

Kratkoročno napovedovanje povprečne dnevne porabe električne energije z uporabo adaptivnega mehkega modeliranja

Gregor Černe, Dejan Dovžan, Igor Škrjanc
Univerza v Ljubljani, Fakulteta za elektrotehniko
Tržaška cesta 25, 1000 Ljubljana
gregor.cerne@fe.uni-lj.si

Short-term load forecasting using adaptive fuzzy modeling

This paper introduces an adaptive fuzzy models in the field of short-term load forecasting (STLF), where electrical load is forecast for a day ahead based on the weather forecast and day properties. The paper focuses on the daily average load forecast, which can be used as the basis for the hourly profile forecast. The motivation behind using fuzzy models for forecasting is the nonlinear dependency of the electrical load on the weather forecast, as fuzzy models are already widely used for modeling non-linear systems. Moreover, using fuzzy clustering lowers the need of expert knowledge for modeling, as different types of days are automatically detected. Firstly, developed adaptive fuzzy model build using Gustafson-Kessel clustering is presented, including adjustments developed specifically for the STLF problem. The model was tested on real data from a Slovenian distributor for years 2010-2012, where it shows significant improvement of forecast accuracy compared to other forecasting methods used in the field of STLF.

Kratek pregled prispevka

Članek vpelje adaptivno mehko modeliranje na področje kratkoročnega napovedovanja porabe električne energije, kjer napovedujemo porabo naslednjega dne na podlagi vremenske napovedi ter lastnosti dneva napovedi. Članek se osredotoči na napovedovanje dnevnega povprečja porabe, katero se lahko uporabi kot osnovo za napoved urnega profila porabe. Motivacija za vpeljavo mehkega modela je nelinearna odvisnost porabe glede na vremensko napoved, saj se mehki modeli že množično uporabljajo na področju modeliranja nelinearnih sistemov. Poleg tega uporabljeno mehko rojenje samodejno detektira roje medsebojno podobnih dni, kar zmanjša potrebo po ekspertnem znanju o sistemu. Najprej bo predstavljen uporabljen adaptivni mehki model zgrajen s pomočjo rojenja Gustafson-Kessel ter razvite prilagoditve za problem napovedovanja porabe električne energije. Model je bil testiran na realnih podatkih slovenskega električnega distributerja za leta 2010-2012, kjer uporaba razvitega modela znatno izboljša natančnost napovedi v primerjavi z ostalimi metodami napovedovanja.

1 Uvod

Dobra in natančna napoved je lahko razlika med stabilnim podjetjem in bankrotom. Eno izmed področij, na katerem večja napaka napovedi neposredno povzroči večje stroške je napovedovanje električne energije naročnikov električnih distributerjev - v primeru premajhne napovedi mora distributer v trenutku porabe manjkajočo energijo kupiti po višji ceni, v primeru prevelike pa odvečno energijo ceneje prodati, hkrati lahko dobi tudi kazen za preobremenitev omrežja.

Problem kratkoročnega napovedovanja električne energije (KNEE), ki se ukvarja z napovedjo urnega profila porabe naročnikov naslednjega dne je pogosto obravnavan v literaturi. Najbolj pogoste uporabljene metode izhajajo iz avtoregresivnega integrirnega modela drsečega povprečja (angl. ARIMA) [1][2], umetnih nevronske mreže (angl. ANN) [3][4] ter mehkih modelov [5] [6] [7] [8].

Cilj raziskovanja predstavljenega v članku je usmerjen v iskanje metode, ki bi imela višjo stopnjo avtomatske izbire parametrov ter njihovo avtomatično prilagajanje glede na aktualne podatke, hkrati pa ponudila dober pogled v strukturo modela. Razvit model napovedovanja bazira na mehkem modelu Takagi-Sugeno (TS), dobljenem z rojenjem Gustafson-Kessel, ki tudi sprotno prilagaja parametre modela z vsakim novim dnevom oziroma vzorcem. Prispevek se osredotoči na kratkoročno napovedovanje povprečne dnevne porabe, katero lahko uporabimo kot osnovo za napovedovanje urnega profila porabe.

2 Kratkoročno napovedovanje povprečne porabe električne energije

Problem KNEE je napovedati urni profil porabe električne energije za naslednji dan $E(k)$ na podlagi vremenske napovedi ter zgodovine podatkov. Sistem porabe električne energije je nelinearen sistem z veliko težko merljivimi dejavniki, saj je odvisen od človeškega obnašanja in potreb. Članek se osredotoči na del obsežnega problema KNEE, in to na napovedovanje dnevnega povprečja porabe električne energije (1). Ta se lahko uporabi

Tabela 1: Merjeni podatki

Ime	Simbol	Enota ali zaloga vrednosti
Električna energija	E	MWh
Temperatura	T	$^{\circ}C$
Sončno sevanje	Γ	W/m^2
Padavine	R	mm
Hitrost vetra	W	m/s
Dan v tednu (1 = ponedeljek)	F	1 – 7
Dan v mesecu	G	1 – 31
Mesec	M	1 – 12
Leto	Y	2010 – 2012
Praznik (1 = praznik, 0 = ostali)	P	0 or 1

kot izhodišče za napovedovanje urnega profila.

$$y(k) = \bar{E}(k) = \frac{1}{24} \sum_{l=1}^{24} E_l(k) \quad (1)$$

Model napovedovanja je razvit in testiran na podatkih velikega slovenskega električnega distributerja za jugo-zahodno regijo Slovenije. Poleg odjema električne energije sta bili na voljo tudi 2 skupini podatkov: vremenski podatki (temperatura, količina padavin, hitrost vetra ter sončno sevanje) in lastnosti dneva (dan v tednu, mesec, praznik ipd.). Vsi razpoložljivi podatki so zbrani v tabeli 1.

Podatkov o vremenski napovedi ni bilo na voljo, zato smo kot vhod v model vzeli kar izmerjene vrednosti vremenskih podatkov.

3 Predprocesiranje podatkov

V poglavju bodo predstavljene tri procedure predprocesiranja podatkov - omejitev padavin, dnevno povprečenje ter vpeljava spremenljivke dneva.

3.1 Omejitev padavin

Pri analizi vpliva padavin na porabo je izstopal vpliv človeških navad - odvisnost porabe od padavin ni bila linearna, ampak je bila razlika le med deževnimi in ne-deževnimi dnevi. Predpostavimo lahko, da padavine ne vplivajo direktno na porabo kot temperatura, ampak vpliva na obnašanje ljudi - ali bodo ostali doma zaradi dežja ali pa šli na sprehod oziroma na izlet na sončni dan.

Omenjeno smo v model vključili tako, da smo urne padavine omejili navzgor na $R_{\max} = \frac{10 \text{ mm}}{24 \text{ h}}$ (2). Vrednost $\frac{10 \text{ mm}}{24 \text{ h}}$ je definirana tako, da je največja vsota padavin v dnevu enaka 10 mm.

$$R_l(k) = \begin{cases} R_l(k), & R_l(k) \leq R_{\max} \\ R_{\max}, & R_l(k) > R_{\max} \end{cases}, \forall (l, k), \quad (2)$$

kjer je $R_l(k)$ količina padavin v k -tem dnevu v l -i uri.

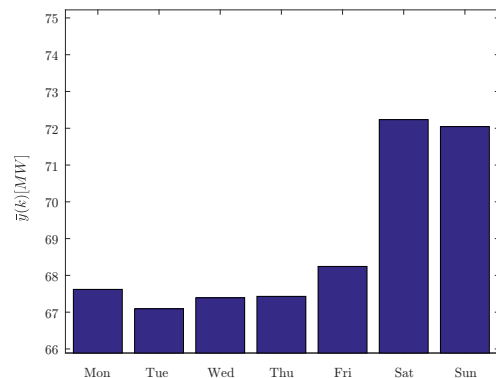
3.2 Dnevno povprečenje

Vsi podatki v tabeli 1 so bili vzorčeni na eno uro. Ker se članek osredotoči na napovedovanje povprečne porabe, smo za vhod v model napovedovanja povprečili tudi vremenske podatke (temperatura, količina padavin, hitrost vetra ter sončno sevanje).

Opomba: podatki o padavinah so bili pred povprečenjem najprej omejeni kot predstavljeno v podpoglavju 3.1.

3.3 Spremenljivka dneva

Mnoge lastnosti dneva so kategorični podatki (dan v tednu, praznik, mesec), katere potrebujemo vključiti v linearni lokalni model. Na sliki 1 je prikazana primerjava povprečne porabe na različne dni v tednu. Opazimo lahko razliko med delavniki in vikendom. Hkrati opazimo tudi vpliv vikenda (zaradi podaljšanih vikendov, ki ji jih ljudje vzamejo), saj na ponedeljek ter petek opazimo višjo povprečno porabo v primerjavi z dnevi v sredini tedna (torek do četrtek). Pri analizi dni, ki ležijo med praznikom ter vikendom, je bilo opaženo da je poraba $\bar{E}(k)$ bolj podobna dnevom tekom



Slika 1: Povprečna dnevna poraba glede na dan v tednu.

vikenda kot delavnikom. Tako lahko podatek o dnevu v tednu ter prazniku združimo v omejeno numerično spremenljivko, ki nam pove koliko je dan podoben delavniku oz. prazniku.

Z namenom upoštevanja naštetih ugotovitev je vpeljana posebnost D^* , ki ima vrednost 1 za praznike, sobote ter nedelje in 0 pri ostalih dneh.

$$D^*(k) = \begin{cases} 1, & F(k) \in \{6, 7\} \vee P(k) = 1 \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

Naslednji korak je vpeljati robne dni (dni pred in po posebnem dnevu) v razviti model. Tako je bila vpeljana spremenljivka vpliva sosednjih dni D^\pm , ki predstavlja vpliv prejšnjega ($k-1$) ter naslednjega ($k+1$) dneva na k -ti dan. Spremenljivka D^\pm je sestavljena iz treh prispevkov: posebnost prejšnjega dneva $D^*(k-1)$ utežena s parametrom g_- , posebnost naslednjega dneva $D^*(k+1)$ utežena s parametrom g_+ ter produkt posebnosti obeh $D^*(k-1)$ ter $D^*(k+1)$ utežena s parametrom g_\pm . Zadnji prispevek $g_\pm D^*(k-1)D^*(k+1)$ je vpeljan za modeliranje dni med dvema posebnimi dnevi.

$$D^\pm(k) = g_+ D^*(k+1) + g_- D^*(k-1) + g_\pm D^*(k-1)D^*(k+1) \quad (4)$$

$$g_+ + g_- + g_\pm \leq 1, \quad g_+, g_-, g_\pm \in [0, 1] \quad (5)$$

Parametri g_- , g_+ ter g_{\pm} so uporabniško določeni, vendar vrednosti parametrov morajo izpolnjevati pogoj (5). Tako je vsiljena predpostavka, da so dnevi kombinacija posebnih ter navadnih dni.

Zadnji korak predstavlja definicija spremenljivke dneve D , ki združi vse omenjene vplive. Definirana je kot večja vrednost med D^{\pm} ter D^* (6). Če je k -ti dan poseben ima vrednost $D(k) = 1$, drugače je pa vrednost izračunana na podlagi vpliva sosednjih dni $D(k) = D^{\pm}(k)$

$$D(k) = \max \left[D^{\pm}(k), D^*(k) \right] \quad (6)$$

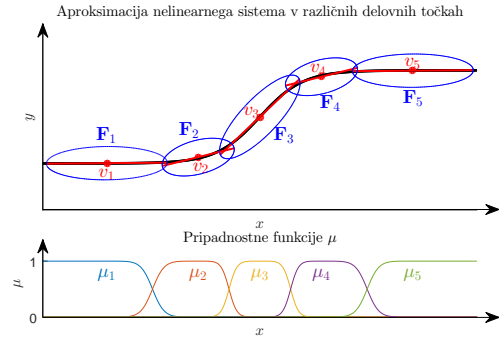
4 Mehki model Takagi-Sugeno in identifikacija

Mehki model Takagi-Sugeno je močno orodje za modeliranje nelinearnih sistemov, saj omogoči izgradnjo kompleksnega modela iz večin enostavnih, kot je na primer linearni model. Princip delovanja je prikazan na sliki 2. Definijsko področje nelinearnega sistema razdelimo na manjša mehka področja, kjer obnašanje sistema lahko dobro aproksimiramo z lokalnim linearnim modelom. Izhod mehkega modela opišemo kot vsoto izhodov lokalnih modelov, uteženih s pripadnostjo vzorca i -temu področju (7).

$$\hat{y} = \sum_{i=1}^m \mu_i(\mathbf{x}_p) \hat{y}_i(\mathbf{x}) \quad , \quad (7)$$

kjer \hat{y} predstavlja izhod modela TS, \mathbf{x} regresijski vektor meritev, \mathbf{x}_p partijski vektor meritev, $\mu_i(\mathbf{x}_p)$ vrednost pripadnostne funkcije i -tega roja pri spremenljivkah \mathbf{x}_p ter $\hat{y}_i(\mathbf{x})$ izhod i -tega lokalnega modela pri regresijskem vektorju \mathbf{x} .

Z definiranjem pripadnostnih funkcij $\mu_i(\mathbf{x}_p)$ določimo velikosti in oblike mehkih področij. Za določitev področij je bilo uporabljen iterativni algoritem rojenja Gustafson-Kessel (GK), s katerim najdemo področja kot roje, opisane s centrom \mathbf{v} ter kovariančno matriko F . Podrobnosti algoritma so opisane v [9]. Uporabljeni sta bila tudi varnostna mehanizma proti singularnosti F_i matrik [10].



Slika 2: Koncept mehkega modela za aproksimacijo nelinearne funkcije. Na zgornjem grafu je s črno prikazana nelinearna funkcija, z modrimi elipsami so obkrožena mehka področja, z rdečo pa izhodi linearnih modelov mehkih področij. Spodnji graf prikazuje funkcije pripadnosti rojem.

En izmed problemov rojenja GK je v neupoštevanju kvalitete modela pri particioniranju prostora. Za model napovedovanja je bila razvita kombinirana pripadnostna funkcija μ_i^* (10), ki jo sestavljata pripadnostna funkcija vhodnega prostora $\mu_i(d_i)$ (8) ter pripadnostna funkcija kvalitete izhoda modela $\mu_i(\zeta_i)$ (9).

$$\mu_i(d_i) = \frac{d_i^{-\frac{2}{\eta-1}}}{\sum_{j=1}^m d_j^{-\frac{2}{\eta-1}}} \quad (8)$$

$$\zeta_i(k) = f_y(\mathbf{y}(k), \hat{\mathbf{y}}_i(k)) \quad (9)$$

$$\mu_i^* = \sqrt{\mu_i(d_i) \cdot \mu_i(\zeta_i)} \quad (10)$$

kjer je d_i Mahalanobisova razdalja do centra roja, ζ_i razdalja med izhodom modela ter izmerjenim izhodom, $\mathbf{y}(k)$ vektor izmerjenega izhoda k -tega vzorca, $\hat{\mathbf{y}}^{(q-1)}(k)$ izhod i -tega lokalnega modela k -tega vzorca, f_y pa funkcija razdalje v izhodnem prostoru, ki je izračunan kot vsota absolutnih razlik komponent.

Z vpeljavo kombinirane pripadnostne funkcije μ_i^* dosežemo, da za upoštevanje določenega vzorca pri računanju lokalnih modelov potrebuje biti in blizu centra roja, kot tudi mora ustrezati trenutnemu lokalnemu modelu. S tem se izognemo upoštevanju napačnih vrednosti (angl. outliers) pri računanju lokalnih modelov.

Za enolično določitev modela TS je potrebno še določiti linearne modele $\hat{y}_i(\mathbf{x})$. Parametre i -tega linearnega modela izračunamo s pomočjo uteženih linearnih kvadratov (11), kjer za uteži vzorcev vzamemo njihove pripadnosti i -temu roju $\mu_i(\mathbf{x}_p)$.

$$\boldsymbol{\theta}_i = [\mathbf{X}^T \mathbf{W}_i \mathbf{X}]^{-1} \mathbf{X}^T \mathbf{W}_i \mathbf{y}, \quad (11)$$

kjer je \mathbf{W} diagonalna matrika z utežmi vzorcev.

Tako parametre lokalnih modelov kot tudi pripadnostne funkcije sprotno prilagajamo glede na napako modela TS, saj je poraba električne energije časovno spremenljiv sistem in se mora model spremembam tudi prilagajati. Prilagoditev parametrov je bila implementirana z rekurzivnim rojenjem GK ter rekurzivno metodo najmanjših kvadratov, kateri sta opisani v [11].

5 Razviti model

Model povprečja dnevne porabe je bil načrtan na podlagi podatkov od 1.1.2010 do 31.12.2011 izključujoč praznike in dneve s 23 meritvami (zaradi menjave ure na poletni čas). Podatki so bili razdeljeni na učno množico (1.1.2010 do 31.12.2010) ter testno množico (1.1.2011 do 31.12.2011).

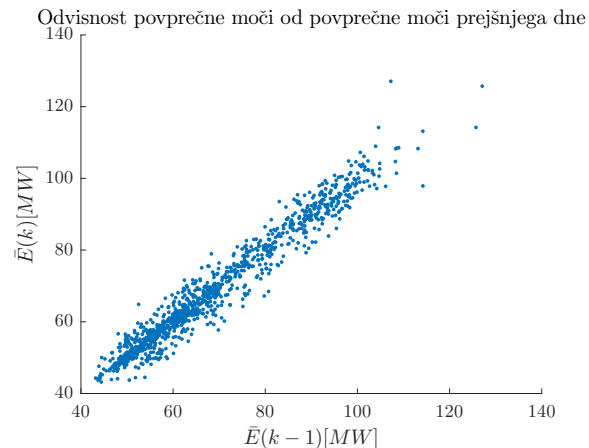
Načrtan model je razložen v poglavju 4. Močna korelacija med porabo k -tega dne $\bar{E}(k)$ ter $(k-1)$ -tega dne $\bar{E}(k-1)$ je razvidna na sliki 3, tako da je poraba prejšnjega dne vzeta kot izhodišče za napoved naslednjega dne. To je realizirano z enačbo (12).

$$\hat{\bar{E}}(k) = \bar{E}(k-1) + \Delta\hat{\bar{E}}(k), \quad (12)$$

kjer je $\hat{\bar{E}}(k)$ povprečna napoved, $\bar{E}(k-1)$ izmerjeno dnevno povprečje in $\Delta\hat{\bar{E}}(k)$ napovedana sprememba dnevnega povprečja med prejšnjim ter napovedovanim dnevom. Z enačbo je doseženo omejitve problema iz napovedovanja povprečja na napovedovanje spremembe med povprečji sosednjih dni.

5.1 Model TS

Za enolično določitev modela TS iz poglavja 4 je potrebno definirati vhodne vektorje \mathbf{x} in \mathbf{x}_p .



Slika 3: Graf povprečja k -tega dneva $\bar{E}(k)$ v odvisnosti od povprečja $(k-1)$ -tega dneva $\bar{E}(k-1)$

Regresor \mathbf{x} vsebuje razlike najbolj vplivnih spremenljivk na porabo med sosednjimi dnevi:

- razlika povprečne temperature $\Delta\bar{T}(k)$,
- razlika povprečnega sončnega sevanja $\Delta\bar{\Gamma}(k)$,
- razlika povprečne količine padavin $\Delta\bar{R}(k)$,
- razlika povprečne hitrosti vetra $\Delta\bar{W}(k)$,
- razlika spremenljivke dneva $\Delta D(k)$.

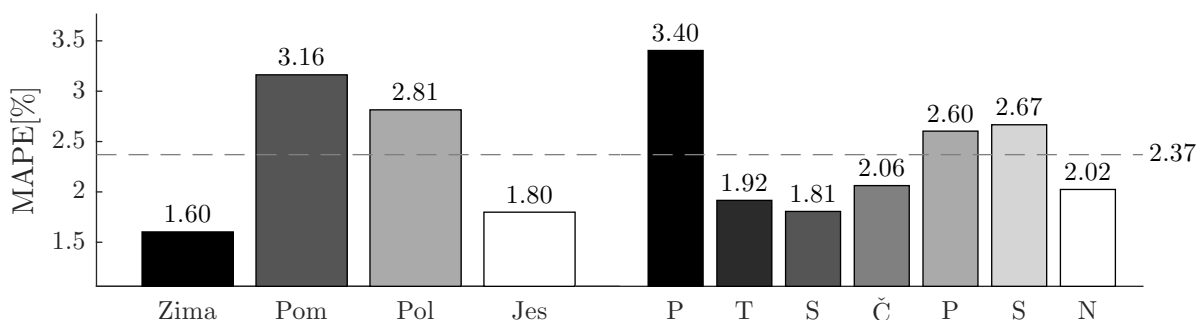
Particioniranje prostora je bilo usmerjeno v ločitev med letnimi časi, zato so bile v particijski vektor vključene naslednje spremenljivke:

- povprečna temperatura $\bar{T}(k)$,
- povprečna hitrost vetra $\bar{W}(k)$,
- povprečna poraba $(k-1)$ -tega dneva $\bar{E}(k-1)$,
- razlika spremenljivke dneva $\Delta D(k)$.

Elementi vektorja $\mathbf{x}_p(k) = [x_{p_1}(k), \dots, x_{p_p}(k)]^T$ so bili normirani glede na celotno zbirko podatkov.

Pri napovedovanju $\bar{E}(k)$ je $\bar{E}(k-1)$ neposredno vključen v regresijski vektor \mathbf{x} za poenostavitev modela, saj lahko pričakujemo njegov linearni koeficient blizu 1. Omenjeno bo rezultiralo v razvitu izhodne funkcije v (12).

Pripadnostne funkcije so definirane z rojenjem GK in adaptirane z rojenjem rGK. Parametri lokalnih modelov so bili izračunani z metodo WLSE



Slika 4: Primerjava MAPE po letnih časih (levo) po dnevu v tednu (desno). Prekinjena črta predstavlja MAPE čez leto 2011.

ter adaptirani z metodo rWLSE. Parametri γ_F , γ_c in γ_θ so bili nastavljeni ročno, m in η pa na podlagi eksperimentalnih rezultatov, kjer smo primerjali rezultate glede na različne kombinacije omejenih parametrov.

6 Rezultati in komentar

Ker dodatni stroški rastejo linearno z absolutno napako napovedi, je bila za ocenjevanje modela uporabljena mera povprečne absolutne procentne napake (angl. mean absolute percentage error oz. MAPE) (13).

$$\text{MAPE}[\%] = 100 * \sum_{k=1}^N \frac{|\hat{E}(k) - \bar{E}(k)|}{|\bar{E}(k)|}, \quad (13)$$

kjer je N število vzorcev v testni množici.

Napaka MAPE celotne validacijske množice je 2,37%. Za podrobno analizo je bila narejena primerjava napake MAPE glede na letni čas ter dan v tednu.

Pri primerjavi napake MAPE glede na dan v tednu (slika 4, desno) je napoved najboljša za srede, kateri sledita napovedi v torek ter četrtek. Predpostavljamo, da je vzrok dobre napovedi v stabilnosti porabe v sredini tedna, saj se nemerljivi dejavniki tekom sredine tedna ne spreminjajo. Hkrati opazimo najslabšo napoved porabo v ponedeljek, kar predpišemo ravno nemerljivim dejavnikom, ki se tekom vikenda spremenijo in jih moramo v ponedeljek spremembe vključiti v model.

Pri primerjavi napake MAPE glede na letni čas (slika 4, levo) je vidno da je napaka pozimi in jeseni manjša vsaj za 36%. To je lahko posledica sestave porabe, saj poraba električne energije za ogrevanje tekom zimskih ter jesenskih mesecev predstavlja večji del celotne porabe kot tekom poletnih in spomladanskih mesecev - le-to je pa lahko vključiti v model zaradi močne korelacije s temperaturo.

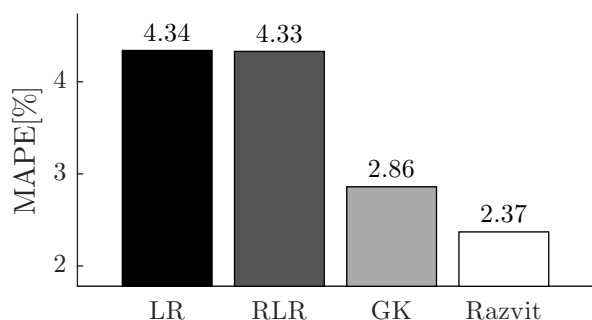
6.1 Primerjava z drugimi metodami

Razvit model napovedovanja je primerjan s pogosto uporabljenima metodama na področju KNEE - linearno regresijo (LR) ter robustno linearno regresijo (RLR). Hkrati je razvit model primerjan tudi z ne-rekurzivno različico algoritma GK.

Lahko vidimo, da že sama vpeljava modela TS in rojenja GK v problem KNEE prinese izboljšanje za 33.9%, dodatne predstavljene izboljšave pa potem skupaj za 45.3%. Po analizi rezultatov se lahko pod vprašanje postavi predpostavka o kvadratni odvisnosti od temperature, ki jo upoštevamo pri LR ter RLR, saj se največja odstopanja modela pojavijo pozimi. To odvisnost porabe od temperature algoritem razvit model zazna avtomatično pri avtomatičnem particioniranju prostora.

7 Zaključek

Prispevek predstavi uporabo mehkih modelov Takagi-Sugeno na področju kratkoročnega napovedovanja električne energije, točneje na problemu napovedovanja dnevnega povprečja porabe.



Slika 5: Primerjava MAPE z ostalimi metodami napovedovanja

Model je bil načrtan in adaptiran z uporabo rekurzivnih metod rojenja Gustafson-Kessel ter uteženih mehkih kvadratov ter omenjenim izboljšavam. Članek predlaga napovedovanje le spremembe povprečne porabe namesto direktno povprečne porabe da model zajame nemerljive vplive čim hitreje.

Napoved pridobljena z uporabo razvitega modela izboljšša napoved za 45.3% v primerjavi z metodami uporabljenimi na tem področju (linearna regresija ter robustna linearna regresija).

Rezultate napovedovanja lahko uporabimo kot osnovo za napoved urnega profila, kjer potrebujemo napovedali le še urna odstopanja.

Literatura

- [1] G. Gross in F. Galiana. Short-term load forecasting. *Proceedings of the IEEE*, zv. 75, št. 12, str. 1558–1573, 1987.
- [2] M. Cho, J. Hwang in C. Chen. Customer short term load forecasting by using ARIMA transfer function model. *International Conference on Energy Management and Power Delivery EMPD '95*, zv. 1, št. 95, str. 317–322, 1995.
- [3] A. S. Khwaja, M. Naeem in sod. Improved short-term load forecasting using bagged neural networks. *Electric Power Systems Research*, zv. 125, str. 109–115, 2015.
- [4] S. Kouhi in F. Keynia. A new cascade NN based method to short-term load forecast in deregulated electricity market. *Energy Conversion and Management*, zv. 71, str. 76–83, 2013.
- [5] D. Ranaweera, N. Hubele in G. Karady. Fuzzy logic for short term load forecasting. *International Journal of Electrical Power & Energy Systems*, zv. 18, št. 4, str. 215–222, 1996.
- [6] A. M. Al-Kandari, S. A. Soliman in M. E. El-Hawary. Fuzzy short-term electric load forecasting. *International Journal of Electrical Power and Energy System*, zv. 26, št. 2, str. 111–122, 2004.
- [7] K.-B. Song, Y.-S. Baek in sod. Short-Term Load Forecasting for the Holidays Using Fuzzy Linear Regression Method. *IEEE Transactions on Power Systems*, zv. 20, št. 1, str. 96–101, 2005.
- [8] A. Kazemi, A. F. A in M. Hosseinzadeh. A Multi-Level Fuzzy Linear Regression Model for Forecasting Industry Energy Demand of Iran. *Procedia - Social and Behavioral Sciences*, zv. 41, str. 342–348, 2012.
- [9] D. Gustafson in W. Kessel. Fuzzy clustering with a fuzzy covariance matrix. *1978 IEEE Conference on Decision and Control including the 17th Symposium on Adaptive Processes*, zv. 2, str. 761–766, 1978.
- [10] R. Babuška, P. V. D. Veen in U. Kaymak. Improved covariance estimation for Gustafson-Kessel clustering. *2002 IEEE World Congress on Computational Intelligence. 2002 IEEE International Conference on Fuzzy Systems. FUZZ-IEEE'02. Proceedings (Cat. No.02CH37291)*, zv. 2, št. usually 2, str. 8–12, 2002.
- [11] D. Dovžan in I. Škrjanc. Recursive clustering based on a Gustafson-Kessel algorithm. *Evolving Systems*, zv. 2, št. 1, str. 15–24, 2011.