

## SATNA PROCENA GUBITAKA ENERGIJE KORIŠĆENJEM TEHNIKA KLASTEROVANJA

### HOURLY ENERGY LOSSES ESTIMATION BY POLYNOMIAL FITTING AND K-MEANS CLUSTERING

Lazar SLADOJEVIĆ, Elektronski Fakultet u Nišu, Srbija  
Aleksandar JANJIĆ, Elektronski Fakultet u Nišu, Srbija

#### KRATAK SADRŽAJ

Ovaj rad predstavlja novi pristup za procenu i prognozu gubitaka u distributivnoj elektroenergetskoj mreži na osnovu podataka koje su najčešće dostupni operateru mreže. Metoda se može koristiti za određivanje satnih gubitaka tokom proteklih meseci ali i za procenu satnih gubitaka za naredna dvadeset i četiri časa, za potrebe tržišta. Predloženi pristup koristi metodu za minimizaciju sume najmanjih kvadrata kako bi se izračunali koeficijenti potrebni za procenu gubitaka. Pored optimizacije, uvedena je i tehnika mašinskog učenja za grupisanje koeficijenata u nekoliko sezona. Količina podataka koja se koristi u proračunima je veoma velika zbog činjenice da se električna energija meri na svakih petnaestak minuta. Dobijeni rezultati prilično su tačni i mogu se koristiti za klasifikaciju gubitaka kao i za procenu budućih gubitaka.

**Ključne reči:** gubici energije, prognoza, klasterovanje

#### ABSTRACT

This paper represents a new approach for the estimation and forecast of losses in a distribution power grid from data that are normally collected by the grid operator. The method is useful for determining hourly losses for past months but also for estimation of hourly losses during the following twenty-four hours period. The proposed approach utilizes the least squares optimization method in order to calculate the coefficients needed for the estimation of losses. Besides optimization, a machine learning technique is introduced for the clustering of coefficients into several seasons. The amount of data used in calculations is very large due to the fact that electrical energy injected in the distribution grid is measured every fifteen minutes. Therefore, this approach is classified as big data analysis. Obtained results are fairly accurate and can be used for losses classification as well as future losses estimation.

**Keywords:** energy losses, forecast, k-means clustering

Lazar Sladojević, Elektronski Fakultet u Nišu, lazar.sladojevic@elfak.rs  
Aleksandar Janjić, Elektronski Fakultet u Nišu, aleksandar.janjic@elfak.ni.ac.rs

#### UVOD

Obrada i analiza podataka velikom brzinom postaje jedan od najznačajnijih alata u mnogim oblastima inženjerstva. Podaci se prikupljaju svuda, a njihova količina i obim prikupljanja su u konstantnom porastu. Iz tog razloga, poslednjih godina je razvijen čitav niz novih metoda za obradu podataka. Ove metode nisu samo pogodne za izvlačenje korisnih informacija iz čitave mase podataka i njihovu vizuelizaciju, već se mogu iskoristiti i za razvoj prediktivnih modela za najrazličitije aplikacije.

Kao i u drugim oblastima, i u elektroenergetici se uočava porast količine dostupnih podataka. Elektroenergetski sistem se modernizuje brže nego ikada, a savremeni pametni senzori se instaliraju u mnogim tačkama mreže. Ovi senzori prikupljaju podatke o svim električnim veličinama koje su važne za normalno funkcionisanje mreže. Ti podaci se koriste svuda, od regulacije proizvodnih pogona do upravljanja potrošnjom. Dobar pregled različitih primena metoda obrade podataka u upravljanju električnom energijom kao i njihov opis se može naći u (1).

Jedna od značajnijih primena se može naći u upravljanju radom solarnih i vetro-elektrana na osnovu raspoloživih podataka o vremenskim prilikama. Kako vreme ima veliki uticaj na proizvodnju energije iz obnovljivih izvora, vrlo je važno posmatrati odnos između ta dva.

Još jedna interesantna aplikacija se ogleda u detekciji različitih profila potrošača električne energije na osnovu relevantnih merenih veličina. Primer se može naći u (2), gde su autori koristili satna očitavanja potrošnje električne energije stambenih objekata i podatke o spoljnoj temperaturi da bi razgraničili razne profile potrošnje.

Kvarovi u elektroenergetskom sistemu mogu predstavljati veliki problem, naročito kada se dogode na geografski udaljenom delu mreže. Detekcija, identifikacija i lokacija kvarova se takođe može sprovesti na osnovu merenja koja se obavljaju u mnogim tačkama u mreži (3).

Krađa električne energije takođe predstavlja veliki problem. Ona u pojedinim delovima sveta dostiže zabrinjavajuće razmere. Zbog toga su autori iz (4) predložili metodu procene količine ukradene energije na osnovu podataka iz pametnih brojila i metode najmanjih kvadrata. Njihova tema je usko povezana sa ovim radom, s obzirom da se krađa posmatra kao netehnički gubitak energije koji se i ovde evaluira.

Predikcija gubitaka u distributivnoj mreži je privukla posebnu pažnju tokom poslednjih godina zbog deregulirane organizacije tržišta električne energije, gde su operatori distributivnog sistema u obavezi da energiju za pokrivanje gubitaka nabavljaju na otvorenom tržištu (5, 6). Povećana deregulacija tržišta (6, 7) i učešće intermitentnih i obnovljivih izvora energije (8) otežava predikciju gubitaka u vodovima. Na ove gubitke utiče mnoštvo faktora i nelinearnih korelacija koji dodatno komplikuju model.

Dizajn predikcionih modela za gubitke u vodovima je postao prioritet i u prenosnoj mreži pa je i tu predloženo nekoliko rešenja (9, 10, 11). U (9), gubici su modelirani kvadratnim relacijama, u (10) su eksplicitno formulisani preko svih prenosnih jednačina, dok se u (11) uvažavaju i vetro-elektrane i promenljiva opterećenja. Ipak, većina ovih modela je dizajnirana u svrhu raspodele gubitaka po vodovima za potrebe tržišta a ne za dan-unapred predikciju koja je potrebna operatorima prenosnog sistema.

I u distributivnim mrežama se koriste različite metode za proračun gubitaka koje uključuju i heurističke algoritme (12) i neuronske mreže (13). Između opterećenja i gubitaka postoji relacija koja je bliska kvadratnoj i ona je iskorišćena u (14) za razvoj empirijskih formula za proračun gubitaka. Ove formule mogu povezivati gubitke sa faktorom opterećenja (15) ili samim opterećenjem (16, 17). U ovim metodama, uz korišćenje pojednostavljenih modela fidera za proračun gubitaka, koristi se i fitovanje krivih zarad estimacije koeficijenata kvadratnih funkcija.

Iako prethodna istraživanja i studije predstavljaju zadovoljavajuće modele za proračun gubitaka energije, ona ne uvažavaju proces nabavljanja energije na otvorenom tržištu. Stoga ovaj rad ima za cilj doprinos na polju pokrivanja gubitaka estimacijom parametara fitovane krive i njihovom klasifikacijom po sezonama. Podaci koji se koriste u radu se prikupljaju na petnaestominutnom nivou te su i proračuni bazirani na ovim merenjima. Tako se rezultati teorijski mogu iskoristiti za procenu gubitaka za svaki petnaestominutni interval u prošlosti koji podleže dobijenim vrednostima. Ipak, ovo ne ograničava model na ovakvu vremensku granulaciju jer se on sa istim koeficijentima može direktno koristiti i za procenu satnih gubitaka za naredni dan.

Metoda najmanjih kvadrata koja se ovde koristi je vrlo slična regresionim algoritmima, koji se koriste kod mašinskog učenja, ali su ovde uvedena i izvesna ograničenja. Metoda je pogodna za numeričke podatke sa linearnim ili kvadratnim odnosom između merenih (ulaznih) i estimiranih (izlaznih) veličina. Metoda najmanjih kvadrata spada u grupu algoritama „nadgledanog“ učenja, koje će detaljnije biti opisano u poglavlju 3. Ipak, iz eksperimentalnih razloga i u cilju poboljšanja modela je ovde iskorišćen još jedan algoritam mašinskog učenja. To je takozvani klasterizacioni model koji spada u kategoriju „nenadgledanog“ učenja. Iako su ovi algoritmi široko primenjivani u elektroenergetici, koliko je autorima poznato, ovde se prvi put se koriste za estimaciju gubitaka.

## **OPIS PROBLEMA**

U ovom radu je predložen novi pristup estimaciji gubitaka u distributivnoj mreži na osnovu analize dostupnih podataka. Gubici se sastoje od dve komponente: tehničkih gubitaka (TG) i netehničkih gubitaka (NTG). Predloženi pristup je baziran na analizi podataka za 2017. godinu jer su oni bili dostupni u trenutku pisanja rada. Ovi podaci će biti iskorišćeni za estimaciju parametara prediktivnog modela za proračun budućih gubitaka.

## **Fizička interpretacija**

Tehnički gubici se mogu podeliti na dva činioca. Prvi predstavlja konstantne gubitke i oni se uglavnom sastoje od gubitaka u gvožđu transformatora, ali doprinose i drugi faktori, kao što su gubici usled korona efekta, konstantno priključena merna oprema, struje curenja i dielektrični gubici. Drugi činilac predstavlja promenljive gubitke. Oni se uglavnom javljaju u provodnicima nadzemnih vodova i kablovima, ali tu su i drugi provodni

elementi kao što su kontakti prekidača i sabirnice. Ovi gubici su proporcionalni kvadratu struje ili, ekvivalentno, kvadratu aktivne snage. Netehnički ili komercijalni gubici se javljaju usled neredovnih ili loših očitavanja merne opreme i usled krađe električne energije. Stoga, ovi gubici su proporcionalni aktivnoj snazi potrošnje.

U elektro-distributivnoj mreži Srbije, energija preuzeta iz prenosne mreže se meri na svakih petnaest minuta tokom cele godine. Sa druge strane, energija isporučena potrošačima se meri jednom mesečno. Ukupni gubici predstavljaju razliku između ukupno preuzete i ukupno isporučene energije tokom celog meseca. Ovi podaci se prikupljaju i mogu se iskoristiti za estimaciju budućih gubitaka kao i njihovu klasifikaciju. Jedna od metoda pogodnih za ovu vrstu estimacije je opisana u narednom odeljku.

## Matematički model

Postoji nekoliko načina za modeliranje gubitaka u distributivnoj mreži ali svi oni zahtevaju podatke o energiji preuzetoj od prenosne mreže i distribuiranih izvora energije i o energiji predatoj krajnjim korisnicima. Model koji je ovde izabran predstavlja gubitke u sledećem polinomskom obliku (18, 19):

$$\Delta W_{c,j} = \sum_i (a_j + b_j \cdot P_i + c_j \cdot P_i^2) \cdot \Delta t_i \quad [1]$$

gde su  $\Delta W_{c,j}$  estimirani (proračunati) ukupni gubici za mesec  $j$ ,  $i$  je indeks petnaestominutnog intervala  $\Delta t_i$  u mesecu  $j$ ,  $P_i$  je srednja vrednost ulazne snage za taj interval,  $a_j$  predstavlja stalne gubitke u mesecu  $j$ ,  $b_j$  je koeficijent komercijalnih gubitaka i proporcionalan je ulaznoj snazi  $P_i$  u mesecu  $j$ , i konačno  $c_j$  je koeficijent promenljivih gubitaka za mesec  $j$ , proporcionalan kvadratu snage. Za sada se koeficijenti  $a$  i  $c$  smatraju konstantnima u toku cele godine, dok se koeficijent  $b$  menja svakog meseca. Ova pretpostavka će biti argumentovana u narednom poglavlju. Sa druge strane, mereni gubici, označeni sa  $\Delta W_{m,j}$  su već unapred dostupni kao razlika merene ulazne i merene izlazne energije.

Za proračun koeficijenata  $a$ ,  $b$  i  $c$  korišćena je metoda najmanjih kvadrata, što znači da je minimizovana suma kvadrata razlika između proračunatih i merenih gubitaka. Ukupan broj promenljivih je 36 (po dvanaest za konstantne gubitke – koeficijenti  $a_j$ , komercijalne gubitke za svaki mesec – koeficijenti  $b_j$  i varijabilne gubitke – koeficijenti  $c_j$ ). Kao što je već rečeno, neki od ovih koeficijenata se smatraju konstantnima pa je efektivan broj promenljivih manji. Opšta funkcija cilja koju treba minimizovati se sada može zapisati na sledeći način:

$$\min F(a,b,c) = \sum_{j=1}^{12} (\Delta W_{c,j} - \Delta W_{m,j})^2 = \sum_{j=1}^{12} \left( \sum_i (a_j + b_j \cdot P_i + c_j \cdot P_i^2) \cdot \Delta t_i - \Delta W_{m,j} \right)^2 \quad [2]$$

Vrednosti koeficijenata su ograničene na odgovarajuće opsege:  $a_{min} \leq a_j \leq a_{max}$ ,  $b_{min} \leq b_j \leq b_{max}$ ,  $c_{min} \leq c_j \leq c_{max}$ . Ograničenja za koeficijente  $a$ ,  $b$  i  $c$  moraju biti pravilno odabrana, u skladu sa njihovom fizičkom interpretacijom koja je bliže objašnjena u narednom poglavlju.

## Ograničenja predloženog modela

Jedan od nedostataka predloženog modela je taj što on koristi mesečna očitavanja potrošnje električne energije. To znači da su vrednosti merenih gubitaka podložne greškama usled lošeg ili nepravovremenog očitavanja. Na primer, u nekim ruralnim sredinama, potrošnja se očitava tek jednom u tri meseca. Ova činjenica dovodi do nešto nižih očitavanja u pojedinim mesecima i nešto viših očitavanja u drugim što na kraju dovodi do greške u proračunima nekih parametara.

Bolji rezultati bi se dobili češćim očitavanjima, po mogućstvu istom frekvencijom kao i očitavanja ulazne snage. Ovo bi, pak, zahtevalo veliki broj pametnih brojlara u svakoj tački mreže, što još uvek nije praktično ostvarivo. Ipak, ovakvi merni instrumenti se svakodnevno primenjuju sve više te se u budućnosti očekuje veći broj pouzdanih podataka za analizu.

## METODA REŠAVANJA MATEMATIČKOG MODELA

Kao što je već ranije napomenuto, matematički model dat sa [2] je rešen korišćenjem optimizacionog algoritma metode najmanjih kvadrata. Ovaj algoritam je vrlo sličan onom koji se koristi za rešavanje problema linearne regresije u polju mašinskog učenja. U linearnog regresiji, izlazne vrednosti su linearna kombinacija konstantnih parametara i merenih vrednosti ulaznih veličina (prediktora). Iz [1] se može zaključiti da bi u ovom slučaju izlazne veličine bile mesečna merenja gubitaka,  $\Delta W_{c,j}$ , prediktori bi bili  $\Delta t_i$ ,  $P_i \cdot \Delta t_i$  i  $P_i^2 \cdot \Delta t_i$  dok bi njihovi koeficijenti bili  $a$ ,  $b$  i  $c$ , respektivno. Ipak, kako su na sve parametre primenjena odgovarajuća ograničenja, a

neki od njih su i promenljivi, ovaj problem je preformulisan u optimizacioni problem nelinearnog programiranja, koji je dat sa [2].

Za rešavanje problema su korišćene standardne matematičke metode, u ovom slučaju „interior-point“ algoritam. Ulazni parametri su petnaestominutna očitavanja ulazne energije u distributivni sistem i merene vrednosti gubitaka na mesečnom nivou. Izlazne vrednosti su koeficijenti  $a$ ,  $b$  i  $c$ .

Radi lakšeg programiranja i formulacije problema, optimizaciona platforma otvorenog tipa, Yalmip (20) je iskorišćena. Sintaksa Yalmip-a omogućava vrlo jednostavno i intuitivno definisanje promenljivih, funkcija cilja, ograničenja i drugih opcija. Yalmip je korišćen sa jednim od Matlab-ovih integrisanih „solver“-a za rešavanje jednačina. Sam Yalmip može automatski odabrati jedan od raspoloživih solver-a na osnovu strukture problema ali dozvoljava i korisniku da sam odabere onaj za koji misli da je najbolji. Ovo omogućava da se razne vrste optimizacionih problema definišu na isti način, za razliku od korišćenja svakog solver-a posebno, kada se svaki problem mora definisati u formi specifičnoj za taj konkretan solver.

Konkretan solver koji je korišćen je Matlab-ov solver za nelinearno programiranje pod nazivom *fmincon*. On koristi nekoliko minimizacionih algoritama a ovde je konkretno korišćen „interior point“ algoritam (21). Promenljive koje se koriste u proračunima su označene sa  $a_j$ ,  $b_j$  i  $c_j$ , gde je  $j = 1 \dots 12$ . Funkcija cilja je data sa [2]. Ograničenja su bazirana na realnim podacima i realističnim vrednostima koeficijenata. Ukupni nominalni gubici u gvožđu svih transformatora u distributivnom sistemu Srbije iznose približno 32 MW. Stoga, usvojeno ograničenje parametra  $a$  iznosi  $30 \leq a \leq 40$ . Komercijalni gubici uvek postoje, ali u srpskoj distributivnoj mreži ne prelaze nekih 10 % te su zbog toga uvedena sledeća ograničenja parametra  $b$ :  $0.01 \leq b \leq 0.1$ . Za razliku od ova dva, parametar  $c$  se ne može tako lako ograničiti. S obzirom da se on množi kvadratom struje, njegova vrednost je nesumnjivo vrlo mala. Na bazi prethodnog iskustva autora, usvojena su ograničenja:  $0.00002 \leq c \leq 0.00006$ . Kao što je već pomenuto, u prvoj iteraciji proračuna se parametri  $a$  i  $c$  smatraju konstantnima dok se parametar  $b$  menja svakog meseca. Ovo se ne treba shvatiti kao ograničenje, koje jednostavno predstavlja minimalne i maksimalne realistične vrednosti rezultata rešavanja problema. Kako se broj transformatora i topologija mreže skoro i ne menjaju u toku godine, ima smisla držati koeficijente  $a$  i  $c$  konstantnima. Sa druge strane, na koeficijent  $b$  utiče mnogo spoljnih faktora te se on smatra promenljivim. Sve ove pretpostavke se koriste za prvu, inicijalnu fazu proračuna.

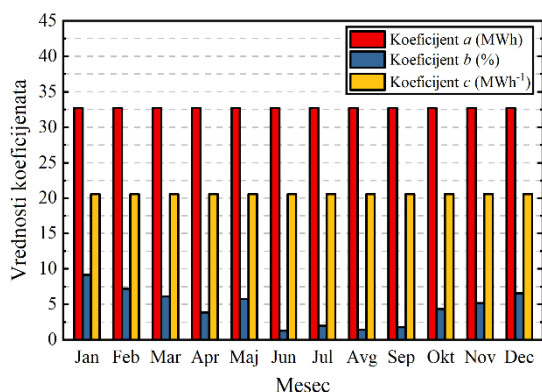
## NUMERIČKI REZULTATI

Dobijeni rezultati su prikazani u tabeli 1 i na slici 1. Tabela 1 sadrži stvarne vrednosti koeficijenata, dok su koeficijenti  $b_j$  i  $c$  na slici 1 skalirani radi bolje vidljivosti pored koeficijenta  $a$ . Nakon skaliranja, koeficijenti  $b_j$  su dati u procentima, dok su koeficijenti  $c$  sada dimenziono jednaki  $W^{-1}$ . Na slici 2 je prikazano poređenje između proračunate i merene vrednosti gubitaka.

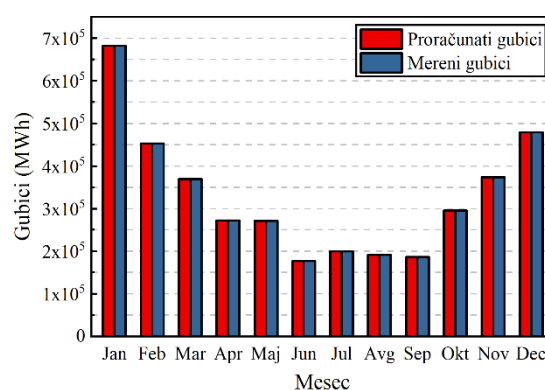
TABELA 1 - IZRAČUNATE VREDNOSTI KOEFICIJENATA

Mesec	Koeficijent $a$ (MW)	Koeficijent $b$ (pu)	Koeficijent $c$ ( $MW^{-1}$ )
Januar	32.739	0.091499	0.000020609
Februar	32.739	0.072052	0.000020609
Mart	32.739	0.061105	0.000020609
April	32.739	0.039002	0.000020609
Maj	32.739	0.057862	0.000020609
Jun	32.739	0.013125	0.000020609
Jul	32.739	0.019671	0.000020609
Avgust	32.739	0.014427	0.000020609
Septembar	32.739	0.018085	0.000020609
Oktobar	32.739	0.043728	0.000020609
Novembar	32.739	0.051645	0.000020609
Decembar	32.739	0.065777	0.000020609

Sa slike 2 se može videti da su proračuni bili uspešni. Izračunate i merene vrednosti gubitaka su jednake, što znači da su koeficijenti dobro estimirani. Govoreći jezikom mašinskog učenja, može se reći da su ulazni podaci, definisani u prethodnom poglavlju u stvari trening podaci za model [1]. Sa slike 2 je očigledno da model ogovara trening podacima. Sada se ovi koeficijenti mogu koristiti za estimaciju gubitaka u budućnosti pod prethodno postavljenim pretpostavkama.



SLIKA 1 - SKALIRANE VREDNOSTI SRAČUNATIH KOEFICIJENATA



SLIKA 2 - POREĐENJE PRORAČUNATIH I MERENIH GUBITAKA

## POBOLJŠANJE MODELA

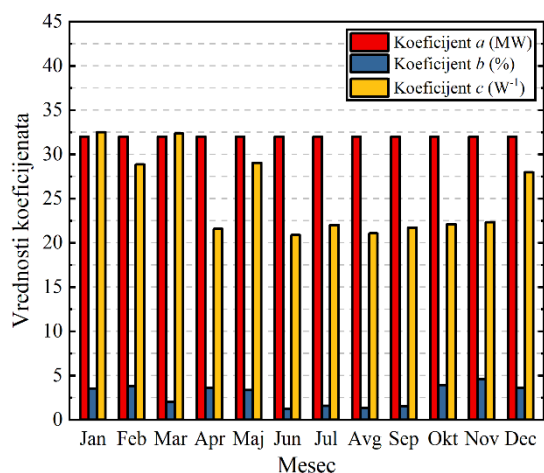
Distributivna preduzeća nabavljaju energiju za pokrivanje gubitaka na otvorenom tržištu električne energije. Da bi se ova procedura pojednostavila, čitav koncept može biti proširen grupisanjem meseci u nekoliko sezona. Prednost ovakvog pristupa se ogleda u samom grupisanju koeficijenata u sezone umesto pojedinačnih koeficijenata za sve sezone i potencijalno bolje tačnosti u pojedinim sezonama. Ipak, tačnost dosta zavisi od kvaliteta prikupljenih podataka. Grupisani koeficijenti su otporniji na razne stohastične varijacije nego pojedinačni. Pored toga, postoji značajna razlika u klimatskim faktorima, kao što su spoljašnja temperatura ili vlažnost vazduha, tokom različitih sezona. Čak ni očitavanje električnih brojlara nije toliko učestalo zimi kao što je leti. Sa druge strane, usvajanje jednog seta koeficijenata za celu godinu bi dovelo do prevelikih grešaka u proračunu budućih gubitaka. Iz svih ovih razloga usvojen je model sa tri sezone, i meseci su klasterovani u zimske, letnje i „prelazne“ sezone.

Klasterizacija je „nenadgledana“ metoda mašinskog učenja, što znači da, za razliku od regresije, nema izlazne podatke sa kojima bi upoređivala ulaz. Umesto toga, ovaj algoritam traži sličnosti između ulaznih podataka. U ovom slučaju, ima dvanaest ulaznih opservacija (merenja), po jedna za svaki mesec, i tri prediktora, koeficijenti  $a$ ,  $b$  i  $c$ . To znači da ima 36 ulaznih vrednosti među kojima klasterizacioni algoritam treba da pronade sličnosti. Za jednostavno skladištenje ovih podata iskorišćena je matrica dimenzija dvanaest puta tri. U ovom konkretnom slučaju, matrica je data u formi tabele 2, gde vrste predstavljaju ulazne opservacije za svaki mesec a kolone predstavljaju prediktore.

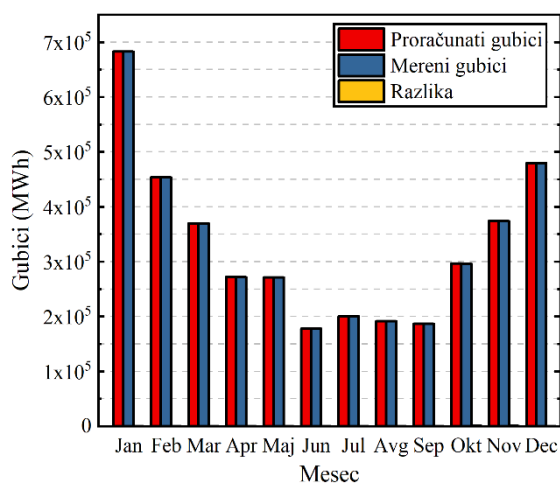
Broj klastera je još jedan od ulaznih parametara klasterizacionog algoritma. Kao što je ranije već naglašeno, odabrani broj klastera je tri, jer je najlogičnije podeliti godinu u tri sezone prema vremenskim uslovima. Ipak, ispitani su i slučajevi sa dva i četiri klastera i rezultati su upoređeni. Kako rezultati, pored broja klastera, zavise i od inicijalne (nasumične) klasterizacije (22), bilo je i nekoliko mogućih rešenja, od kojih je odabrano ono koje ima najviše smisla. U opštem slučaju nema jednog „najboljeg“ načina odabira broja klastera. Umesto toga, potrebno je izvesno poznavanje oblasti iz koje potiču podaci (22).

Još jedan potencijalni problem kod klasterizacije je različiti opseg vrednosti prediktora. Oni mogu da se razlikuju i za nekoliko redova veličine, kao što je slučaj sa tabelom 2. Jasno je da je u ovom slučaju neophodno izvršiti normalizaciju vrednosti da bi se dobili smisljeni rezultati.

Za ovaj konkretan slučaj klasterizacije primenjen je nešto drugačiji pristup nego ranije. Prethodno uvedena ograničenja i dalje važe, ali se sada samo koeficijent  $a$  smatra konstantnim i njegova vrednost je fiksirana na 32 MW. Svi ostali koeficijenti su promenljivi i imaju različite vrednosti svakog meseca. Na ovaj način, koeficijenti su optimizovani tako da su proračunati gubici jednaki ili približno jednaki merenim, slično prethodnom slučaju sa slike 2. Ove vrednosti koeficijenata služe kao ulazni podaci za proces klasterizacije pa će oni na dalje biti nazivani inicijalnim koeficijentima. Dobijene vrednosti ovih koeficijenata su date u tabeli 2, dok slika 3 prikazuje njihove skalirane vrednosti (slično slici 1). Poređenje izračunatih i merenih gubitaka je dato na slici 4. Vrednosti grešaka, tj. razlike sa slike 4 su date i u apsolutnim i u relativnim jedinicama u tabeli 3.



SLIKA 3 - SKALIRANE VREDNOSTI PRORAČUNATIH KOEFICIJENATA NAKON INICIJALNE OPTIMIZACIJE



SLIKA 4 - POREĐENJE IZRAČUNATIH I MERENIH GUBITAKA NAKON INICIJALNE OPTIMIZACIJE

Iz tabele 2 se vidi da su vrednosti koeficijenata  $b_j$  drugačije od onih iz tabele 1. Razlog leži u činjenici da su sada i drugi koeficijenti promenjeni (iako se koeficijent  $a$  promenio vrlo malo). To znači da su i koeficijenti  $b_j$  morali da se promene da bi se dobio najbolji mogući fit ulaznim podacima.

Sa slike 4 i tabele 3 je očigledno da koeficijenti daju skoro savršene vrednosti jer su žuti stubovi na slici 4 skoro nevidljivi za sve mesece.

Koeficijenti  $b$  i  $c$  su sada promenljivi po mesecima, i otkrivanje sličnosti između njih je ključ klasterizacije po sezonama. Ovakav pristup teorijski daje bolje rezultate od pristupa sa dva konstantna parametra jer se u prvom slučaju klasterizacija obavlja na osnovu dva koeficijenta ( $b$  i  $c$ ) dok bi se u drugom slučaju moralo raditi samo sa jednim, koeficijentom  $b$ . Ova teorija je i eksperimentalno dokazana. Na osnovu ovih vrednosti, izvršena je klasterizacija meseci u sezone i rezultati su prikazani u tabeli 4.

TABELA 4 - MESECI KLASTEROVANI U SEZONE

Mesec	Jan	Feb	Mar	Apr	Maj	Jun	Jul	Avg	Sep	Okt	Nov	Dec
Indeks klastera	1	1	1	3	1	2	2	2	2	3	3	1
Sezona	Zima	Zima	Zima	Prelaz	Zima	Leto	Leto	Leto	Leto	Prelaz	Prelaz	Zima

Nazivi sezona u tabeli 4 su jasni – zimska, letnja i prelazna sezona. Iz tabele se vidi da su meseci klasterovani skoro u potpunosti onako kako se očekivalo. Jedini izuzetak je Maj, koji je svrstan u zimske mesece. Ako se pogleda unazad, i prethodni proračuni pokazuju da se Maj ne ponaša prema očekivanom obrascu kao ostali meseci. Na slici 1, svi meseci prate opšti obrazac da koeficijenti  $b$  opadaju tokom leta i rastu tokom zime, u karakterističnom obliku „lakta“. Jedino Maj odstupa od ove karakteristike jer je njegova vrednost izrazito veća od vrednosti okolnih meseci. Takođe je uočljiva i razlika u koeficijentu  $c$  za Maj na slici 3 u odnosu na okolne mesece. Ova vrednost više odgovara zimskim mesecima. Razlog tome može da leži u značajno manjim

TABELA 2 - SRAČUNATE VREDNOSTI KOEFICIJENATA

Mesec	Koeficijent a (MW)	Koeficijent b (pu)	Koeficijent c (MW <sup>-1</sup> )
Jan	32	0.0353882	0.000032503
Feb	32	0.0380628	0.000028874
Mar	32	0.0198292	0.000032351
Apr	32	0.0360867	0.000021577
Maj	32	0.0337267	0.000029041
Jun	32	0.0125835	0.000020878
Jul	32	0.0158188	0.000021984
Avg	32	0.0132383	0.000021086
Sep	32	0.0150446	0.000021727
Okt	32	0.039316	0.000022101
Nov	32	0.0455431	0.000022306
Dec	32	0.0357529	0.00002796

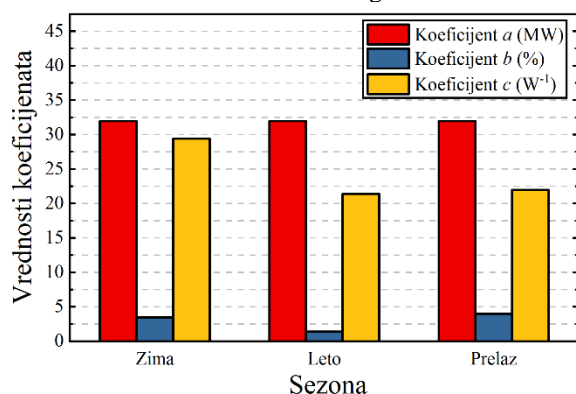
TABELA 3 - APSOLUTNE I RELATIVNE GREŠKE NAKON INICIJALNE OPTIMIZACIJE

Mesec	Apsolutna razlika (MWh)	Relativna razlika (%)
Jan	0.0000	0.0000 %
Feb	0.0000	0.0000 %
Mar	0.0000	0.0000 %
Apr	117.5637	0.0432 %
Maj	0.0000	0.0000 %
Jun	0.0000	0.0000 %
Jul	0.0000	0.0000 %
Avg	0.0003	0.0000 %
Sep	0.0000	0.0000 %
Okt	784.1185	0.2644 %
Nov	407.3877	0.1088 %
Dec	0.0000	0.0000 %

očitanjima u Maju, nepravovremenim prikupljanjem podataka i kreiranjem izveštaja, grešci u statističkoj obradi podataka ili jednostavno većoj stopi krađe električne energije u tom delu godine.

Konačni koeficijenti su izračunati kao centriodi svakog klastera iz tabele 4. Centroid klastera je u stvari vektor srednjih vrednosti svih prediktora za sve opservacije koje pripadaju tom klasteru. Ovi koeficijenti su dati u tabeli 5 a njihove skalirane vrednosti su prikazane na slici 5, dok se novo poređenje merenih gubitaka i gubitaka izračunatih primenom ovih koeficijenata može videti na slici 6. Apsolutne i relativne razlike gubitaka su date u tabeli 6.

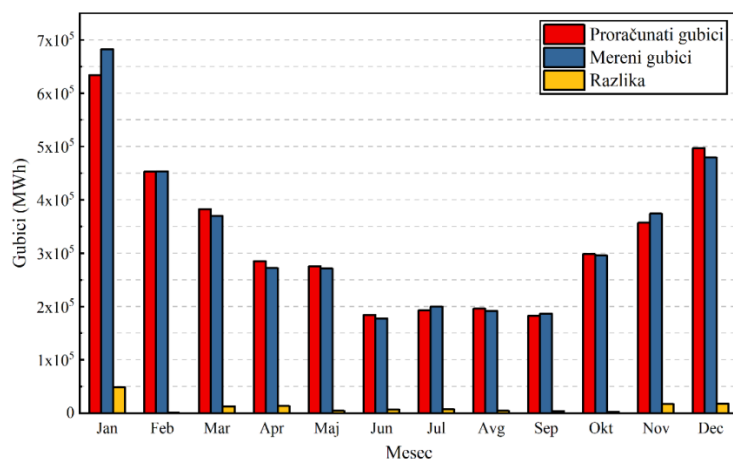
Može se videti da se javljaju male greške u proračunima. Ove greške su posledica činjenice da jedan set koeficijenata ne može savršeno da odgovara svim mesecima, već je cilj da se ukupna greška smanji. Veće vrednosti grešaka se mogu uočiti samo za zimske mesece, i to zbog toga što je i Maj među njima. Ovo se donekle manifestuje i na veće vrednosti koeficijenta  $c$  za zimsku sezonu. Ipak, prosečna vrednost greške je vrlo mala u odnosu na vrednosti samih gubitaka.



SLIKA 5 - GRUPISANI KOEFICIJENTI ZA SVE TRI SEZONE

TABELA 5 - GRUPISANI KOEFICIJENTI ZA SVE TRI SEZONE

Indeks klastera	Sezona	Koeficijent a (MW)	Koeficijent b (pu)	Koeficijent c (MW <sup>-1</sup> )
1	Zima	32	0.0350917	0.00002942712
2	Leto	32	0.0141719	0.00002141825
3	Prelaz	32	0.0403153	0.00002199484



SLIKA 6 - POREĐENJE IZRAČUNATIH I MERENIH VREDNOSTI GUBITAKA POSLE KLASTERIZACIJE

TABELA 6 - APSOLUTNE I RELATIVNE RAZLIKE IZMEĐU IZRAČUNATIH I MERENIH GUBITAKA NAKON KLASTERIZACIJE

Mesec	Apsolutna razlika (MWh)	Relativna razlika (%)
Jan	48709.5891	7.6852 %
Feb	697.9128	0.1541 %
Mar	12719.8945	3.3250 %
Apr	13128.5232	4.6069 %
Maj	4248.3817	1.5425 %
Jun	6624.1044	3.5945 %
Jul	7211.0681	3.7413 %
Avg	4227.7345	2.1589 %
Sep	3686.2140	2.0147 %
Okt	2343.3930	0.7859 %
Nov	16880.7436	4.7272 %
Dec	17489.1562	3.5188 %

## ZAKLJUČAK

U ovom radu je predstavljen novi pristup proračunu gubitaka u elektro-distributivnoj mreži. Prikazane su dve opšte klase gubitaka: tehnički i netehnički. I jedni i drugi su neizbežni, ali je važno znati u kojoj meri oni utiču na ukupne gubitke tj. treba ih klasifikovati. To je učinjeno analizom podataka koji su dostupni operatoru mreže. Podaci sadrže merene vrednosti ulazne energije u distributivni sistem za svaki petnaestominutni interval u toku godine i ukupnu količinu energije isporučene korisnicima. Razlika između ove dve vrednosti predstavlja stvarnu, merenu vrednost gubitaka. Sa druge strane, predstavljena je polinomska relacija uz pomoć koje je moguće te iste gubitke sračunati i samo na osnovu podataka o ulaznoj energiji. Koeficijenti ovog polinoma su estimirani primenom metode najmanjih kvadrata, minimizujući razliku između merenih i izračunatih gubitaka. Vrednosti koeficijenata su ograničene na osnovu njihove fizičke interpretacije i iskustva autora rada. Rezultati pokazuju da je minimizacija bila uspešna i da se gubici mogu jasno klasifikovati na ovaj način. Pored toga, izračunati koeficijenti se mogu koristiti za estimaciju gubitaka u budućnosti. Koncept je proširen klasterizacijom meseci u sezone. Dobijeni rezultati pokazuju očekivanu raspodelu meseci, sa izuzetkom meseca Maja, koji je ovde klasifikovan kao zimski. Ova činjenica, zajedno sa rezultatima prethodnih proračuna, dovodi do zaključka da je u toku ovog meseca verovatno došlo do greške u očitavanju potrošnje električne energije. Ipak, ukupni rezultati proračuna su vrlo dobri, i ceo koncept je moguće dalje unaprediti. Buduća istraživanja će se fokusirati na obradu podataka i od prethodnih godina. Ova obimnija obrada bi mogla pomoći istraživačima da uoče neke specifične trendove koji se iz jedne godine ne vide, kao i da dovede do bolje klasterizacije meseci s obzirom da jedna sezona može početi u jednoj godini a završiti se u drugoj.

## LITERATURA

- [1] Zhou K, Fu C, Yang S, 2016, "Big data driven smart energy management: From big data to big insights", *Renewable and Sustainable Energy Reviews* 56, pp. 215–225, DOI: 10.1016/j.rser.2015.11.050
- [2] Koochakzadeh O N A, Singh R. P, Golab L, Keshav S, 2014, "Computing Electricity Consumption Profiles from Household Smart Meter Data," in *EDBT Workshop on Energy Data Management*, pp. 140–147.
- [3] Jiang H, Zhang J, Gao W, and Wu Z, 2015, "Fault detection, identification, and location in smart grid based on data-driven computational methods," *IEEE Transaction on Smart Grid*, vol. 5, no. 6, pp. 2947–2956, DOI: 10.1109/TSG.2014.2330624
- [4] Nikovski D. N, Wang Z, Esenther A, Sun H, Sugiura K, Muso T, Tsuru K, 2013, "Smart meter data analysis for power theft detection", in *Proc. Int. Workshop Machine Learning and Data Mining in Pattern Recognition*, ser. LNCS, vol. 7988. Springer, pp. 379–389.
- [5] Nallagownden P, Hong T. P, 2011, "Development of a new loss prediction method in a deregulated power market using proportional sharing," in *Proceedings of Power Engineering and Optimization Conference (PEOCO), 2011 5th International*. Shah Alam, Selangor, Malaysia: IEEE, June 2011, pp. 48–53.
- [6] Petrušić A, Janjić A, 2016, "Economic Regulation of Electric Power Distribution Network" in *Proceedings of 2<sup>nd</sup> Virtual International Conference on Science, Technology and Management in Energy*, Niš, Serbia, pp. 25-32
- [7] Liu W. D, Zeng T. J, Jun T, Shang Guan L. L, Li B. J, 2014, "Research on Electric Power with Development and Application of Line Loss Rate Forecasting Software Based on MLRM-GM," *Advanced Materials Research*, vol. 977, pp. 182–185.
- [8] Li K, Sun Z. Q, Wang M, 2014, "Theoretical Line Loss Calculation of Distribution Network Considering Wind Turbine Power Constraint," *Advanced Materials Research*, vol. 986-987, pp. 630–634.
- [9] Ding Q, Abur A, 2006, "Transmission Loss Allocation Based on a New Quadratic Loss Expression," *IEEE Transactions on Power Systems*, vol. 21, pp. 1227 – 1233.
- [10] Gross G, Tao S, 2000, "A Physical-Flow-Based Approach to Allocating Transmission Losses in a Transaction Framework," *IEEE Transactions on Power Systems*, vol. 15, pp. 631 – 637.
- [11] Elmitwally, Eladl A, Abdelkader S. M, 2015, "Efficient algorithm for transmission system energy loss allocation considering multilateral contracts and load variation," *IET Generation, Transmission & Distribution*, vol. 9, pp. 2653 – 2663
- [12] Tian L, Wang Q. Q, Cao A. Z, 2013, "Research on SVM Line Loss Rate Prediction Based on Heuristic Algorithm," *Applied Mechanics and Materials*, vol. 291-294, pp. 2164–2168.
- [13] Ren Y, Zhang X. G, Huang X. C, 2014, "Study on the Prediction of Line Loss Rate Based on the Improved RBF Neural Network," *Advanced Materials Research*, vol. 915-916, pp. 1292–1295.
- [14] Nagendra Rao P. S, Deekshit R, 2006, "Energy Loss Estimation in Distribution Feeders," *IEEE TRANSACTIONS ON POWER DELIVERY*, VOL. 21, NO. 3.
- [15] Gustafson M.W, Baylor J. S, Mulnix S. S, 1988, "Equivalent hours loss factor revisited," *IEEE Trans. Power Syst.*, vol. 3, no. 4, pp. 1502–1507, Nov.
- [16] Gustafson M. W, 1983, "Demand, energy and marginal electric system losses," *IEEE Trans. Power App. Syst.*, vol. PAS-102, no. 9, pp. 3189–3195, Sep.
- [17] Gustafson W, Baylor J. S, 1989, "Approximating the system losses equation," *IEEE Trans. Power Syst.*, vol. 4, no. 3, pp. 850–855, Aug. 1989
- [18] Järvinen M, 2013, "Developing network loss forecasting for Distribution System Operator", Master of Science Thesis, Tampere University of Technology.
- [19] <https://www.ofgem.gov.uk/ofgem-publications/43519/sohn-overview-losses-final-internet-version.pdf>, accessed on 26.2.2019.
- [20] Löfberg J, 2004, "YALMIP: A Toolbox for Modeling and Optimization in MATLAB", In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2–4 September 2004, DOI: 10.1109/CACSD.2004.1393890
- [21] Waltz R. A, Morales J. L, Nocedal J, Orban D, 2006, "An interior algorithm for nonlinear optimization that combines line search and trust region steps," *Mathematical Programming*, Vol 107, No. 3, pp. 391–408
- [22] James G, Witten D, Hastie T, Tibshirani R, 2015, "An Introduction to Statistical Learning with Applications in R", *Springer Texts in Statistics*, Springer Science + Business Media New York, 2013 (Corrected at 6th printing 2015), DOI: 10.1007/978-1-4614-7138-7