



Ognjen Ivić¹

Originalni naučni rad
DOI: 10.5937/VIK25345I

PROJEKCIJA I ANALIZA POTROŠNJE ELEKTRIČNE ENERGIJE U VODOSNABDEVANJU NASELJA BAČKI VIÑOGRADI KORIŠĆENJEM ISTORIJSKIH PODATAKA I MODELA MAŠINSKOG UČENJA (2022–2024)

Rezime: Rad prikazuje projekciju i analizu potrošnje električne energije u vodosnabdevanju naselja Bački Vinogradi, Subotica (2022–2024). Korišćeni su istorijski podaci o VTA, NTA, VTR, NTR mesečnom obračunu struje i potisnute količine vode (m³). Primenjeni su modeli Linearna i Polinomijalna regresija, K-Nearest Neighbors, Decision Tree, Random Forest, kao i ARIMA modeli vremenskih serija. Regresioni i mašinski modeli omogućavaju predviđanje mesečne potrošnje na osnovu parametra m³ i prethodnih vrednosti reaktivne struje, dok ARIMA model koristi vremenske serije za dinamičku prognozu. Analiza uključuje upoređivanje predviđenih i realnih vrednosti, identifikaciju najpreciznijih modela i vizualizaciju rezultata radi optimizacije potrošnje.

Кljučне речи: potrošnja energije, vodosnabdevanje, mašinsko učenje, regresioni modeli, ARIMA prognoza

PROJECTION AND ANALYSIS OF ELECTRICITY CONSUMPTION IN THE WATER SUPPLY SYSTEM OF BAČKI VIÑOGRADI USING HISTORICAL DATA AND MACHINE LEARNING MODELS (2022–2024)

Abstract : This study presents the projection and analysis of electricity consumption in the water supply system of Bački Vinogradi, Subotica (2022–2024). Historical data were used for VTA, NTA, VTR, NTR, monthly electricity billing and pumped water volumes (m³). For prediction, Linear and Polynomial Regression, K-Nearest Neighbors, Decision Tree, Random Forest, and ARIMA time series models were applied. Regression and machine learning models enable forecasting of monthly consumption based on the m³ parameter and previous reactive power values, while the ARIMA model uses time series for dynamic forecasting. The analysis includes a comparison of predicted and actual values, identification of the most accurate models, and visualization of results aimed at optimizing consumption.

Key Words: Energy consumption, Water supply, Machine learning, Regression models, ARIMA forecasting

1. Uvod

Ovaj rad analizira, predviđa i planira potrošnju parametara procesa proizvodnje vode na primeru manje Fabrike za preradu pitke vode u naselju Bački Vinogradi u Subotici, u periodu

¹ JKP „Vodovod i kanalizacija“ Subotica, ognjen@vodovodsu.rs, ORCID: 0009-0005-3005-512X



2022–2024. Istorijski podaci koji se koriste uključuju potrošnju električne energije cele fabrike, izraženu u kilovat-satima (kWh) za aktivnu struju i kilovolt-amper reaktivne (kVArh) za reaktivnu struju, obračun troškova (u rsd) i količinu potisnute vode (m^3). Aktivna struja (viša/niša tarifa -VTA/NTA) predstavlja energiju koja se stvarno troši za rad celokupne fabrike i pokretanje procesa potiskivanja vode u naselje, dok reaktivna struja (viša/niša tarifa - VTR/NTR) ne vrši koristan rad već je potrebna za održavanje elektromagnetnih polja u motorima i transformatorima. Praćenje oba tipa struje u odnosu na potisnutu količinu vode omogućava analizu efikasnosti fabrike i optimizaciju potrošnje energije. Podaci se kombinuju u jedinstvene vremenske serije, kako bi se omogućila precizna analiza.

Korišćenje istorijskih podataka omogućava modelima mašinskog učenja da iz stvarnih obrazaca potrošnje, identifikuju sezonske varijacije i predvide buduće vrednosti sa većom preciznošću nego jednostavne statističke metode. U radu se primenjuju različiti modeli, linearni i polinomski regresioni modeli, „stabla“ odluke, RandomForest, KNN i ARIMA model za vremenske serije. Ovo omogućava poređenje performansi, identifikaciju najboljih metoda za svaki tip podataka i smanjenje greške.

Algoritam je razvijen i implementiran u okviru programskog jezika Python, gde je moguće vršiti manipulaciju podacima i izvođenje matematičkih operacija nad nizovima podataka. Za izradu, treniranje i evaluaciju modela mašinskog učenja primenjene su različite metode iz „biblioteke“. Za analizu vremenskih serija korišćen je ARIMA model. Modeli se kasnije čuvaju i učitavaju, omogućavajući ponovno korišćenje najboljih modela bez potrebe za ponovnim treniranjem. Ovaj pristup omogućava autoru da kombinuje različite modele, analizira njihove performanse i obezbedi precizna predviđanja potrošnje energije i optimizaciju upravljanja resursima [1, 2, 3, 4].

2. Prikupljanje i priprema podataka

U pripremi podataka, prvobitno se kreiraju posebni skupovi podataka za svaku godinu (2022, 2023, 2024), gde su vrednosti mesečno raspoređene po kolonama: VTA, NTA, VTR, NTR, OBRACUN i potisnuta količina vode (m^3). Za period 2022. godine prisutni su nedostajući podaci, označeni kao *NaN*, koji se naknadno popunjavaju tehnikama poput „backfill“ (popunjavanje vrednostima iz narednih meseci) ili zamenom nulom, kako bi se obezbedila potpuna i stabilna matrica podataka pogodna za treniranje modela. Mesečni podaci se kombinuju u jedinstvenu vremensku seriju koja obuhvata sve tri godine. Svakom mesecu dodeljuje se numerički kod (MesecNum od 1 do 12), što omogućava lakše korišćenje u regresionim i modelima vremenskih serija, jer algoritmi mogu direktno uočiti sezonske obrasce [2].

Za reaktivnu energiju, kreirane su i lag vrednosti prethodnog meseca (VTR_lag1 i NTR_lag1), koje se koriste kod ARIMA i regresionim modelima zavisnim od prošlih vrednosti. Lag vrednosti omogućavaju modelima da predvide buduće potrošnje na osnovu istorijskih obrazaca, čime se poboljšava preciznost sezonskih i vremenskih prognoza. Nakon konsolidacije, podaci se razdvajaju nazad u godišnje skupove (df_2022, df_2023, df_2024), što omogućava i uporedivu analizu po godinama, ali i pripremu ulaznih matrica (X) i izlaznih vektora (Y) za svaki cilj: VTA, NTA, VTR, NTR i OBRACUN. Za modele reaktivne energije (VTR, NTR) ulazne matrice uključuju i prethodnu vrednost (lag kolone), dok je za aktivnu



energiju i obračun dovoljan samo mesec i količina potisnute vode (m^3). Kombinacija više modela omogućava upoređivanje njihove tačnosti, uočavanje najbolje strategije u planiranju i optimizaciji troškova potrošnje u procesu vodosnabdevanja. Sve ove pripreme procedure osiguravaju da podaci budu čisti, kompletni i spremni za preciznu analizu.

3. Analiza i poređenje modela predviđanja potrošnje energije

Modeli predviđaju potrošnju energije i obračun na osnovu mesečnih podataka i potrošnje vode. Najbolji model se bira po najmanjem apsolutnom odstupanju predviđanja od realnih vrednosti.

3.1. Linear Regression (Linearna regresija)

Linearna regresija se koristi za direktne proporcionalnosti između nezavisnih promenljivih (mesec, m^3 vode) i zavisne promenljive (VTA, NTA, VTR, NTR, OBRACUN). Model se opisuje formulom: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \epsilon$, gde su x_i ulazne promenljive, β_i koeficijenti regresije, a ϵ greška. Linearni model je pogodan za stabilne, linearnim obrascima potrošnje, posebno kod aktivne energije i obračuna.

3.2. Polynomial Regression (Polinomska regresija)

Polinomska regresija proširuje linearnu regresiju uključivanjem viših stepena promenljivih, što omogućava kreiranje nelinearnih obrazaca potrošnje: $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_d x^d + \epsilon$, gde d predstavlja stepen polinoma. Ovaj model se koristi kada se potrošnja menja sezonski ili nelinearno u zavisnosti od meseca i volumena potisnute vode.

3.3. Decision Tree (Drvo odlučivanja)

Decision Tree model pravi hijerarhijsku podelu podataka prema vrednostima ulaznih promenljivih i predviđa izlaz u listovima. Algoritam minimizira varijansu u listovima i može uhvatiti kompleksne nelinearne obrasce kada se koristi posebno za reaktivnu energiju (VTR, NTR), gde se odnosi ne manifestuju linearnom formom i prednost je jednostavnost interpretacije i sposobnost hvatanja interakcija između meseca i m^3 vode.

3.4. Random Forest (Slučajna šuma)

Random Forest kombinuje više Decision Tree modela kako bi smanjio varijansu i poboljšao tačnost. Svako stablo trenira se na nasumičnom uzorku podataka i podskupu promenljivih, a konačna vrednost se računa kao prosek predviđanja svih stabala:

$$\hat{y} = \frac{1}{T} \sum_{t=1}^T h_t(x)$$

gde je T broj stabala, a $h_t(x)$ predviđanja unutar stabla t . Ovaj model je robustan i koristi se za sve varijable sa većom složenosti i sezonskim fluktuacijama.

3.5. K-Nearest Neighbors (KNN)

KNN predviđa vrednost zavisne promenljive kao prosečnu vrednost K najbližih suseda u prostoru ulaznih promenljivih:



$$\hat{y} = \frac{1}{K} \sum_{i \in Nk(x)} y_i$$

gde $Nk(x)$ označava skup K najbližih suseda tačke x . KNN se koristi za reaktivnu energiju, gde obrasci zavise od lokalnih istorijskih vrednosti i teško ih je aproksimirati linearnim modelom.

3.6. ARIMA (AutoRegressive Integrated Moving Average)

ARIMA model predviđa buduće vrednosti vremenskih serija na osnovu prethodnih vrednosti (autoregresija), srednjih promena (integracija) i prosečnih grešaka (*moving average*):

$$y_t = c + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q} + \epsilon_t$$

ARIMA se koristi za VTR i NTR sa *lag* vrednostima, kako bi se iskoristili sezonski i trendovski obrasci u potrošnji. Za analizu potrošnje električne energije i obračuna troškova fabrike vode primenjeni su različiti regresioni i modeli vremenskih serija, sa ciljem predviđanja buduće potrošnje i optimizacije troškova.

4. Prednosti i mane različitih modela predviđanja

Linearna regresija koristi se za podatke sa stabilnim, linearnim obrascima, poput aktivne energije i ukupnog obračuna. Prednost je jednostavnost i lakoća interpretacije, dok ne hvata nelinearne ili sezonske obrasce. Polinomska regresija proširuje linearnu regresiju i omogućava modelovanje sezonskih ili nelinearnih obrazaca potrošnje. Prednost je preciznije praćenje krive, a mana je osetljivost na preterano prilagođavanje podacima. Decision Tree modeli se mogu koristiti za reaktivnu energiju (VTR, NTR), hvatajući kompleksne interakcije između meseca i m^3 vode.

Prednost je interpretabilnost, a mana je osetljivost na male promene u podacima. Random Forest kombinuje više stabala radi veće tačnosti i otpornosti na šum, ali je slabije pregledan i sporije se trenira. K-Nearest Neighbors (KNN) predviđa vrednosti na osnovu preseka najbližih suseda, pogodan za lokalne istorijske obrasce. Prednost je fleksibilnost i jednostavnost, a mana spora obrada kod velikih skupova podataka. ARIMA se koristi za vremenske serije, predviđajući buduće vrednosti reaktivne energije i hvata sezonske i trendovske obrasce, ali zahteva stacionarnost podataka i pažljiv izbor parametara.

5. Princip rada modela i algoritma

Kada korisnik unese mesec, količinu vode (m^3) i stvarne vrednosti za svaku stavku (VTA, NTA, VTR, NTR, OBRACUN), program priprema podatke u odgovarajući oblik za svaki model. Za promenljive koje zavise od prethodnog meseca (VTR i NTR) dodaju se i *lag* vrednosti iz prethodnog perioda. Sklearn modeli se uče nad istorijskim podacima, gde ulaz predstavlja kombinaciju broja meseca (1–12) i potrošnje (m^3) za taj mesec. ARIMA modeli posebno uzimaju u obzir prethodne vrednosti ciljne promenljive za predviđanje vremenske serije. Evaluacija, ocena modela, sprovodi se deljenjem istorijskih podataka na trening i test skup. Trening skup služi za učenje modela, dok test skup omogućava procenu tačnosti predviđanja. Za kvantitativnu ocenu koriste se vrednosti greške:



- MAE (Mean Absolute Error) – prosečna apsolutna razlika između predviđenih i stvarnih vrednosti,
- RMSE (Root Mean Squared Error) – koren srednje kvadratne greške,
- MAPE (Mean Absolute Percentage Error) – procentualna greška.

Pored numeričkih, vizualna evaluacija sprovodi se kroz grafički prikaz predviđanja, što omogućava procenu da li model prati trend stvarnih podataka. Za svaki cilj (VTA, NTA, VTR, NTR, OBRACUN) može postojati više modela. Odabir najboljeg modela vrši se na dva načina:

- Na osnovu apsolutne razlike između predviđene i stvarne vrednosti – model sa najmanjom apsolutnom razlikom koristi se za direktnu primenu na zadate ulazne vrednosti.
- Na osnovu ocene-evaluacije (MAE, RMSE, MAPE) – model sa najmanjim vrednostima predstavlja najbolji generalni model za predviđanje budućih vrednosti.

U prvom delu algoritma, svi modeli se prethodno treniraju i čuvaju zasebno, a ARIMA modeli kao nizovi sa istorijskim vrednostima. Korisnik bira ciljnu promenljivu i vidi predviđanja svih modela, pri čemu su najbolji modeli označeni i dostupni za dalje planiranje. Ovo omogućava kako trenutnu ocenu na osnovu stvarnih vrednosti, tako i korišćenje prethodno ocenjenih najboljih modela za automatsko predviđanje planiranja potrošnje u budućim periodima. U drugom delu koda implementiran je i grafički interfejs. Korisnik bira mesec, ciljnu promenljivu i model i unosi vrednosti planirane potrošnje vode za budućih 12 meseci. Planiranje i rezultati predviđanja se prikazuju tekstualno i grafički, omogućavajući vizualnu procenu trenda i odstupanja [3]. Ovaj pristup omogućava jednostavnu ocenu različitih modela i praktičnu primenu u praćenju i planiranju potrošnje željenog parametra, uz mogućnost proširenja ciljanih promenljivih i dodavanja novih modela bez značajnijih izmena koda.

5.1. Dijagram predviđanja

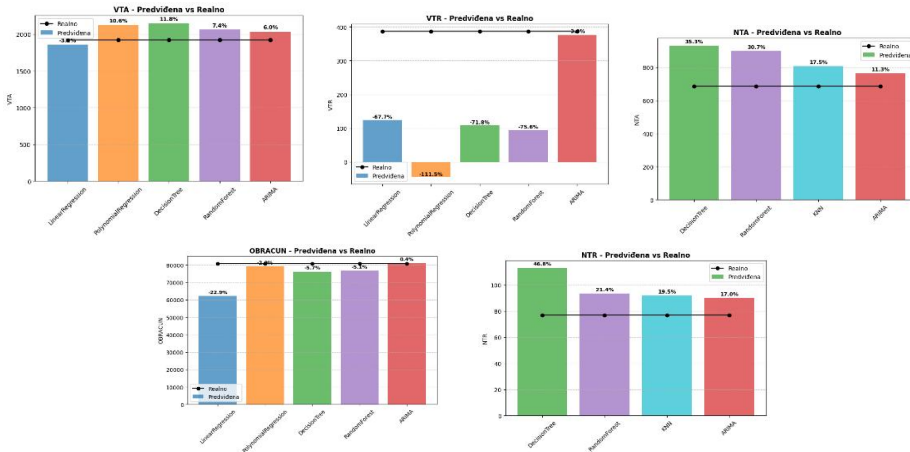
Grafikoni „predviđanja/realno“ (slika 1), prikazuju predviđene vrednosti za različite modele u poređenju sa stvarnim vrednostima za odabrani mesec i količinu m³ (prime meseca januar 2025). Crna linija označava stvarnu vrednost koliko je očitano, dok ostale kolone („bar-ovi“) predstavljaju predviđanja pojedinih modela. Iznad svakog prikazana je procentualna razlika u odnosu na stvarnu vrednost (realnu).

Analiza procentualne apsolutne razlike, pokazuje koliko predviđanje svakog modela odstupa od stvarnih vrednosti u procentima. Za VTA, ARIMA model ima najmanju % razliku (6,03 %), što znači da je najprecizniji, dok Decision Tree i PolynomialRegression odstupaju preko 10 %, dajući manje tačne rezultate.

Kod NTA, ARIMA takođe ima najmanju % razliku (11,32 %), dok modeli poput Decision Tree i RandomForest pokazuju odstupanja preko 20–30 %. Kod VTR i NTR, linearni i složeni regresioni modeli pokazuju velike % razlike (VTR LinearRegression - 67,74 %, PolynomialRegression -111 %), dok ARIMA model odstupa minimalno (-2,97 % za VTR, 16,98 % za NTR), što ukazuje na njegovu sposobnost u predviđanju promenljivih sa izraženim trendovima i sezonalnošću.

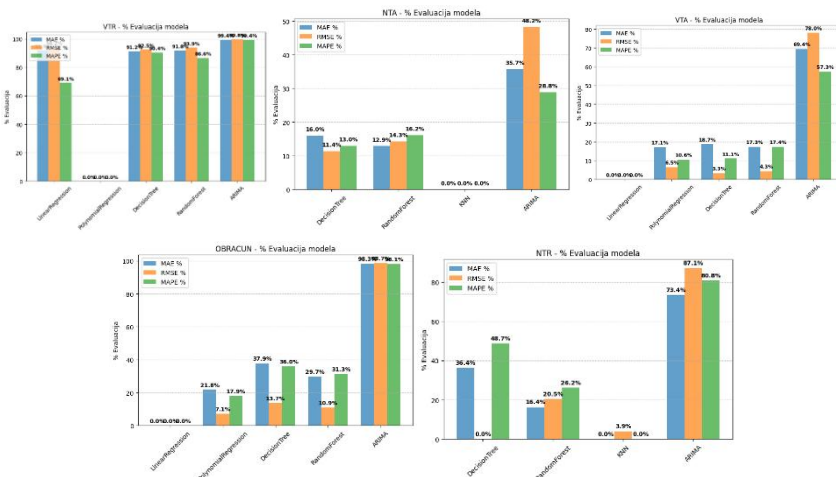


Za OBRACUN, ARIMA model odstupa samo 0,37 %, dok LinearRegression ima čak - 22,89 %, što jasno pokazuje da % razlika omogućava brzo vizuelno i numeričko poređenje preciznosti modela — manja % razlika znači bolje predviđanje [4].



Slika 1. Dijagram predviđanja i realnih vrednosti na primeru mesec januar 2025.
Figure 1. Chart of predicted and actual values for the month of January 2025.

Ocena modela predviđanja (slika 2), rađena je pomoću MAE , RMSE i MAPE. Za VTA, Linear Regression ima najniže greške (MAE 0.0%, RMSE 0.0%, MAPE 0.0%), što pokazuje da su njegova predviđanja najbliže realnim vrednostima. Polynomial Regression i Decision Tree imaju veće greške (MAE 17–19%, RMSE 3–6%, MAPE 10–11%), dok ARIMA ima relativno visoke procenete kod MAE i RMSE (69–78%), ali nizak MAPE (57.3%).



Slika 2. Dijagram rezultata evaluacije modela za mesec januar 2025.
Figure 2. Chart of model evaluation results for the month of January 2025.



Kod NTA, ARIMA ima MAE 35.7%, RMSE 48.2% i MAPE 28.8%, što je bolje od KNN kod naglih promena vrednosti (MAE 0%, RMSE 0%, MAPE 0% zbog specifične vrednosti), dok Decision Tree i Random Forest pokazuju umerene greške (MAE 12–16%, RMSE 11–14%, MAPE 13–16%). Za VTR, ARIMA je najprecizniji (MAE 99.4%, RMSE 99.8%, MAPE 99.4%), dok Linear Regression i Random Forest pokazuju velike odstupanja (MAE 91–92%, RMSE 94–94%, MAPE 69–86%).

5.2. Uticaj stabilnosti istorijskih podataka na preciznost prognoze

Modeli poput Linear Regression i Random Forest daju preciznija predviđanja i planiranja kod stabilnih istorijskih podataka, dok nagla oscilovanja ili manji setovi podataka zahtevaju pažljiv izbor modela, jer greške mogu značajno varirati (% greške jasno pokazuju koliko je model precizan).

Zaključak je da modeli linearnog tipa i ARIMA generalno bolje predviđaju kada istorijski podaci imaju stabilne trendove i sezonske obrasce, dok složeniji modeli ili stabla odluke mogu biti osetljivija na šum, što rezultira većim % razlikama i većim vrednostima MAE, RMSE i MAPE. U praksi, izbor modela zavisi od karaktera podataka: za stabilne istorijske serije može se preporučiti Linear Regression ili ARIMA, dok za nelinearne i promenljive podatke ensemble modeli mogu biti više korisniji, ali uz pažljivo podešavanje kako bi se izbeglo prekomerno prilagođavanje šumu.

Analizom istorijskih podataka za 2022–2024. godinu, primećuje se jasna razlika u ponašanju pojedinih serija podataka. Stabilne serije: VTA, NTA i OBRACUN pokazuju relativno stabilne mesečne trendove i sezonske obrasce. Ove karakteristike omogućavaju linearnim modelima i ARIMA da daju precizna predviđanja. Kod ovih serija, % razlike između predviđenih i stvarnih vrednosti su male, a ocena matrica (MAE, RMSE, MAPE) su niske, što potvrđuje visoku preciznost modela. VTR i NTR imaju velike skokove i nagle skokove u određenim mesecima (npr. VTR u aprilu 2024. je 1446, dok je prethodnog meseca samo 95).

Za ovakve podatke, linearni modeli često zanemaruju ekstremne vrednosti, što rezultira većim % razlikama. Sa druge strane, modeli stabala i ensemble modeli (Decision Tree, Random Forest, KNN) mogu bolje pratiti kratkoročne promene, ali su osetljivi na šum, što povremeno povećava MAE, RMSE i MAPE. Korišćenje prethodnih vrednosti (lag) kod VTR i NTR pomaže modelima da prepoznaju kratkoročne trendove, što naročito poboljšava performanse ARIMA i stabala odluke.

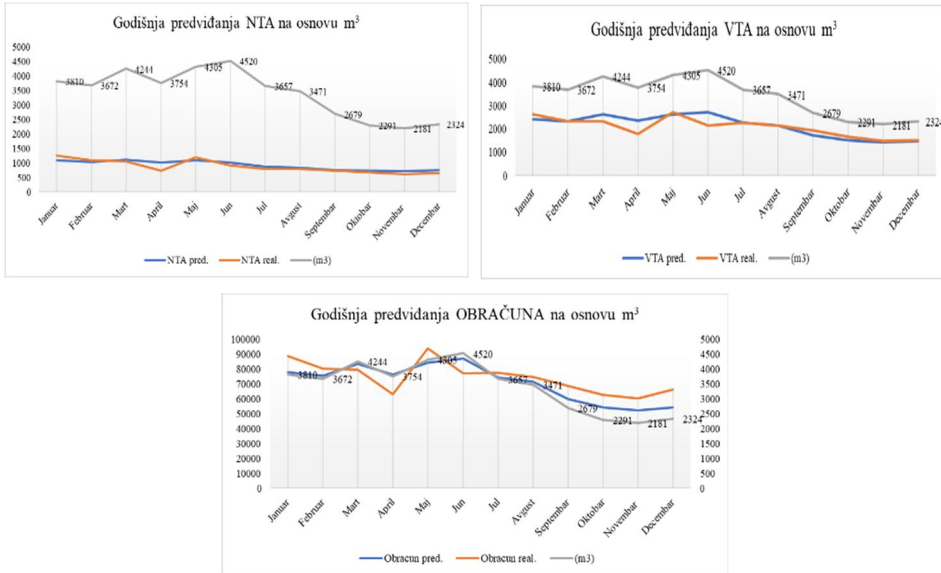
5.3. Projekcija parametara na osnovu m^3 i realnih vrednosti

Poređenjem predviđenih vrednosti za VTA, NTA i OBRACUN sa realnim vrednostima u odnosu na mesečnu potrošnju (m^3) vidi se da modeli generalno prate sezonske trendove. U mesecima sa stabilnom potrošnjom, poput februara, jula i decembra, predviđanja su vrlo precizna, dok u mesecima sa naglim promenama, poput aprila i maja, modeli često precenjuju ili zanemaruju stvarne vrednosti.

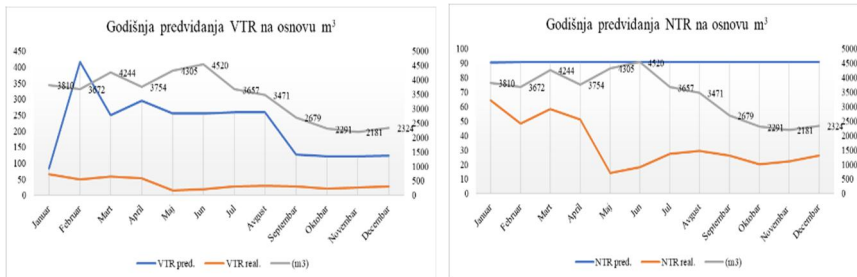
Ovo pokazuje da količina vode (m^3) dobro predviđa utrošak i potrošnju, ali da bi preciznost mogla biti poboljšana uključivanjem dodatnih faktora ili prethodnih vrednosti. Linear Regression i ARIMA najbolje predviđaju stabilne trendove, dok složeniji modeli mogu biti osetljiviji na šum i kratkoročne promene.



Zaključak je da za stabilne istorijske podatke linearnim modelima možemo dobiti pouzdane vrednosti za planiranje, dok promenljivi meseci zahtevaju dodatnu pažnju pri modeliranju.



Slika 3. Dijagram rezultata predviđanja VTA,NTA i Obracuna u odnosu na realnu za 1 godinu
Figure 3. Diagram of the predicted results for VTA, NTA, and OBRACUN compared to the actual values over one year.



Slika 4. Dijagram prikaza rezultata predviđanja VTR,NTR za 1 godinu
Figure 4. Chart displaying the prediction results of VTR and NTR for one year.

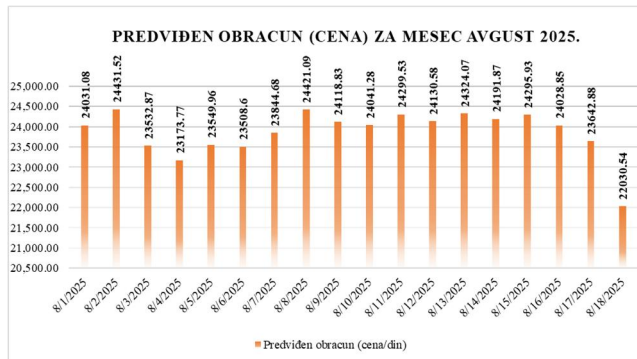
Analiza VTR i NTR pokazuje velika odstupanja između predviđenih i stvarnih vrednosti. Realne vrednosti su male (14–64 m³), dok predviđanja često precenjuju ili ostaju konstantna (npr. NTR ≈ 90 m³), ovo ukazuje da modeli ne prate dinamiku i sezonske promene. Problem je delimično u ulaznim podacima (nula kao početna vrednost) i pristranosti zbog treninga na većim vrednostima. Za preciznije vrednosti, preporučuje se korišćenje stvarnih prethodnih vrednosti, dodatnih feature-a (sezonski indikatori, trend), skaliranje



podataka ili razvoj specifičnih modela za male vrednosti. Planiranje obračuna potrošnje električne energije pokazuje relativno dobro slaganje sa stvarnim vrednostima. Ukupna predviđena planirana godišnja vrednost iznosi 849.993,59 din, a stvarna 891.321,16 din (primer 2023.godina), što predstavlja odstupanje od oko 4,6%. Model uspešno prati sezonski trend potrošnje, pa su mesečne oscilacije u predviđanjima uglavnom u skladu sa realnim promenama. Ipak, u pojedinim mesecima primećuju se veća odstupanja, poput aprila i maja, gde su razlike izraženije, što može ukazivati na uticaj specifičnih faktora (npr. vremenski uslovi, promena u režimu potrošnje) koji nisu obuhvaćeni postojećim modelom. Uprkos tim odstupanjima, ovakva tačnost je dovoljna za potrebe planiranja budžeta i proceni troškova, posebno u situacijama gde je važnija procena trenda nego potpuno precizna mesečna vrednost. Predviđanja za VTA (viša tarifa aktivne struje) i NTA (niža tarifa aktivne struje) pokazuju relativno dobro slaganje sa stvarnim vrednostima, uz mala ukupna godišnja odstupanja – za VTA razlika iznosi oko 699 kWh ($\approx 2,8\%$), a za NTA oko 508 kWh ($\approx 4,9\%$). Model uspešno prati sezonski obrazac potrošnje, pri čemu su predviđene i realne vrednosti bliske u većini meseci.

5.4. Automatizovana predviđanja obračuna (cene) koristeći (m^3) iz SCADA podataka

Ovaj algoritam takođe omogućava praćenje i predviđanje obračuna (cene) potisnute količine vode po danima i mesecu koristeći podatke iz SCADA sistema, slika 5. Iz fajlova se izdvajaju vrednosti protoka vode (Modbus1.Izlazni_protok) i izračunava se zapremina u m^3 po danima. Za svaki dan se koristi sačuvani najbolji željeni model (LinearRegression), kako bi se predvideo obračun za taj dan. Na osnovu ukupne zapremine za ceo mesec računa se i sumarni predviđeni obračun. Algoritam omogućava stalno osvežavanje vrednosti za današnji fajl (18.08.2025.), što omogućava praćenje u realnom vremenu.



Slika 5. Dijagram prikaza dnevnih predviđanja obračuna (cene) za mesec avgust 2025.

Figure 5. Diagram showing daily predicted billing (prices) for August 2025.

6. Zaključak

Kombinovanjem istorijskih podataka i modela mašinskog učenja moguće je predvideti i planirati potrošnju ključnih parametara u procesu prerade vode. Regresioni i vremenski modeli omogućavaju bolje planiranje, optimizaciju troškova i identifikaciju odstupanja u



potrošnji. Za stabilne i sezonske serije (VTA, NTA, OBRACUN) najefikasniji su Linear Regression i ARIMA, jer daju niske greške i precizne rezultate. Za promenljive i naglo promenljive serije (VTR, NTR) bolje funkcionišu ensemble modeli i stabla odluke, jer linearni modeli često ne prate nagle skokove. Pri izboru modela važno je analizirati karakter podataka: stabilne serije zahtevaju jednostavnije modele, a promenljive serije modele koji prate kratkoročne fluktuacije uz kontrolu šuma.

7. Literatura

- [1] Alba E. L., Ribeiro M. H. D. M., Adamczuk G, Trojan F, & Rodrigues E. O. *Water and electricity consumption forecasting at an educational institution using machine learning models with metaheuristic optimization*. <https://arxiv.org/abs/2410.19709>, 2024.
- [2] García-Soto, C. G. Water consumption time series forecasting in urban centers. *Journal of Water Resources Management*, 38(5), 1123–1138. <https://link.springer.com/article/10.1007/s13201-023-02072-4>, 2024.
- [3] Kontopoulos I, Makris A, Tserpes K, & Varvarigou T. *An evaluation of time series forecasting models on water consumption data: A case study of Greece*. <https://arxiv.org/abs/2303.17617>, 2023.
- [4] Ugbehe, P. O. (2025). Electricity demand forecasting methodologies and applications. *Sustainable Energy Research*, 11(2), 145–167. <https://sustainenergyres.-springeropen.com/articles/10.1186/s40807-025-00149-z>, 2025.