Broj rada: R4-10
DOI:10.46793/CIRED2026.R4-10_110

KRATKOROČNA PROGNOZA PROIZVODNJE ELEKTRIČNE ENERGIJE SOLARNE ELEKTRANE

SHORT-TERM FORECAST OF ELECTRICITY PRODUCTION OF A SOLAR POWER PLANT

Stevan Božić, Elektrodistribucija Srbije d.o.o, Srbija

KRATAK SADRŽAJ

U radu će biti prikazani modeli mašinskog učenja za prognozu proizvodnje solarne elektrane. Ponašanje proizvodnje solarne elektrane zavisi od meteo situacije i rezultati prognoze biće tačniji ukoliko prognoza meteo podataka bude najbolje prikazana i najbliža stvarnim vremenskim uslovima rada elektrane. Modeli koji će biti prikazani koristiće istorijske podatke o proizvodnji jedne solarne elektrane i realnim (izmerenim) vrednostima. Metode koje su korišćenje za analizu pripadaju širokoj porodici mašinskog učenja zasnovanog na stablima odlučivanja ali se nalaze u dve različite podgrupe ansambl metoda. Prognoza proizvodnje električne energije iz solarne elektrane je izabrana kao jedan globalni izazov u njihovoj integraciji u distributivni elektroenergetski sistem (ali i širem kontekstu) i u svetlu izazova njihovog priključenja i balansiranja energije u odnosu na potrošnju električne energije. Cilj rada je da upravo rezultati prognoze budu uspešni i najbliži stvarnoj prognozi.

Ključne reči: Balansiranje energije, mašinsko učenje, obnovljivi izvori energije, prognoza proizvodnje solarne elektrane, stabla odlučivanja

ABSTRACT

This paper presents machine learning models for forecasting solar power plant production. The output of a solar power plant is highly dependent on meteorological conditions; therefore, the forecasting results achieve higher accuracy when the weather data is precisely represented and closely aligns with the actual operating conditions of the plant. The models presented utilize historical production data from a specific solar power plant alongside real-world measured values. The methodologies used for the analysis belong to the broad family of decision-tree-based machine learning but are categorized into two distinct subgroups of ensemble methods. Solar power production forecasting is identified as a significant global challenge regarding integration into the electrical distribution system, particularly in the context of grid connection challenges and energy balancing relative to consumption. The primary objective of this research is to ensure successful forecasting results that achieve the highest possible proximity to actual production values.

Key words: Decision tree-based models, electrical distribution system, energy balancing, machine learning, renewable energy sources (RES), solar power forecasting

* Stevan Božić, Elektrodistribucija Srbije d.o.o, Srbija

1. UVOD

Uvođenje negativnih cena na SEEPEX (South East European Power Exchange) predstavlja istorijsku prekretnicu za naše energetsko tržište. Od maja ove godine SEEPEX omogućava da cena struje padne ispod nule, što je direktna posledica rasta broja solarnih elektrana. Negativna cena znači da proizvođač plaća kupcu da preuzme proizvedenu električnu energiju. Iako zvuči nelogično, to se dešava kada je ponuda proizvedene električne energije drastično veća od potražnje. Glavni uzroci ove pojave su vršna proizvodnja solarnih elektrana i niska potrošnja. Vlasnici elektrana koji prodaju struju direktno na tržištu mogu ostati bez prihoda ili čak biti u minusu. Sa druge strane veliki potrošači mogu da prilagode svoju proizvodnju tako da najviše troše kada im se za to plaća. Najveći pobednici u ovoj konstelaciji stvari su skladištari energije jer mogu da prodaju struju kada je ona skuplja. U trenutku pisanja rada cena električne energije u toku perioda od 01.04-07.04.2026. godine (10:00 h-16:00 h) iznosila je 0 €/MWh. Pored ovih činjenica danas prognoza proizvodnje iz solarnih elektrana ključna je za stabilnost elektroenergetskog sistema i optimizaciju potrošnje. Savremeno upravljanje solarnih elektrana nezamislivo je bez precizne prognoze. Sa druge strane proizvodnja zavisi od atmosferskih prilika i njihove promenjive i stohastičke prirode. Cilj ovog rada je da pokaže jedan pristup rešavanja problema kratkoročne prognoze proizvodnje jedne solarne elektrane u Srbiji i dovede do najboljih rezultata uključujući sve faktore koji na njih utiču.

2. UOPŠTENO O KRATKOROČNOJ PROGNOZI PROIZVODNJE

Uobičajeni vremenski okviri za kratkoročne prognoze proizvodnje su:

- Unutrasatna prognoza (Intra-hour): koja se radi za intervale od 15 do 30 minuta. Ovi modeli su korisni za balansiranje mreže u realnom vremenu.
- Satna prognoza (Hour-ahead): predviđanje za narednih 1 do 6 sati. Koristi se za operativno planiranje rada EEO.
- Dan unapred (Day-ahead): prognoza za narednih 24 do 48 sati. Ovo je ključni horizont za učesnike na berzi.

Što je period prognoze kraći prognoza je preciznija. Fokus u radu će biti na prognozi dan unapred (day-ahead) na dve vremenske rezolucije (15- minutnoj i satnoj).

2.1 Metode kratkoročne prognoze proizvodnje električne energije

Metode koje se koriste u prognozi proizvodnje električne energije mogu se podeliti u nekoliko grupa:

- Fizičke metode koje se oslanjaju na numeričke modele predviđanja vremena, satelitske snimke i fizičke karakteristike panela. Ove metode simuliraju kretanje oblaka i atmosferske promene kako bi procenile sunčanu iradijaciju (zračenje).
- Statistčke metode koje koriste istorijske podatke o proizvodnji i pronalaženje linearnih vremenskih serija i kraće periode prognoze.
- Mašinsko učenje (AI) trenutno najnapredniji pristup koji koriste algoritme za modelovanje nelinearnih zavisnosti između meteoroloških parametara i izlazne snage
- Hibridne metode kao kombinacija gore navedenih metoda.

3. OPŠTI PODACI O ELEKTRANI KOJA SE ANALIZIRA

Na slici br. 1 prikazana je geo-lokacija solarne elektrane koja je predmet predviđanja proizvodnje električne energije.



Slika 1 – Lokacija solarne elektrane (izvor: Geosrbija)

Elektrana se nalazi na teritoriji katastarske opštine Beloljin, opština Prokuplje. Instalisana snaga elektrane je 155 kW i nalazi se na zemlji. Predmet ovog rada je prognoza dan-unapred (day-ahead) sa rezolucijom proizvodnje na 15 minuta i 60 minuta. Testni period za model je u vremenskom opsegu 01.01.2026- 14.04.2026. godine i bilo je ukupno 104 dana ove godine.

4. IZBOR METODE I ULAZNIH PODATAKA (PREDIKTORA)

Za prognozu proizvodnje električne energije iz solarne elektrane primarno biće korišćena metoda Random Forest (u daljem tekstu RF) koja spada u tehnike mašinskog učenja i ima široku primenu u prognozi. Sekundarna metoda je LightGBM (u daljem tekstu LGBM). [1].
RF pravi stotine stabala odlučivanja (Decision Trees). Svako stablo daje svoju procenu proizvodnje, a finalni rezultat je prosek svih tih procena. Da bismo napravili dobru prognozu pomoću oba modela, potrebni su nam ulazni podaci i to su podaci o proizvodnji predmetne elektrane sa obračunskog merenja (brojila električne energije). Ovi podaci su "sirovi" i ne mogu se direktno koristiti.
LGBM (Light Gradient Boosting Machine) takođe formira stabla odlučivanja ali na način da svako sledeće stablo ispravlja greške prethodnog (više stotina stabala međusobno „sarađuju").

4.1 Ulazni podaci i prediktori za modele

Podaci o proizvodnji biće ulazni podaci. Za analizu se koriste merenja izlazne aktivne snage na broju električne energije (izražene u kW) odnosno 15-minutni profili iz sistema Elektrodistribucije Srbije d.o.o. Period treniranja modela je 01.01.2025- 31.12.2025, što obezbeđuje pun sezonski ciklus u trening skupu. Testni period iznosi 104 dana i traje od 01.01.2026.-14.04.2026. godine.
U model prognoze uključuju se i meteo podaci. Meteo podaci preuzeti su sa open-source platforme Open-Meteo. [2]. Open-Meteo nudi dva ključna endpoint-a. i to su Histrorical WEATHER API (koji daje stvarne podatke bazirane na osmatranjima) i Historical FORECAST API (koji čuva arhivu onoga što je model predvideo). Open-Meteo daje satne ERA5 reanalitičke podatke i oni predstavljaju reanalizu istorijskih podata o vremenu u kombinaciji sa numeričkim modelima (ovo nisu prognozirane vrednosti). ERA5 daje optimističniju gornju granicu sposobnosti modela i kao takav daće bolju prognozu nego u slučaju korišćenja numeričkih modela prognoze NWP (Numerical Weather Prediction).

Za preciznu prognozu proizvodnje ključno je prikupiti meteo podatke koji direktno utiču na apsorbciju energije u panelima. Najvažniji parametar je solarna iradijacija (zračenje) dok ostali faktori poput temperature i vetra koriguju prognozu. Solarna iradijacija je najvažniji faktor za proizvodnju. Postoje četiri ključna termina koja treba razlikovati:

- Global Horizontal Irradiance (u daljem tekstu: GHI) je ukupno sunčevo zračenje koje pada na horizontalnu površinu zemlje. Ono se sastoji od zbira direktnog i difuznog zračenja.
- Direct Normal Irradiance (u daljem tekstu: DNI) je zračenje koje dolazi od sunca pod pravim uglom i ključno je za solarne trekere i koncentrisanu solarnu energiju.
- Diffuse Horizontal Irradiance (u daljem tekstu: DHI) je zračenje koje se raspršilo u atmosferi (zbog oblaka, vlage ili prašine) i dopire do panela iz svih pravaca a ne direktno od sunca.
- Global Tilted Irradiance (u daljem tekstu: GTI) je najprecizniji pokazatelj solarne energije koja zapravo pada na površinu solarnih panela a uvažavajući nagib panela.

GTI nije jedan izolovan podatak, već zbir tri različita izvora svetlosti koji stižu do nagnutog panela:

1. Direktno zračenje ili DNI je podatak o zračenju koje pada pod pravim uglom na površinu.
2. Difuzno zračenje: predstavlja svetlost rasejanu u atmosferi koja stiže do horizontalne površine. Nagnut panel vidi malo manji deo neba nego horizontalni, pa je ovaj udeo blago modifikovan.
3. Reflektovano zračenje: svetlost koja se odbila od zemlje (trave, betona) ispred panela i pogodila ga. Horizontalna površina ovo nema.

Pomenuti GTI u sebi već sadrži informacije iz GHI, DNI i DHI, preračunate kroz geometrijske modele za zadati nagib panela i ključna je promenjiva za prognozu. GTI je ključni meteo podatak za prognozu jer predstavlja energiju sunca koja pada na panel. .
U analizu će biti uključena i temperatura vazduha (odnosi se na temperaturu merenu na standardnoj visini od 2 m iznad tla) koja utiče na efikasnost panela i oblačnost (predstavlja procenat pokrivenosti neba oblacima) a sve to radi preciznije potrošnje. Zanimljivo će biti upravo koliko poslednja dva faktora utiču na krajnji rezultat prognoze.

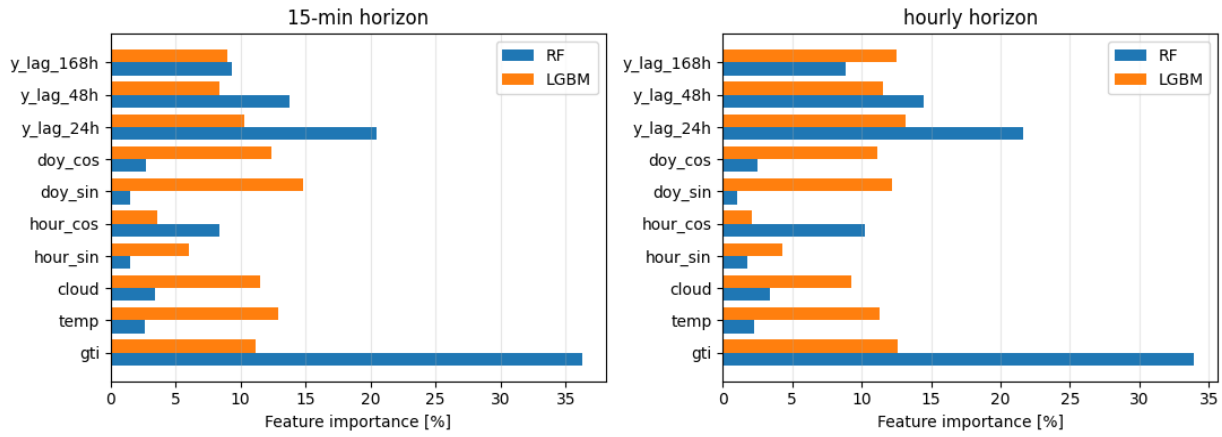
I na kraju formira se konstrukcija predikatora (features) za modele i oni se direktno koriste u RF i LGBM algoritmima i to su:
- Kalendarski podaci: hour_sin i hour_cos (doba dana) i day_sin i day_cos (dan u periodu godine- sezonalnost),
- Lag prediktori (proizvodnja) : y_lag_24h , y_lag_48h i y_lag_168h,
- Meteo podaci: GTI, temp i cloud.

U tabeli I prikazana je važnost prediktora uz obrazloženje i njihov značaj. Dominantni su lag prediktori o prozvodnji jer prate ponašanje potrošnje. Jučerašnji profil proizvodnje najviše liči onome što će se dogoditi i toga dana.

Tabela I – Važnost prediktora (features)

RANG	PREDIKTOR	GRUPA	OBRAZLOŽENJE
1	y_lag_24h	Lag prediktori	Model koristi proizvodnju iz istog sata pre 24 sata; najveći doprinos zbog dnevne perzistentnosti vremenskih obrazaca
2	y_lag_48h	Lag prediktori	Model definiše isti sat dva dana ranije
3	hour_cos	Kalendarski podatak	Model definiše osnovni profil (dan/noć).
4	y_lag_168h	Lag prediktori	Model hvata nedeljnu sezonalnost.
5	gti	Meteo prognoza	Prognoza sunčevog zračenja (iradijacije)
6	doy_cos	Kalendarski podatak	Definiše sezonske varijacije (leto/zima).
7	temp	Meteo prognoza	Temperatura vazduha utiče na efikasnost solarnih panela.
8	Ostali (doy_sin, hour_sin, cloud)	Mešovito	Parametri sa manjim pojedinačnim doprinosom modelu.

Na slici br. 2 prikazano je učešće prediktora (features) u arhitekturi oba metoda. RF posmatra sistem koji je primarno vođen sunčevim zračenjem (iradijacijom) dok LGBM to posmatra malo drugačije. Sezonalnost i lag prediktori proizvodnje LGBM imaju veću važnost. Takođe i važnost temperature i oblačnosti je veća. U globalu meteo podaci uzimaju učešće do najviše 40.5 % (kod RF) i nešto niže (ali ujednačenije) kod LGBM. Podaci o proizvodnji (lag features) idu do 44 % (kod RF) i 33 % (kod LGBM) učešća u predikciji. Vremenski (kalendarski) podaci svoje učešće imaju do 33.5 % (kod LGBM) i upola manji procenat (kod RF) u svojoj važnosti.



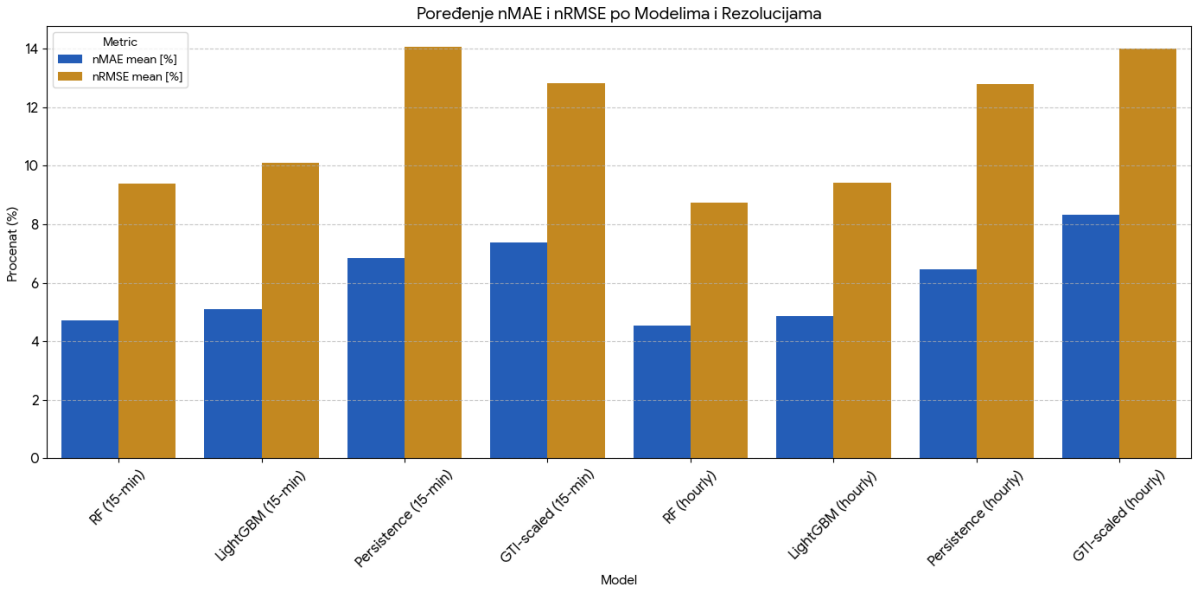
Slika 2 – Procentualno učešće prediktora (features) u prognozi

Dok RF model temperaturu i oblačnost tretira kao sekundarne faktore, LGBM integriše ove parametre ujednačenom težinom što ukazuje na potencijalno bolju adaptilnost modela na nagle promene atmosferskih prilika koje nisu direktno povezane sa sunčevim zračenjem. Promena rezolucije sa 15-minutne na satnu ne menja značajno hijerarhiju bitnih faktora kod oba modela, što potvrđuje stabilnost odabranih skupova.

5. ANALIZA REZULTATA I DISKUSIJA

Za analizu dobijenih rezultata koristiće se poznate metrike greške (videti tabelu br. II). A to su:

- Normalized Mean Absolute Error (u daljem tekstu: nMAE) je metrički podatak koji se koristi za procenu preciznosti modela. Greška u prognozi normalizuje se u odnosu na instalisani kapacitet elektrane a ne u odnosu na trenutnu proizvodnju.
- Normalized Root Mean Square Error (u daljem tekstu: nRMSE) predstavlja kvadrat razlike između prognoziranih i stvarnih vrednosti. Koristiće ovu ometriku model nam opisuje stabilnost prognoze.
- Koeficijent determinacije (u daljem tekstu: R^2) pokazuje koliko dobro model prati varijacije u podacima i prati dinamiku. R^2 pokazuje koji procenat varijabilnosti u stvarnim podacima model uspeva da objasni. [3].



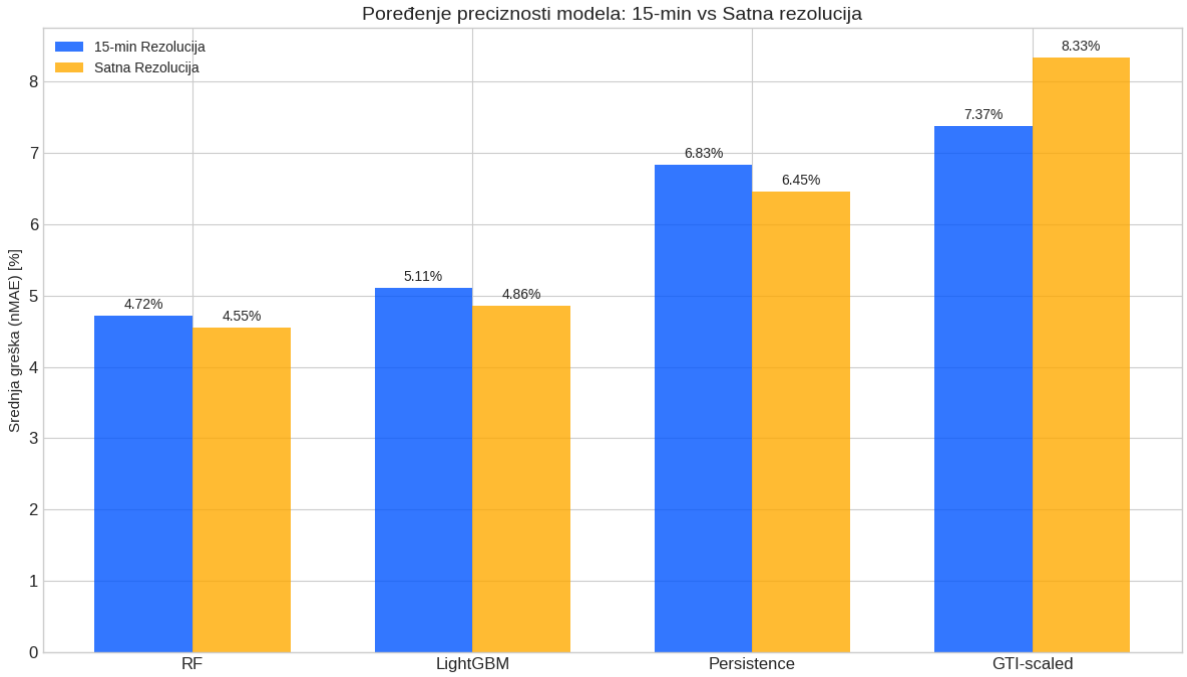
Slika 3 – Poređenje grešaka nMAE i nRMSE za sve modele i rezolucije

Na slici br. 3 mogu se primetiti i rezultati dveju metode (pored RF i LGBM) koje se koriste kao referentne u odnosu na greške prilikom predikcije. To su Persistence odnosno model kod koga se za svaki sat sutrašnjeg dana preuzima izmerena vrednost iz istog sata današnjeg dana. Ovaj model implicitno pretpostavlja da se vremenski uslovi ne menjaju iz dana u dan.. GTI-scaled nam sa druge strane daje teorijski maksimum proizvodnje a da na proizvodnju ne utiču i drugi meteo podaci. Ovaj model koristi samo prognoziranu iradijaciju, bez lag prediktora i kalendarskih obeležja, te služi kao donja referentna granica koju svaka učena metoda mora nadmašiti.

***Tabela II* – Direktno poređenje metrika grešaka**

MODEL	REZOLUCIJA	nMAE (%)	nRMSE (%)	R^2 (mean)
RF	**Satna**	**4.55 %**	**8.75 %**	**-1.78**
RF	**15-min**	**4.72 %**	**9.38 %**	**-1.82**
LGBM	**Satna**	**4.86 %**	**9.41 %**	**-2.27**
Persistence	**Satna**	**6.45 %**	**12.78 %**	**-1.50**
GTI-scaled	**Satna**	**8.33 %**	**14.00 %**	**-5.11**

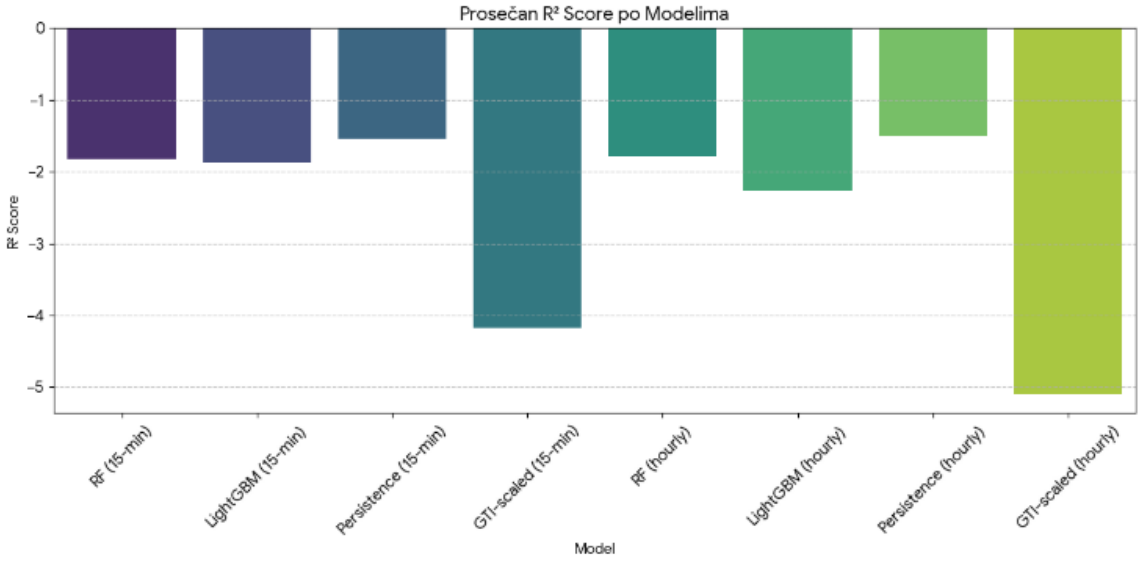
Maksimalne vrednosti za nMAE i nRMSE za satnu rezoluciju za RF su iznosile 13.27 % i 21.79 %. a minimalne vrednosti su bile 1 % i 1.68 %.
Vidimo da su RF i LGBM značajno bolji od fizičkog modela zasnovanog samo na zračenju (GTI-scaled), što znači da algoritmi uspešno uče druge faktore koji utiču na proizvodnju.
Oba modela mašinskog učenja iz prethodnih poglavlja pokazuju niže vrednosti greške na 15-minutnom i satnom novou. RF (na satnom nivou) daje najbolje rezultate prognoze jer su greške metrike najniže (slika br. 4). Na slici br. 4 uočava se razlika u dobijenim rezultatima u zavisnosti od rezolucije prognoze i svih metoda prognoze. RF pokazuje blagu prednost u odnosu na LGBM.



Slika 4 – Poređenje preciznosti modela (15-min i satna rezolucija)

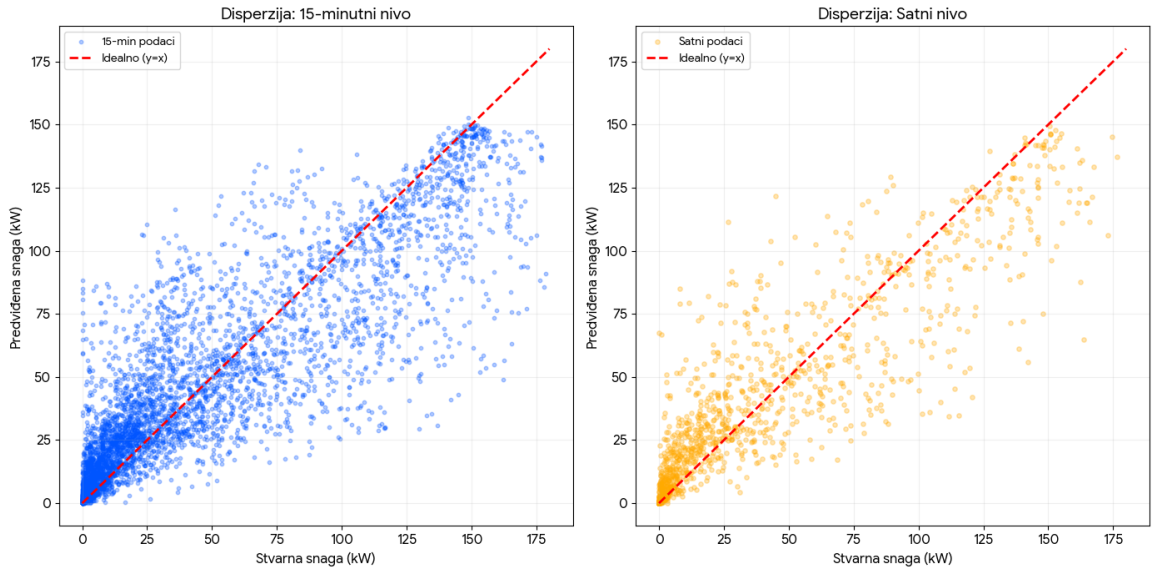
Na slici br. 5 prikazane su vrednosti za R^2 za sve modele i rezolucije predviđanja. To ukazuje da, bez obzira na preciznost modela, postoje dani sa odstupanjima koji kvare ukupnu statistiku. Negativne vrednosti R^2 predstavljaju prosek dnevnih R^2 vrednosti.

Ako bi se posmatrale sve test tačke odjednom (umesto prosečnih dnevnih) vrednosti R^2 postaju pozitivne i kreću se u granicama (0,625-1,847) zavisno od modela.



Slika 5 – Poređenje R^2 za sve modele i rezolucije (15-minutna i satna rezolucija)

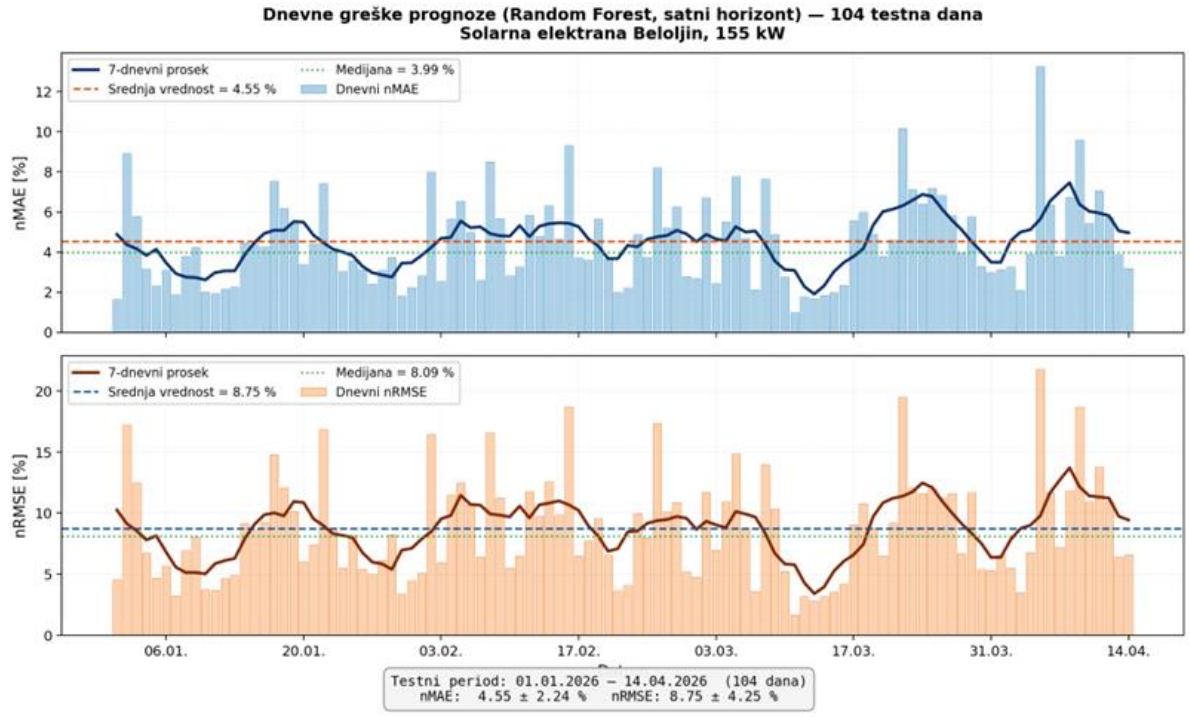
Na slici br. 6 je dijagram disperzije ili Scatter plot koji vizuelno objašnjava preciznost modela na svakom uzorku prognoze. Horizontalna osa prestvalja stvarnu snagu koju je elektrana zaista proizvela (izmerena vrednost). Vertikalna osa je snaga koju je model (RF) predvideo. Crvena isprekidana linija je linija idealne prognoze.



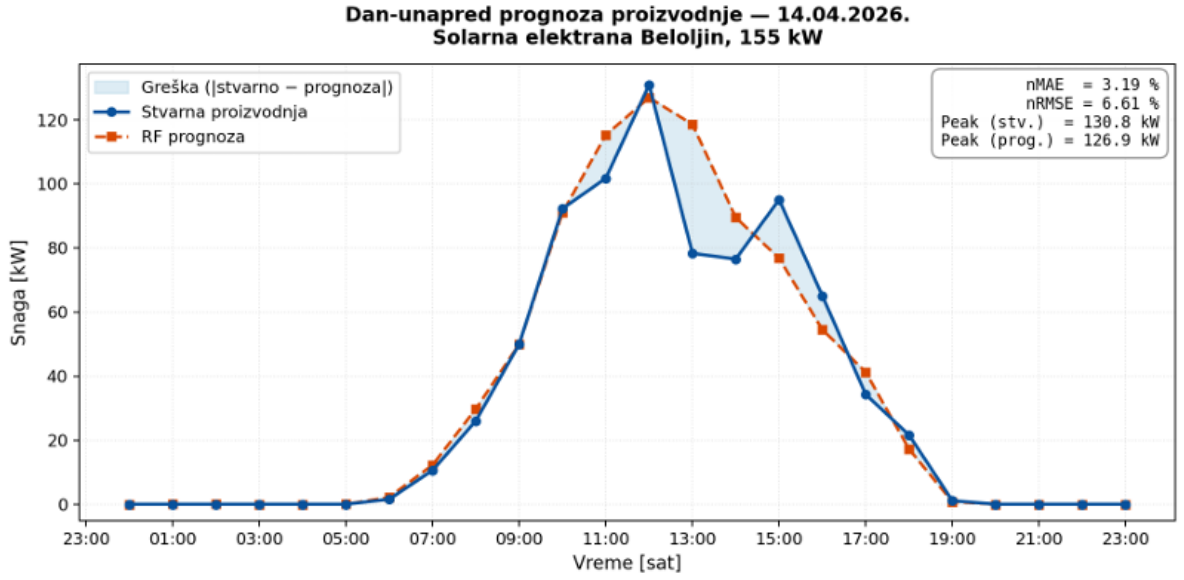
Slika 6 – Disperzija rezultata prognoze (15-minutna i satna rezolucija)

Dnevni nMAE i nRMSE za RF (za satni horizont) kroz 104 testna dana biće prikazani na slici br. 7. Stupčići prikazuju pojedinačne dane, pune linije su 7-dnevni pokretni proseci, a isprekidane/tačkaste linije označavaju ukupnu srednju vrednost. Skokovi greške odgovaraju danima sa jakom ili brzo menjajućom oblačnošću. 7-dnevni pokretni prosek se kroz veći deo perioda zadržava u pojasu 3–6 %, čime se potvrđuje stabilnost RF metode.

Sa druge strane dve jasne epizode nepovoljnog vremena: oko 16.02. i u periodu 22.03.– 05.04. (višednevno nestabilno vreme) su uticale na povećanje greške prognoze.
Na slici br. 8 je dijagram prognoze za 14.04.2026. godine i poslednji dan u nizu za prognozu. nMAE za ovaj konkretni dan je 3.19 % a nRMSE 6.61 %. Osenčeni region između dve linije je apsolutna greška u svakoj tački.

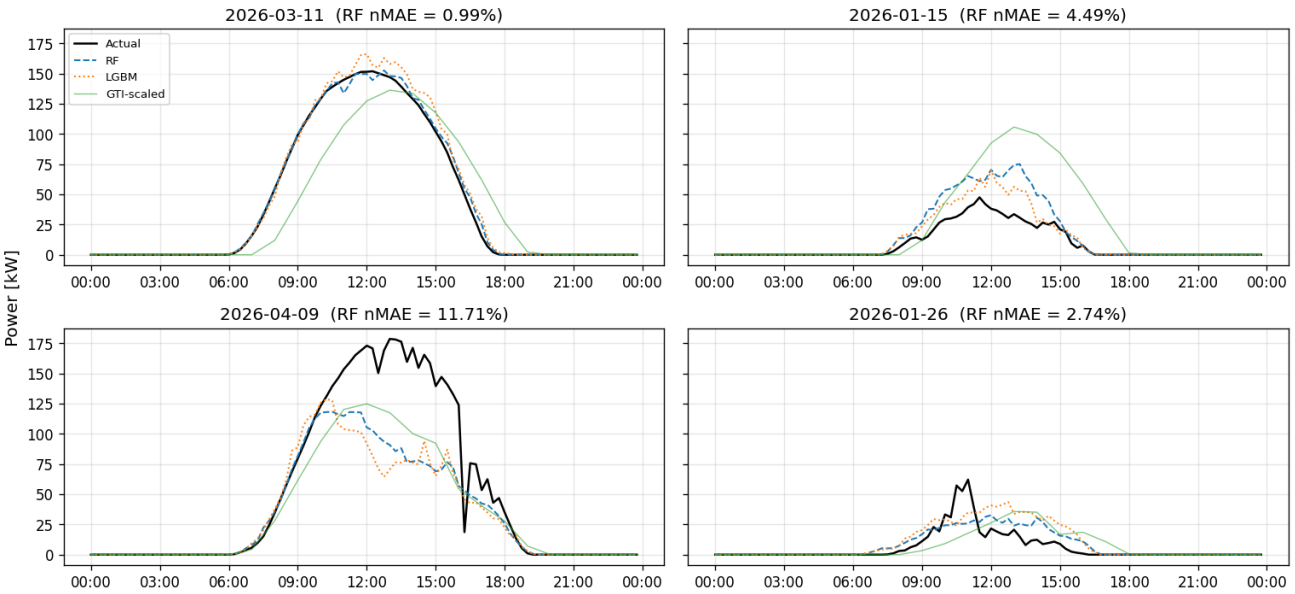


Slika 7– Dijagram dnevnih nMAE i nRMSE za RF (satni horizont) kroz 104 testna dana.



Slika 8– Dijagram prognozirane proizvodnje za 14.04.2026. godine (RF metoda)

I za kraj analize biće prikazana četiri dana prognoze sa karakterističnom proizvodnjom i vrednostima nMAE i to su dani kada su to bili sa najnižom vrednošću, medijanom, najvišom vrednošću i jedan nasumično odabran dan za prognozu (slika br. 9).
Za analizu tipičnog dnevnog profila proizvodnje korišćena je medijana kao mera centralne tendencije, s obzirom na njenu robusnost na ekstremne varijacije uzrokovane naglim promenama oblačnosti, čime se dobija realniji prikaz stabilnog rada elektrane u poređenju sa aritmetičkom sredinom. U kontekstu istraživanja i rada sa vremenskim serijama, medijana za izabrani dan predstavlja statističku vrednost koja se nalazi tačno u sredini niza podataka za taj dan, kada bi sve vrednosti poređao od najmanje do najveće.



Slika 9– Dijagram prognozirane proizvodnje za 4 karakteristična slučaja (RF i LGBM metoda)

6. ZAKLJUČAK

U ovom radu sprovedena je komparativna analiza savremenih modela mašinskog učenja, zasnovanih na ansambl metodama Random Forest i LGBM, u svrhu kratkoročne prognoze proizvodnje solarne elektrane. Analiza rezultata na 15-minutnom i satnom nivou otkrila je kompleksnost modelovanja stohastičke prirode solarnog zračenja, naročito u uslovima promenljive oblačnosti. Rezultati sugerišu da bi za sisteme sa visokom penetracijom obnovljivih izvora energije RF model mogao pružiti visoku preciznost u stabilnim uslovima. LGBM se nameće kao bolje rešenje za modele koji zahtevaju dublje razumevanje sezonskih trendova i kompleksnih nelinearnih veza između meteoroloških varijabli. Oba modela mašinskog učenja su jasno nadmašila referentne beseline modele. Prikazane vrednosti grešaka u radu treba tumačiti uz dva važna ograničenja. Prvo, meteo podaci su preuzeti iz ERA5 reanalize, što predstavlja odličnu istorijsku rekonstrukciju vremena, a ne stvarnu numeričku prognozu. U operativnoj primeni sa realnim NWP (numerička prognoza vremena) prognozama greške bi bile veće za 2-5 %. Druga činjenica je da testni period pokriva samo zimu i proleće pa se ponašanje elektrane u letnjem periodu ne može direktno ekstrapolirati iz rezultata ovog rada. Za operatera distributivnog sistema rezultati pokazuju da se kvalitetna prognoza (dan-unapred) može ostvariti relativno jednostavnim modelima mašinskog učenja bez složene NWP infrastrukture.

LITERATURA

[1] Mladen Apostolović, Miroslav Divčić, Marijan Rančić, Ristov Vasiljević „Izbor odgovarajućih vrsti greške prognoze za adekvatnu procenu uspešnosti operativnog planiranja proizvodnje iz različitih tipova obnovljivih izvora energije", *energija, ekonomija, ekologija,* broj 4/XXVI, 2024,

[2] Open Meteo API, https://open-meteo.com/

[3] Ali Jassim lari, Antonio P. Sanfilippo, Dunia Cachour and Daniel Perez- Astudillo, Using Machine Learning Algorithms to Forecast Solar Energy Power Output, Electronics 2025, MDPI 14,886, https://doi.org/10.3390/electronics14050866