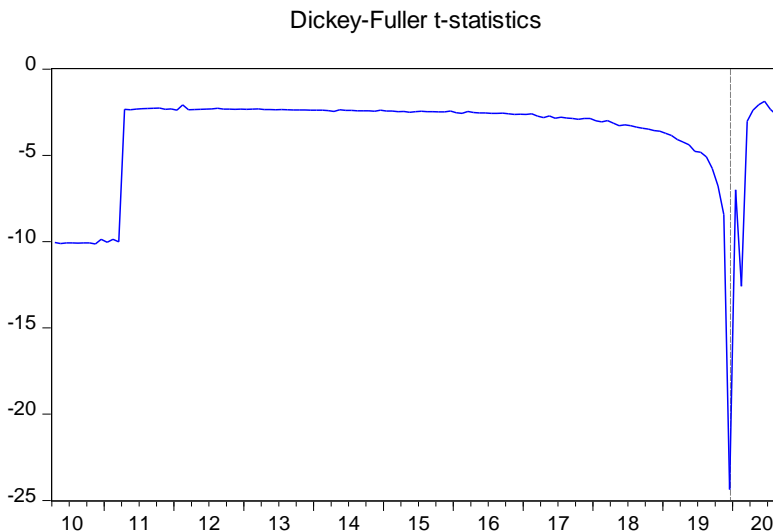
Specifikacija prekida: samo presretanje

Tip preloma: inovativni izuzetak

Datum prekida: 2019M12

Izbor pauze: minimalizovana Diki-Fuler t-statistika

Dužina kašnjenja: 10 (automatski – na osnovu Švarcovog kriterijuma info maksimalno kašnjenje = 12)



Slika 3. Grafički prikaz strukturnih promena prema Diki-Fulerovoj t-statistici

Iz rezultata datih u Tabeli 2, može se zaključiti da postoji strukturna promena serije na kraju 2019. godine, odnosno detektovana je strukturna promena serije nakon 2019. godine.

Nakon stacioniranja serije i testiranja postojanja strukturnih promena, modeliranje se može nastaviti. Za određivanje nezavisnih varijabli serije napravljena je korelacija serije. Vrednosti korelograma su date u Tabeli 3.

Tabela 3. Korelogram modifikovane serije

Sample: 2010M01 2020M09
Included observations: 128

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.100	0.100	1.3171	0.251
		2	0.191	0.183	6.1325	0.047
		3	-0.003	-0.039	6.1336	0.105
		4	-0.160	-0.200	9.5639	0.048
		5	-0.229	-0.209	16.671	0.005
		6	-0.585	-0.553	63.330	0.000
		7	-0.180	-0.165	67.799	0.000
		8	-0.127	0.027	70.022	0.000
		9	-0.066	-0.103	70.630	0.000
		10	0.070	-0.168	71.316	0.000
		11	0.137	-0.201	74.001	0.000
		12	0.603	0.365	126.16	0.000

Korelogram pokazuje da šesto i dvanaesto kašnjenje imaju najveće vrednosti, koje ukazuju na sezonsku komponentu serije. Neophodno je testirati serijsku korelaciju u seriji, ali to je malo verovatno jer prvo kašnjenje ima malu vrednost.

Za seriju je napravljeno nekoliko različitih konkurentskih modela, a u radu su prikazani samo oni koji daju najbolje rezultate. Prvi model ima samo jednu nezavisnu varijablu, a to je dvanaesto kašnjenje u seriji. Rezultati su prikazani u Tabeli 4.

Tabela 4. Model serije turističkih noćenja u Republici Srbiji

Zavisna varijabla: NOĆENJA_STRANI_DLOG

Metod: ARMA maksimalna verovatnoća (OPG – BHHH)

Datum: 7. 11. 2021. Vreme: 09.21

Uzorak: 2010M02 2020M09

Broj obuhvaćenih opservacija: 128

Konvergenција je postignuta nakon 11 iteracija

Kovarijansa koeficijenta izračunata korišćenjem spoljašnjeg proizvođa gradijenata

Varijabla	Koeficijent	St.greška	T-statistika	Verovatnoća
AR (12)	0,877828	0,072841	12,05135	0,0000
SIGMASQ	0,023934	0,001189	20,12629	0,0000
R-kvadrat	0,642378	Srednja zavisna var.		0,009632

Rezultati pokazuju da je nezavisna varijabla validna, tj. da ima visoku vrednost t-statistike i malu verovatnoću da će ovaj parametar biti uklonjen iz modela. Stepenn varijanse originalne serije iznosi oko 64%, a vrednost Durbin-Vatson statistike blizu je 2, što ukazuje da ne postoji značajna serijska korelacija reziduala.

Da bi se model poboljšao, napravljena je lažna varijabla koja ima za cilj da modelira strukturnu promenu. Ova promenljiva ima vrednost 1 za mesec 2021. godine, a za podatke prethodnih godina ima vrednost 0. Rezultati modelovanja ovog modela prikazani su u Tabeli 5.

Iz rezultata datih u Tabeli 5 može se zaključiti:

- Lažna varijabla je relevantna za model, a verovatnoća izbacivanja ovog parametra iz modela manja je od 1%.
- Ovaj model pokriva većinu varijanse modela.
- Informacioni kriterijum ima nižu apsolutnu vrednost u odnosu na prethodni model.

Vrednost Durbin-Vatson statistike ima vrednost bližu 2 u poređenju sa prethodnim modelom. Za model je napravljena korelacija reziduala

koja pokazuje da su svi reziduali unutar intervala poverenja $\pm 2se$. U nastavku je data prognoza budućih vrednosti vremenske serije, sa prognozom budućih vrednosti na osnovu izabranog modela.

Tabela 5. Model serije turističkih noćenja u Republici Srbiji sa dodatnom lažnom varijablom

Zavisna varijabla: NOĆENJA_STRANI_DLOG

Metod: ARMA maksimalna verovatnoća (OPG – BHHH)

Datum: 7. 11. 2021. Vreme: 09.21

Uzorak: 2010M02 2020M09

Broj obuhvaćenih opservacija: 128

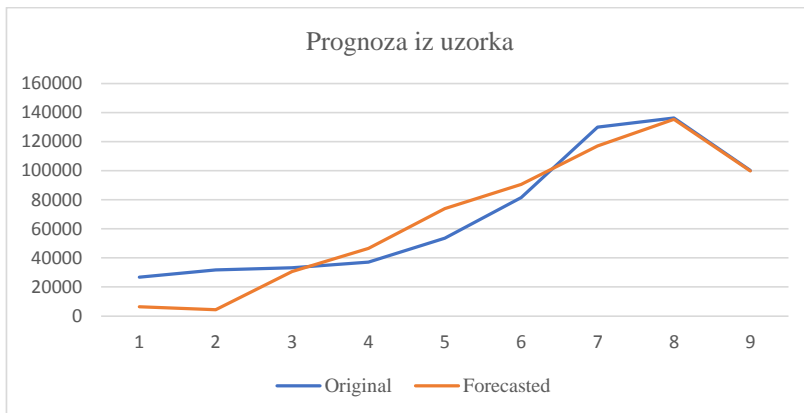
Konvergencija je postignuta nakon 7 iteracija

Kovarijansa koeficijenta izračunata korišćenjem spoljašnjeg proizvoda gradijenata

Varijabla	Koeficijent	St. greška	T-statistika	Verovatnoća
DUMMY	-0,083368	0,026179	-3,184512	0,0018
AR(12)	0,882392	0,071973	12,26012	0,0000
SIGMASQ	0,023367	0,001553	15,04329	0,0000
R-kvadrat	0,650839	Srednja zavisna var.		0,009632
Prilagođeni R-kvadrat	0,645253	Stand. devijacija zavisne var.		0,259714
Stand. greška regresije	0,154687	Akaike info kriterijum		-0,730304
Ukupni kvadrat rezid.	2,991016	Schwarz kriterijum		-0,663460
Log verovatnoće	49,73946	Hannan-Quinn kriterijum		-0,703145
Durbin-Watson stat.	2,264859			

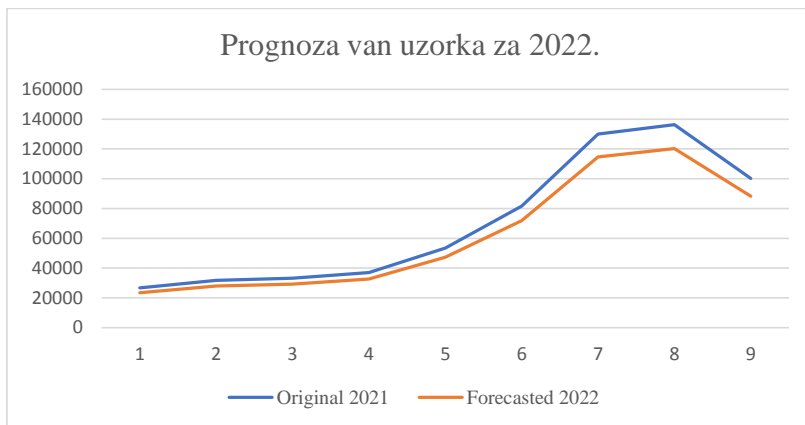
3. Predviđanje serije

Da bi se izvršilo predviđanje, serija je produžena za 12 meseci i izvršeno je vraćanje na predviđene vrednosti kako bi odgovarale vrednostima originalne serije. Za potrebe istraživanja, napravljena je *in-sample* prognoza vrednosti iz uzorka (*in-sample*), a uzete su poznate vrednosti iz 2021. godine, kao što je prikazano na Slici 4.



Slika 4. Prognoza iz uzorka vrednosti za 1. 9. 2021.

Na kraju, napravljena je prognoza vrednosti očekivanog broja stranih turista za 2022. godinu, opet za prvih devet meseci. Ovi rezultati su prikazani na Slici 5. Isti grafikon prikazuje seriju vrednosti za 2021. godinu, radi poređenja dve serije.



Slika 5. Prognoza van uzorka vrednosti za 1. 9. 2021.

Lažna varijabla nije korišćena za izračunavanje vrednosti u 2022. godini, jer nije verovatno da će doći do još neke strukturne promene u

ovoj godini. U suprotnom, nikakve prognoze se ne bi mogle napraviti. Kao što pokazuje grafik, za 2022. godinu predviđa se manji broj turista u odnosu na 2021. godinu, što je malo verovatno. Međutim, model vrši proračun na osnovu celokupne istorije za analiziranu seriju i vrednost parametra α (12) manja je od 1, što znači da buduće vrednosti treba da imaju niže vrednosti od prethodnih. Procenjuje se da je pad noćenja oko 8%.

4. Zaključna razmatranja

U analizi se koristi vremenska serija podataka o noćenjima stranih turista u Republici Srbiji. Ovo je relevantan model vremenske serije, koji se može koristiti za praćenje pandemije KOVID-19 na razvoj turizma u bilo kojoj zemlji. Serija ima strukturnu promenu koja čini njeno modeliranje složenijim. Pravilnom prethodnom obradom serije i dodavanjem lažne promenljive u modeliranje, kreiran je validan model. Konačno, prognoza budućih vrednosti serije pokazuje niže vrednosti za narednu godinu (2022) u odnosu na prethodnu (2021), što će se verovatno i dogoditi. Međutim, značaj ovog modela je u tome što je pokazao da je moguće napraviti validan model čak i za serije koje predstavljaju izazov za modeliranje. Kada bude dovoljno podataka iz serije nakon strukturne promene, biće moguće napraviti pouzdaniji model zasnovan na podacima nakon pandemije. Očekuje se da će ovaj model pružiti preciznije predviđanje budućih vrednosti.

Literatura

- Andreeski, C. & Mechkaroska, D., (2020). Modelling, Forecasting and Testing Decisions for Seasonal Time Series in Tourism. *Acta Polytechnica Hungarica* 17 (10), pp. 149–171.
- Andreeski, C. & Petrevska, B., (2021). *FORECASTING DYNAMIC TOURISM DEMAND USING ARTIFICIAL NEURAL NETWORKS*. Skopje, s.n.
- Asghar, Y. & Amena, U., (2012). Structural Breaks, Automatic Model Selection and Forecasting Wheat and Rice Prices for Pakistan. *Pakistan Journal of Statistics and Operation Research*, pp. 1–20.
- Baldigara, T. & Mamula, M., (2015). MODELLING INTERNATIONAL TOURISM DEMAND. *Tourism and Hospitality Management*, Vol. 21, No. 1., pp. 19–31.
- Boot, Tom; Pick, Andreas (2019). Does modeling a structural break improve forecast accuracy?. *Journal of Econometrics*, (), S0304407619301824-. doi:10.1016/j.jeconom.2019.07.007
<https://data.stat.gov.rs/Home/Result/220203?languageCode=en-US>. [Online]
<https://data.worldbank.org/indicator/ST.INT.TRNX.CD?locations=RS>. [Online]
- Jian, B. W., Hui, L. & Zhi, F. P., (2021). Tourism demand forecasting with time series imaging: A deep learning model. *Annals of Tourism Research*, Volume 90.
- Kaushnik, R., Jain, S. & Dash, T., (2019). Performance evaluation of deep neural networks for forecasting time-series with multiple structural breaks and high volatility. *Computer Science, Mathematics (arXiv preprint arXiv:1911.06704)*.
- O’Hare, C. & Li, Y., (2015). Identifying Structural Breaks in Stochastic Mortality Models. *SSRN electronic Journal*.
- Petrevska, B., (2017). Predicting tourism demand by ARIMA models. *Economic research-Ekonomska istraživanja*, pp. 939-950.
- Provenzano, D. & Volo, S., (2021). Tourism recovery amid COVID-19: The case of Lombardy, Italy. *Tourism Economics*, <https://doi.org/10.1177/135481662111039702>.
- Šenkova, A. et al., (2021). Time Series Modeling Analysis of the Development and Impact of the COVID-19 Pandemic on Spa Tourism in Slovakia. *Sustainability* 13(20) 11476.
- Shi, X., (2019). Tourism culture and demand forecasting based on BP neural network. *Personal and Ubiquitous Computing*, pp. 299–308.