

**KRATKOROČNA PROGNOZA POTROŠNJE ELEKTRIČNE ENERGIJE PUTEM
MODIFIKOVANOG KNN ALGORITMA I ARIMA STATISTIČKOG MODELA****SHORT-TERM FORECAST OF ELECTRICITY CONSUMPTION THROUGH MODIFIED KNN
ALGORITHM AND ARIMA STATISTICAL MODEL**

Vladimir UROŠEVIĆ, Univerzitet u Beogradu, Generali Development, Srbija

Željko MARKOVIĆ, Univerzitet u Beogradu – Elektrotehnički fakultet, Deloitte d.o.o. Beograd, Srbija

KRATAK SADRŽAJ

Prognoza potrošnje električne energije predstavlja prvi i najvažniji korak u upravljanju potrošnjom, donošenju investicionih odluka u trgovini električnom energijom i planiranju sistema u uslovima liberalizovanog tržišta električne energije. U zavisnosti od vremenskog horizonta na kom se vrši prognoza, prognozi se pristupa putem različitih metoda od jednostavnih algoritma za klasifikaciju i predviđanje, različitih statističkih metoda regresije, „fuzzy“ logike, neuronskih mreža, do hibridnih sistema koji kombinuju više pristupa.

U radu se ispituju dva metodološka pristupa. U okviru prvog pristupa razmatra se klasičan KNN (k-Nearest Neighbors) algoritam čiji je ulaz samo prognozirana temperatura. Potom se ispituje predlog modifikovanog KNN (k-Nearest Neighbors) algoritma koji osim temperature uzima i ostvarene dijagrame potrošnje iz bliske prošlosti. Drugi pristup podrazumeva primenu ARIMA statističkog modela. Ova dva pristupa se koriste za kratkoročnu prognozu satne potrošnje električne energije, na nivou dva dana unapred, na konzumu Republike Srbije. U radu se, na osnovu konkretnih rezultata, obrazlažu prednosti i nedostaci svakog od izabranih modela.

Ključne reči: Prognoza potrošnje električne energije, KNN, ARIMA**ABSTRACT**

Forecasting of electricity demand is the first and the most important step in energy management, investment decisions in electricity trading, and planning of electrical systems in a liberalized electricity market. Depending on the forecast time horizon, the different methods are in use: from simple classification and prediction algorithms, different statistical regression methods, fuzzy logic, neural networks, to hybrid systems that combine multiple approaches.

This paper examines two methodological approaches. The first approach considers the classical KNN (k-Nearest Neighbors) algorithm whose input is only the predicted temperature. Then the proposal of the modified KNN (k-Nearest Neighbors) algorithm is examined, which, in addition to temperature, takes into account the actual consumption diagrams from the recent past. The second approach involves applying the ARIMA statistical model. These two approaches are used for the short-term forecast of hourly electricity consumption, at the level of two days ahead, for the electricity demand of the Republic of Serbia. Based on concrete results, the paper, discusses the advantages and disadvantages of each of the selected models.

Key words: Electricity Demand Forecast, KNN, ARIMA

Željko Marković, zmarkovic@deloittece.com

UVOD

Uvođenjem tržišta električne energije, tačna procena, odnosno prognoza, kako kratkoročnih, tako i srednjoročnih i dugoročnih potreba za električnom energijom, postaje naročito važna u cilju optimizacije i održivosti poslovanja. Iz tog razloga, u poslednje dve decenije, utemeljen je veliki broj metoda za prognozu, koje se koriste različitim pristupima, a u cilju modeliranja i predviđanja potražnje za električnom energijom, kako analizom istorijskih podataka o potrošnji, tako i na bazi korelacije sa drugim istorijskim parametrima.

Većina tih metoda se zasniva na analizi vremenskih serija i stohastičkih procesa kao što su: autoregresivni pokretni proseci (ARMA) (1), sezonski autoregresivni integrisani pokretni proseci (SARIMA) (2), jednostruka i višestruka regresiona analiza (3), (4), splajn interpolacija (5), modeli sa jednostrukim i višestrukim eksponencijalnim izravnavanjem (6), metoda haotičnih vremenskih serija (7), linearni regresioni modeli (8), sezonski hibridni nelinearni modeli (9) i slično.

Određeni broj metoda koristi i upotrebu veštačke inteligencije (10), neuronske mreže (11), kao i integraciju veštačkih neuronskih mreža i drugih algoritama kao što su: genetički algoritam (12), fazi – logika (13), ekspertski sistemi (14) i slično.

Sama činjenica da je razvijen veliki broj pristupa za modelovanje i prognozu potrošnje električne energije, dokazuje da je veoma komplikovano i teško vršiti prognozu potrošnje uz očekivanje visoke tačnosti. Svaka kategorija metoda ima svoje teorijske prednosti i ograničenja i nijedna od njih nije identifikovana kao generalno superiorna u odnosu na druge. Na dalje, zajednička karakteristika većine ovih metoda je upotreba komplikovanog nelinearnog strukturnog ili parametarskog modeliranja vremenskih serija, što u samom startu, usled složenosti implicira i nedostatak razumevanja i transparentnosti metode, a samim tim i verodostojnosti prognoze. Takođe, izbor metode zavisi i od vremenskog intervala prognoze; tako, neke metode daju bolje rezultate ukoliko se vrši kratkoročno planiranje, kao što su satne prognoze na nivou dana unapred, dok druge pak daju bolje rezultate pri srednjeročnim ili dugoročnim prognozama. Važno je naglasiti da na preciznost prognoze utiče i veličina potrošnje; naime, jasno je da greška prognoze opada se veličinom konzuma za koji se prognozira potrošnja.

Izbor metode zavisi i od vrste i karakteristika potrošnje, odnosno od kategorije kupaca čija se potrošnja električne energije prognozira. Tako potrošnja kupaca iz kategorije domaćinstava u velikoj meri zavisi od temperature i doba dana. Slično važi, doduše u nešto manjoj meri i za privredne subjekte koji obavljaju administrativne poslove, odnosno čiji se poslovni proces odvija u kancelarijama. Sa druge strane, potrošnja kupaca iz kategorije privrednih subjekata koji imaju industrijsku proizvodnju skoro uopšte ne zavisi od temperature, već zavisi od karakteristika tehnoloških procesa, kao i od doba dana u skladu sa tim da li se radi u jednoj, dve ili tri smene. U pojedinim industrijama, veoma često potrošnja električne energije može imati sve odlike čisto stohastičkih procesa.

Najčešći slučaj, naročito kada je u pitanju deo distributivnog elektroenergetskog sistema, je potrošnja električne energije koja predstavlja složeni skup više različitih potrošnji, pa tako u sebi sadrži i komponentu koja je temperaturno zavisna, kao i komponentu koja ne zavisi od temperature. U takvim slučajevima, u cilju poboljšanja prognoze, a ukoliko se raspolaže sa drugim neophodnim podacima, može se izvršiti i dekompozicija potrošnje na više komponenti. Potom se vrši prognoza za svaku komponentu odgovarajućim metodama a potom dolazi do ukupne prognoze.

U ovom radu autori ispituju dva metodološka pristupa, koje autori imaju nameru da u narednim istraživanjima koriste u razvoju hibridnog modela koji će koristiti i pristup putem neuralnih mreža. U okviru prvog pristupa razmatra se klasičan KNN (k-Nearest Neighbors) algoritam čiji je ulaz samo prognozirana temperatura i sagledavaju se njegove slabosti, pa se radi prevazilaženja nedostataka predlaže i ispituje predlog modifikovanog KNN (k-Nearest Neighbors) algoritma, koji osim temperature uzima i ostvarene dijagrame potrošnje iz bliske prošlosti, kao i druge meteorološke parametre koji bi mogli biti od značaja za preciznost prognoze. Drugi pristup podrazumeva primenu autoregresivnih integrisanih pokretnih proseka, odnosno primenu ARIMA statističkog modela. Ova dva pristupa se koriste za kratkoročnu prognozu satne potrošnje električne energije, na nivou dva dana unapred, na konzumu Republike Srbije.

U radu će se najpre dati kratak opis ulaznih podataka koji se koriste i navesti izvori iz kojih su podaci preuzeti. Potom će biti opisana oba pristupa koji se koriste – algoritam K-najbližih suseda sa modifikacijama koje su uneli autori i metoda autoregresivnih integrisanih pokretnih proseka (ARIMA). Na kraju, daće se zapažanja o preciznosti prognoze na testnom uzorku za oba ova pristupa.

OPIS ULAZNIH PODATAKA KOJI SE KORISTE

Kao ulazni set podataka, korišćene su vrednosti ostvarenih satnih potrošnji električne energije u Srbiji, koje su ostvarene u periodu od 01.01.2015. godine do 31.12.2019. godine. Podaci su preuzeti sa sajta Evropske mreže operatora prenosnih elektroenergetskih sistema, ENTSO-e Transparency platform¹, dok su odgovarajući meteorološki podaci preuzeti sa sajta RHMZ². Od navedenih podataka je strukturirana tabela koja sadrži: sat, tip dana, dan, mesec, godinu, minimalnu, maksimalnu i srednju temperaturu, temperaturu subjektivnog osećaja, brzinu vetra, vlažnost, pritisak, tačku rosišta i satne vrednosti snage i potrošnje električne energije. Podaci u periodu od 01.01.2015. godine do 30.06.2018. godine su korišćeni kao trening podaci, dok su podaci u periodu

¹ <https://transparency.entsoe.eu/>

² <http://www.hidmet.gov.rs/>

od 01.07.2018. do 31.12.2019. godine korišćeni kao set za validaciju, s tim što su prognoze rađene na period od 48 časova, tako da je za oba algoritma trening set sukcesivno proširivan sa podacima za sate, za koje je prethodno urađena prognoza.

ALGORITAM K-NAJBLIŽIH SUSEDA

Algoritam k-najbližih suseda, kraće KNN (k-Nearest Neighbors) se koristi za prepoznavanje obrazaca i predstavlja neparametarski metod koji se koristi za klasifikaciju i regresiju (15). Kako u slučaju klasifikacije, tako i u slučaju regresije, ulaz se sastoji od k najbližih test primera u funkciji prostora, dok izlaz zavisi upravo od toga, da li se KNN koristi za klasifikaciju ili regresiju. Kada se KNN koristi za klasifikaciju, izlaz je član klase. Objekat se klasifikuje glasovima većine svojih suseda, tako da objekat bude raspoređen u klasu najčešću među svojim k najbližim susedima (k je pozitivan ceo broj, najčešće mali broj). Ako je $k = 1$, onda se objekat jednostavno dodeljuje klasi tog jednog najbližeg suseda. U KNN regresiji, izlaz predstavlja vrednost objekta. Ova vrednost predstavlja prosek vrednosti njegovih k najbližih suseda.

Dalje se kod KNN algoritma može vršiti i dodela težine doprinosima suseda, što može biti od koristi i za klasifikaciju i za regresiju, tako da bliži susedi doprinose proseku više u odnosu na one udaljenije. Susedi se uzimaju iz skupa objekata za koje je klasa (za KNN klasifikaciju) ili vrednost objekta (za KNN regresiju) poznata. Ovo se može posmatrati kao test skup za algoritam iako eksplicitni test koraci nisu neophodni.

U fazi klasifikacije, k je konstanta definisana od strane korisnika, a neoznačeni vektor (upit ili test primer) se klasifikuje dodelom oznake koja je najučestalija među k test primerima najbližim datom upitu. Kada se vrši klasifikacija gde su promenljive kontinualne, odnosno neprekidne, za ocenu se najčešće koristi Euklidovo rastojanje.

Mana klasične klasifikacije "glasanje većine" dešava se kada je klasna distribucija iskrivljena. Tako primeri učestalijih klasa teže dominaciji pri predviđanju novih primera, jer teže da, uprkos njihovom velikom broju, budu česti među k najbližim susedima. Jedan od načina da se prevaziđe ovaj problem jeste da se doda težina klasifikaciji, uzimajući u obzir rastojanje od test tačke do svakog od k najbližih suseda. Klasa (ili vrednost, u problemima sa regresijom) svake od k najbližih tačaka množi se težinom koja je obrnuto proporcionalna rastojanju od date tačke do test tačke. Još jedan od nedostataka KNN algoritma je poznat pod nazivom "prokletstvo dimenzionalnosti". Fenomen prokletstva dimenzionalnosti govori da kod visoko dimenzionalnih prostora, udaljenosti između najbližih i najudaljenijih tačaka postaju gotovo jednake (16), pa se u cilju izbegavanja ovog problema primenjuju mnogobrojne metode redukcije dimenzija ili indeksiranja atributa.

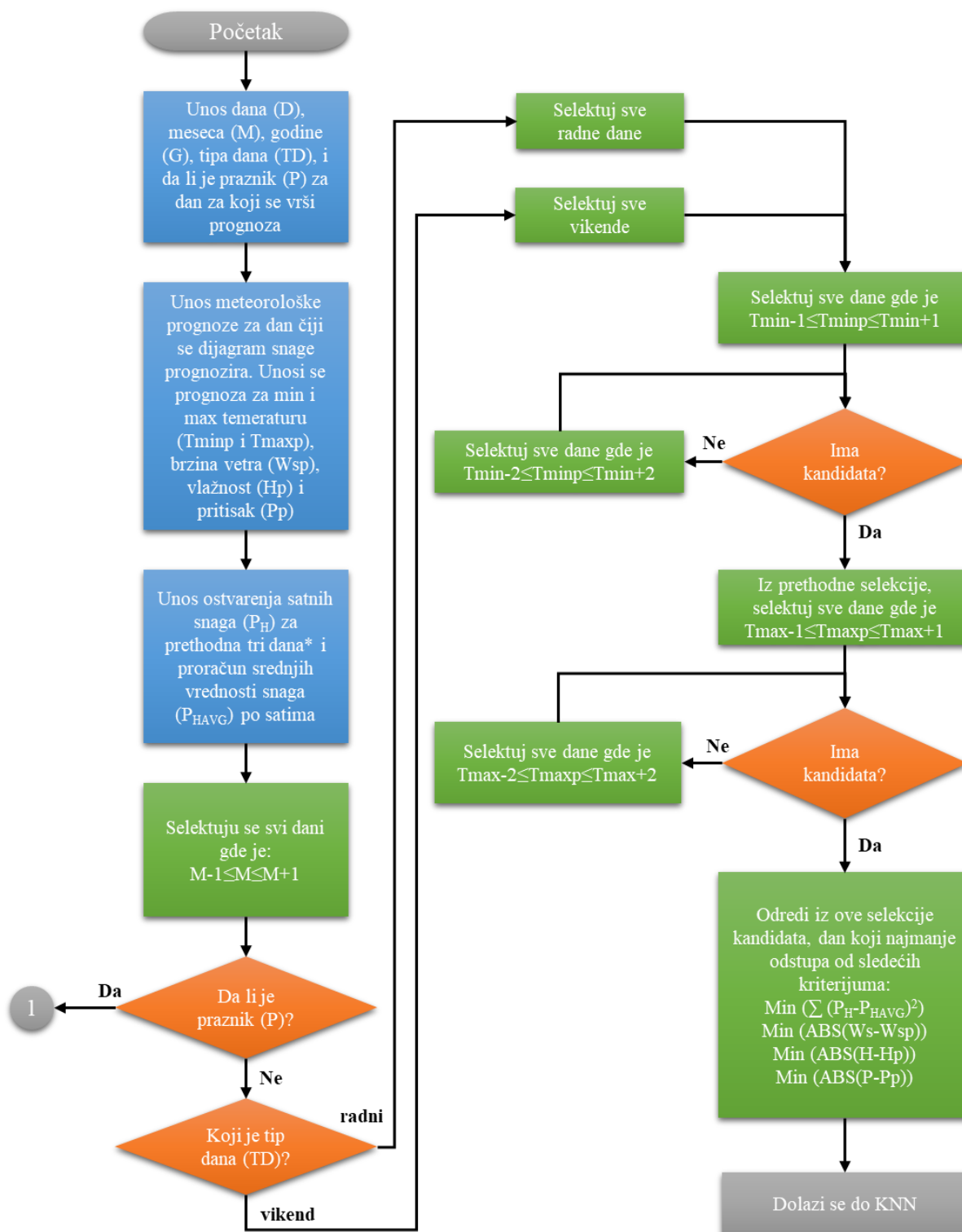
Za potrebe prognoze satnih potrošnja jednog dnevnog dijagrama, pri klasičnom pristupu primene KNN algoritma, kao ulazni podatak bi se koristila prognozirana prosečna temperatura za dan za koji se izrađuje prognoza satnih potrošnji (ujedno i snaga). Algoritam potom bira dnevni dijagram koji je ostvaren pri identičnoj ili najbližoj temperaturi u skupu poznatih dnevnih dijagrama i taj dnevni dijagram se uzima kao prognoza za taj dan.

Mana ovakvog pristupa je upravo u tome što vrlo često možemo imati nekoliko kandidata za dnevni dijagram, koji su nastali u različitim vremenskim periodima. Imajući u vidu da se potrošnja menja i da se iz godine u godinu beleži rast potrošnje, jasno je da se osim temperature, u obzir mora uzeti i vremenska bliskost ostvarenog dijagrama. Zbog toga su autori izvršili modifikaciju KNN algoritma koji osim temperature uzima u obzir i druge meteorološke parametre (kao što su na primer vlažnost i pritisak), kao i ostvarene dijagrame potrošnje iz bliske prošlosti, odnosno za tri dana unazad, vodeći računa o tipu dana (da li je u pitanju radni dan, vikend i praznik), imajući u vidu da se potrošnje razlikuju radnim i neradnim danima. Stoga se klasifikacija vršila uzimajući u obzir tip dana, a shodno tabeli 1.

TABELA 1. NAČIN ODABIRANJA DANA ZA POTROŠNJU IZ BLISKE PROŠLOSTI

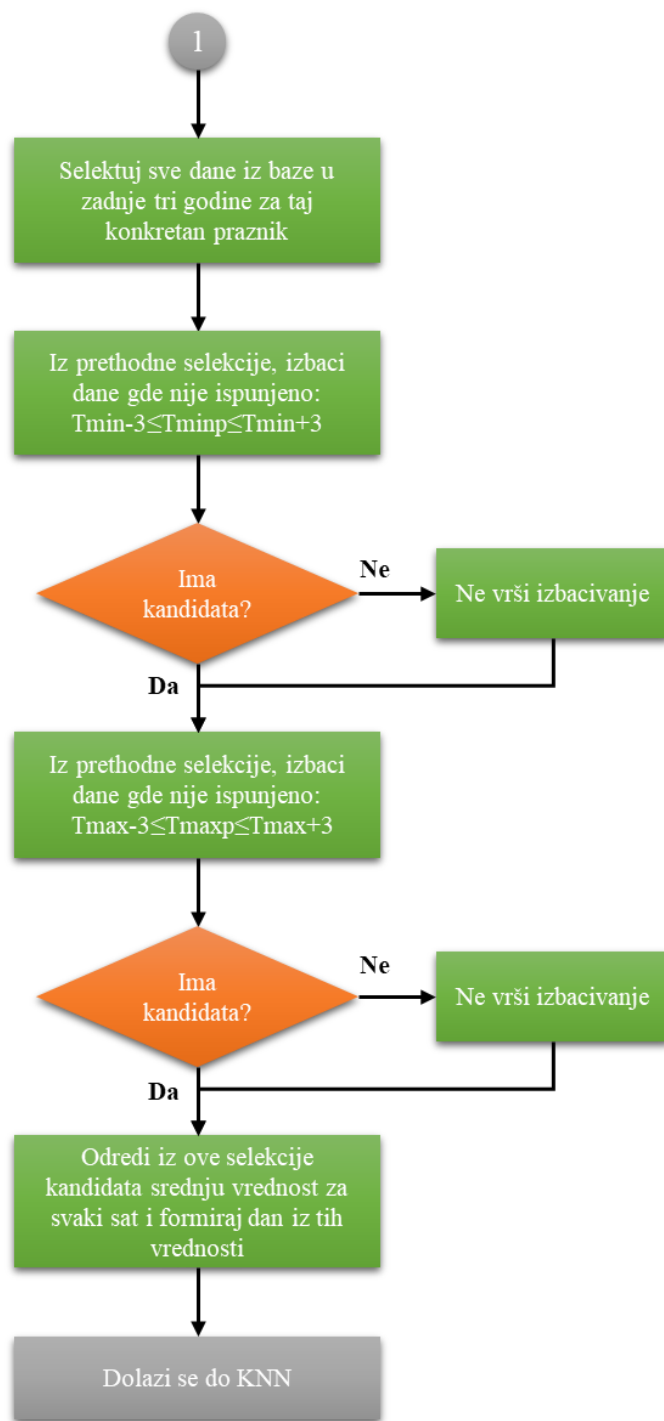
Dan prognoze	Prethodna tri dana koja se uzimaju u obzir
Ponedeljak	Sreda, Četvrtak, Petak
Utorak	Četvrtak, Petak, Ponedeljak
Sreda	Petak, Ponedeljak, Utorak
Četvrtak	Ponedeljak, Utorak, Sreda
Petak	Utorak, Sreda, Četvrtak
Subota	Nedelja pre 14 dana, Subota i Nedelja pre 7 dana
Nedelja	Subota i Nedelja pre 7 dana, Subota

Detaljan opis algoritma sa svojim dijagramom toka je dat na slikama 1 i 2. Algoritam je potom realizovan u obliku računarskog programa urađenog na JAVA programskom jeziku.



* Unose se dani prema tabeli 1

SLIKA 1. ALGORITAM MODIFIKOVANOG KNN ZA PROGNOZU DNEVNOG DIJAGRAMA SATNE POTROŠNJE (UJEDNO I SNAGE) ZA DAN ZA KOJI SE VRŠI PROGNOZA



SLIKA 2. ALGORITAM MODIFIKOVANOG KNN ZA PROGNOZU DNEVNOG DIJAGRAMA SATNE POTROŠNJE (UJEDNO I SNAGE) ZA DAN ZA KOJI SE VRŠI PROGNOZA (NASTAVAK)

MODEL AUTOREGRESIVNIH INTEGRISANIH POKRETNIH PROSEKA (ARIMA)

U analizi vremenskih serija, kao što je slučaj i sa dnevnim dijagramima satne potrošnje, posmatramo serije koje su uređene u odnosu na vreme, pa samim tim posmatramo skup opservacija koje su međusobno zavisne. Vremenska serija je stacionarna ukoliko njena svojstva ostaju nepromenjena tokom vremena, pa samim tim ispoljava određeni obrazac koji nam služi kao osnova za predviđanje budućih vrednosti. U praksi većina vremenskih serija ne ispunjava uslove stacionarnosti, a to važi i za vremenske serije satnih potrošnji. Model autoregresivnih integrisanih pokretnih proseka (ARIMA), koji je predložen od Box i Jenkins (17), predstavlja

modifikaciju ARMA modela i upravo je namenjen modelovanju vremenskih serija koje se ne mogu smatrati stacionarnim. Ipak mnoge vremenske serije imaju neke trendove, pa se stoga njihovim diferenciranjem mogu transformisati u stacionarne. ARIMA model predstavlja vremensku seriju koja koristi diferenciranje i ARMA model koji uključuje proces pokretnih proseka (MA) autoregresivni (AR) model.

Ako posmatramo vremensku seriju podataka X_t gde je t celobrojni indeks a X_t su realni brojevi, ARMA (p', q) model je dat sa:

$$X_t - \alpha_1 X_{t-1} - \dots - \alpha_{p'} X_{t-p'} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q},$$

odnosno sa:

$$\left(1 - \sum_{i=1}^{p'} \alpha_i L^i\right) X_t = \left(1 + \sum_{i=1}^q \theta_i L^i\right) \varepsilon_t$$

gde je L operator vremenskog kašnjenja, α_i su parametri autoregresivnog dela modela, θ_i su parametri dela modela koji opisuje pokretne proseke i ε_t su greške. Pretpostavlja se da su ε_t nezavisne (nekorelisane), ravnomerno distribuirane promenljive u skladu sa normalnom distribucijom i nultom prosečnom vrednošću („beli šum“).

Ako se pretpostavi da polinom $\left(1 - \sum_{i=1}^{p'} \alpha_i L^i\right)$ ima višestruke nule u jedinici, odnosno jedinične korene, reda d , onda se može napisati:

$$\left(1 - \sum_{i=1}^{p'} \alpha_i L^i\right) = \left(1 - \sum_{i=1}^{p'-d} \phi_i L^i\right) (1 - L)^d$$

ARIMA (p, d, q) model se tada, koristeći se prethodnim izrazom uz $p=p'-d$ može prikazati kao:

$$\left(1 - \sum_{i=1}^p \phi_i L^i\right) (1 - L)^d X_t = \left(1 + \sum_{i=1}^q \theta_i L^i\right) \varepsilon_t$$

ARIMA model je veoma značajan u analizi vremenskih serija, naročito u prognozi u slučaju nestacionarnih vremenskih serija. Box i Jenkins su predložili sledeće korake pri analizi i predikciji:

1. Testirati stabilnost izlazne serije iz modela (odziva modela). Ukoliko uslov stacionarnosti nije ispunjen, nestacionarnost se može ukloniti diferenciranjem;
2. Izračunati statističke parametre koji opisuju karakteristike odziva, na primer ACF (autokorelaciona funkcija) i PACF (parcijalna autokorelaciona funkcija), na osnovu kojih se mogu odrediti parametri p i q ;
3. Estimirati nepoznate parametre modela i testirati veličine serije reziduala, i
4. Izvršiti analizu modela radi njegove verifikacije i utvrđivanja da li model korespondira sa karakteristikama podataka.

Imajući u vidu prethodno, pri izboru modela, vodi se računa o sledećem:

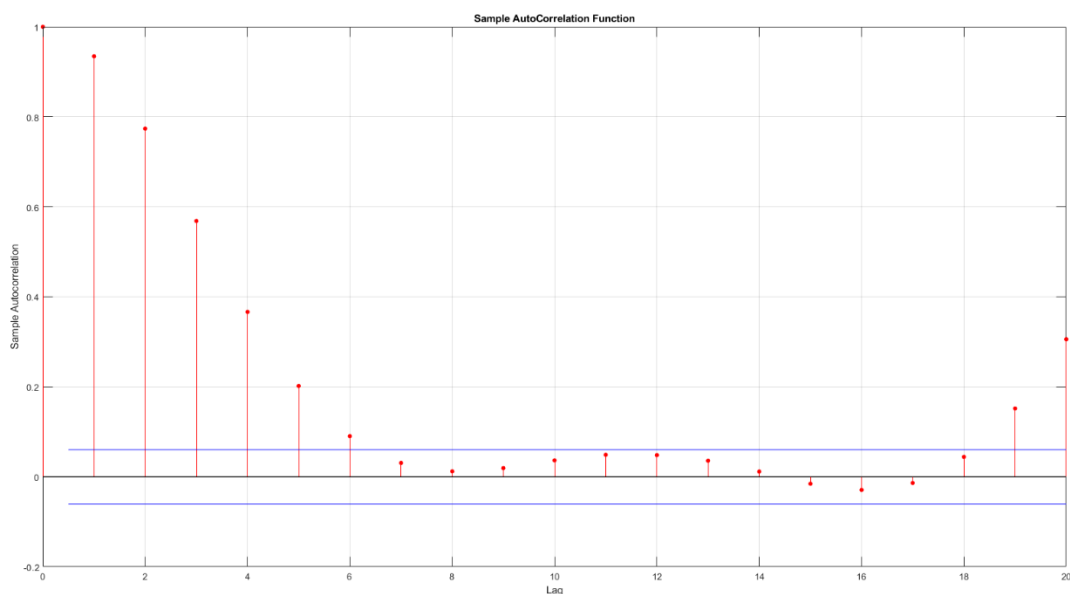
1. ekonomičnost – opisati pojavu što jednostavnijim modelom, koji će istaći suštinsku karakteristiku izučavane pojave;
2. identifikabilnost – bez identifikacije modela postoje bar dva skupa vrednosti koeficijenata koji su u saglasnosti sa podacima;
3. konzistentnost sa podacima i teorijom – testovima se utvrđuje adekvatnost modela, tj. njegova usaglašenost sa podacima i apriornom znanju. Jedan od uslova koji model treba da zadovoljava da bi se smatrao dobrim modelom jeste da se dobro prilagođava podacima. Pored toga od modela se očekuje da ima male rezidualne sa odlikama potpuno slučajnog procesa;
4. prihvatljivost podataka – model ne sme da predviđa vrednosti koje ne zadovoljavaju neka ograničenja koja su prirodna za posmatranu pojavu (npr. model ne sme davati negativne vrednosti za veličine koje ne mogu biti negativne);
5. uspešnost prognoziranja – kriterijum uspešnosti prognoziranja proverava se tako što se koriste opservacije van uzorka za ocenjivanje u cilju provere stepena preciznosti prognoze modela. Takođe, ako jedan od modela ima manju srednje kvadratnu grešku prognoze, a sve ostale karakteristike su jednake, tada se taj model kategoriše kao prihvatljiviji za korišćenje, i

- obuhvatnost – model treba ne samo da objasni, odnosno opiše podatke, već i da objasni uspeh ili propašaj konkurentskog modela u objašnjenju istih podataka.

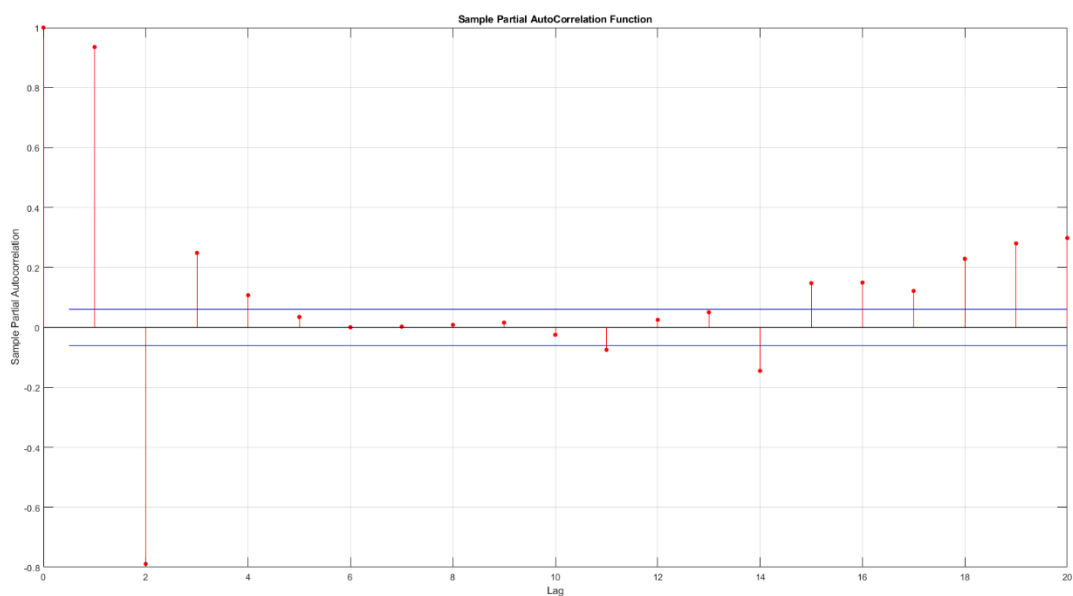
Razvoj i testiranje modela je rađeno sa alatom „EconometricModeler“, a konačni model je potom urađen kao samostalna aplikacija u JAVA programskom jeziku.

IZBOR ARIMA MODELA

Kao što je već rečeno u prethodnom delu, potrebno je izvršiti izbor parametara: p , d i q , $ARIMA(p, d, q)$ modela. Stoga je primenom ADF testa (Augmented Dickey-Fuller unit root test) (18) testirana stacionarnost vremenske serije i utvrđeno je da se nestacionarnost može ukloniti diferenciranjem prvog reda. Zbog toga je za vrednost parametra d izabrana vrednost 1 ($d=1$). Potom su izračunati statistički parametri koji opisuju karakteristike odziva, pa su koristeći se Akaike informacionim kriterijumom (AIC) (19) upotrebene ACF (autokorelaciona funkcija) i PACF (parcijalna autokorelaciona funkcija), da bi se odredili parametri p i q modela, što je prikazano slikama 3 i 4:



SLIKA 3. DIJAGRAM REZULTATA AUTOKORELACIONE FUNKCIJE ACF



SLIKA 4. DIJAGRAM REZULTATA PARCIJALNE AUTOKORELACIONE FUNKCIJE PACF

Vrednosti ACF od prvog do šestog reda su veće od intervala pouzdanosti, pa je izbor $q=6$ prihvatljiv za MA deo ARIMA modela (slika 3). Saglasno tome, na slici 4 se vidi da su najizraženije vrednosti od 1. do 4. reda, pa je stoga za parametar p u AR delu modela, izabrana vrednost 4, $p=4$, tako da je za modelovanje ove vremenske serije izabran model *ARIMA (4, 1, 6)*.

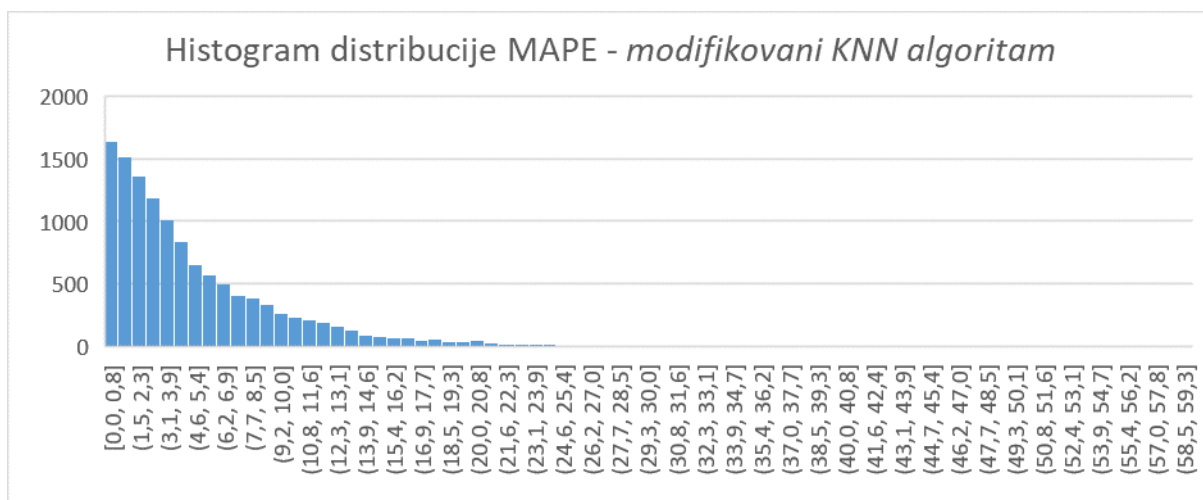
REZULTATI PRIMENE MODIFIKOVANOG KNN ALGORITMA I ARIMA ALGORITMA NA 48 ČASOVNU PREDIKCIJU SATNIH VREDNOSTI POTROŠNJE

Ocene kvaliteta prognoze oba primenjena algoritma su izvršena upoređivanjem maksimalnih (Max APE) i minimalnih (Min APE) apsolutnih procentualnih vrednosti odstupanja (greške) prognoziranih od ostvarenih vrednosti na satnom nivou, potom medijane apsolutnih procentualnih grešaka (MdAPE – Median Absolute Percentage Error), kao i srednje apsolutne procentualne greške (MAPE – Mean Absolut Percentage Error). Rezultati su dati sledećom tabelom:

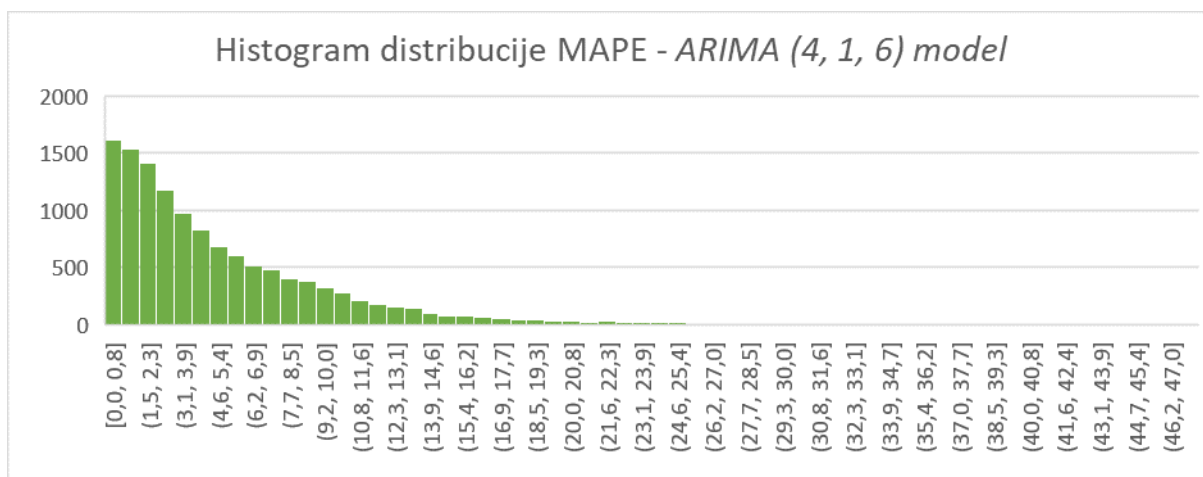
TABELA 2. PROCENTUALNE GREŠKE IZMEĐU PROGNOZIRANIH I OSTVARENIH SATNIH REDNOSTI POTROŠNJE

Vrsta algoritma	Max. APE (%)	Min. APE (%)	MdAPE (%)	MAPE (%)
Modif. KNN	59,20114	0,000000	3,467574	5,074155
ARIMA (4,1,6)	47,61356	0,000318	3,499393	5,071455

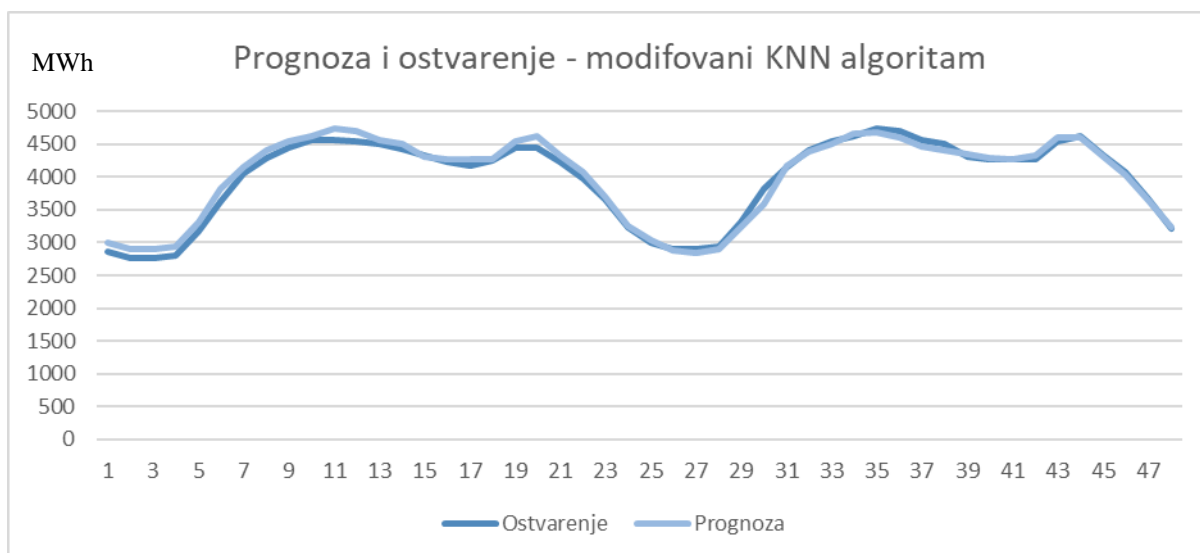
Histogrami distribucije (frekvencije) MAPE za svaki od primenjenih algoritama za prognozu su dati slikama 5 i 6, dok su dijagrami prognoziranih i ostvarenih vrednosti u 48 časovnom periodu, za period od 01.08.2018. – 03.08.2018. godine, dati slikama 7 i 8.



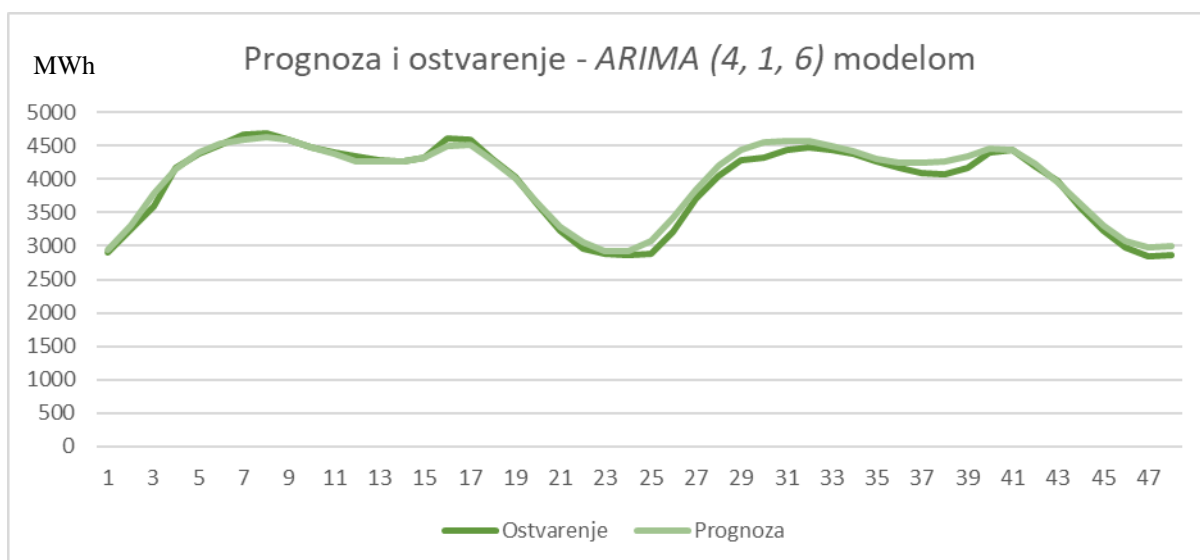
SLIKA 5. HISTORAM DISTRIBUCIJE MAPE ZA MODIFIKOVANI KNN ALGORITAM PROGNOZE



SLIKA 6. HISTORAM DISTRIBUCIJE MAPE ZA PROGNOZU ARIMA (4, 1, 6) MODELOM



SLIKA 7. PROGNOZA I OSTVARENJE SATNE POTROŠNJE ZA PERIOD OD 01.08. DO 03.08.2018 MODIFOVANIM KNN ALGORITMOM



SLIKA 8. PROGNOZA I OSTVARENJE SATNE POTROŠNJE ZA PERIOD OD 01.08. DO 03.08.2018 ARIMA (4, 1, 6) MODELOM

ZAKLJUČAK

Nijedan od primenjenih modela za prognozu na ispitivanom uzorku nije pokazao veću prednost u odnosu na drugi model. Naime kod oba modela, što se tiče grešaka prognoze, kako MAPE, tako i MdAPE, se beleže skoro identične vrednosti. Manja greška ARIMA modela u pogledu maksimalne apsolutne procentualne vrednosti daje blagu prednost prognozi sa ARIMA modelom, dok sa druge strane prednost modifikovanog KNN algoritma se ogleda u većoj transparentnosti i intuitivnosti samog modela.

Autori su ove algoritme isprobavali i na drugim setovima podataka, za okolne elektroenergetske sisteme, i greška prognoze se nije značajno menjala i na tim setovima podataka. Da bi se smanjile greške u prognozi, čija bi prihvatljiva vrednost po ocenu autora bila uslovljena sa $MAPE \leq 2\%$, nameće se potreba kombinovanja navedenih algoritama sa drugim algoritmima u hibridnim modelima, tako da će se dalji rad autora u ovoj oblasti zasnivati na kombinaciji ovih alogritama sa neuralnim mrežama kao i razvoju ARIMAX i SARIMAX modela.

LITERATURA

1. Magnano L, Boland J, 2007, "Generation of synthetic sequences of electricity demand: application in South Australia", "Energy";32, 2230–43.
2. Soares LJ, Souza LR, 2006, "Forecasting electricity demand using generalized long memory", "International Journal of Forecasting", 22, 17–28.
3. Ramanathan R, Engle R, Granger CWJ, Vahid-Araghi F, Brace C, 1997, "Short-run forecasts of electricity loads and peaks", "International Journal of Forecasting", 13, 161–74.
4. Thatcher MJ, 2007, "Modelling changes to electricity demand load duration curves as a consequence of predicted climate change for Australia", "Energy", 32, 1647–59.
5. Harvey A, Koopman SJ, 1993, "Forecasting hourly electricity demand using time-varying splines", "Journal of American Statistical Association", 88, 1228–36.
6. Taylor J, 2003, "Short-term electricity demand forecasting using double seasonal exponential smoothing", "Journal of the Operational Research Society", 54, 799–805.
7. Wang JZ, Chi DZ, Wu J, Lu HY, 2011, "Chaotic time series method combined with particle swarm optimization and trend adjustment for electricity demand forecasting", "Journal of Expert System with Application", 38, 8419–29.
8. Bianco V, Manca O, Nardini S, 2009, "Electricity consumption forecasting in Italy using linear regression models", "Energy", 34, 1413–21.
9. Pao HT, 2009, "Forecasting energy consumption in Taiwan using hybrid nonlinear models", "Energy", 34, 1438–46.
10. Metaxiotis K, Kagiannas A, Askounis D, Psarras J, 2003, "Artificial intelligence in short term electrical load forecasting: a state-of-the-art survey for the researcher", "Journal of Energy Conversion and Management", 44, 1525–34.
11. Chang PC, Fan CY, Lin JJ, 2011, "Monthly electricity demand forecasting based on a weighted evolving fuzzy neural network approach", "International Journal of Electric Power Energy Systems", 33, 17–27.
12. Azadeh A, Ghaderi F, Tarvardiyan S, 2007, "Integration of artificial neural network and genetic algorithm to predict electricity energy consumption", "Journal of Applied Mathematics and Computing", 186(2):1731–41
13. Kucukali S, Baris K, 2010, "Turkey's short-term gross annual electricity demand forecast by fuzzy logic approach", "Energy Policy", 38, 2438–45.
14. Dash PK, Liew AC, Rahman S, Ramakrishna G, 1995, "Building a fuzzy expert system for electric load forecasting using a hybrid neural network", "Journal of Expert System with Application", 9, 407–21.
15. Altman, N. S, 1992, "An introduction to kernel and nearest-neighbor nonparametric regression", "The American Statistician", 46, 175–185.
16. Kouiroukidis, Nikolaos, G. Evangelidis, 2011. "The effects of dimensionality curse in high dimensional knn search", Proceedings of 15th Panhellenic Conference on Informatics IEEE, 421-442
17. George E.P. Box, Gwilym M. Jenkins, Gregory C. Reinsel, 2008. "Time Series Analysis – Forecasting and Control (Fourth Edition)", John Wiley & Sons.
18. W. A. Fuller, 1976. "Introduction to Statistical Time Series", John Wiley and Sons.
19. H. Akaike, 1974. "A new look at the statistical model identification", IEEE Transactions on Automatic Control, 19 (6): 716–723