

dr  
498



0052618

COBISS

KNJIŽNICA FE IN FRI

UNIVERZA V LJUBLJANI  
Fakulteta za elektrotehniko

DAVID GERBEC

**METODE ZA UGOTAVLJANJE PORABE ELEKTRIČNE  
ENERGIJE ODJEMALCEV BREZ SPROTNIH MERITEV**

Doktorska disertacija

Ljubljana, junij 2005

52618



UNIVERZA V LJUBLJANI

Fakulteta za elektrotehniko

DAVID GERBEC

**Metode za ugotavljanje porabe električne energije odjemalcev  
brez sprotnih meritev**

Doktorska disertacija

mentor: prof. dr. Ferdinand Gubina

Ljubljana, junij 2005

## ZAHVALA

Zahvaljujem se mentorju prof. dr. Ferdinandu Gubini za strokovno pomoč, nasvete in mentorstvo na podiplomskem študiju in pri izdelavi disertacije.

Za mnenja, nasvete, pomoč in prijetno vzdušje se zahvaljujem sodelavcem v Laboratoriju za elektroenergetske sisteme in Laboratoriju za elektroenergetske strategije. Hvala Tadeji, Robiju, Gregorju, Samu, Milošu, Ivu, Martinu, Ludviku, Tomažu in Andreju. Prav tako se zahvaljujem bivšim sodelavcem Davidu, Dejanu, Andreju, Boštjanu in Tomažu.

Zahvaljujem se tudi svojim staršem in vsem najbližjim, ki so me vedno spodbujali in podpirali na moji študijski poti. Posebna zahvala gre Vladu, ki je pregledal besedilo.

Moje strokovno usposabljenje je potekalo v okviru projekta mladih raziskovalcev, ki ga je financiralo Ministrstvo za šolstvo znanost in šport Republike Slovenije.

Hvala vam.



# Kazalo

<b>KAZALO</b> .....	<b>I</b>
<b>SEZNAM UPORABLJENIH SIMBOLOV IN OZNAK</b> .....	<b>V</b>
<b>POVZETEK</b> .....	<b>IX</b>
<b>ABSTRACT</b> .....	<b>XI</b>
<b>1. UVOD</b> .....	<b>1</b>
1.1. Predstavitev tematike .....	1
1.2. Namen disertacije .....	2
1.3. Potek dela .....	3
<b>2. PREGLED STANJA</b> .....	<b>5</b>
2.1. Dejavniki, ki vplivajo na porabo električne energije .....	5
2.1.1. Dejavniki odjemalca .....	6
2.1.2. Časovni dejavniki .....	6
2.1.3. Klimatski dejavniki .....	6
2.1.4. Ostali dejavniki .....	7
2.2. Zajem podatkov o porabi električne energije .....	7
2.2.1. Razporeditev in vrsta merilne opreme .....	7
2.2.2. Načini merjena porabe in njihov vpliv na obliko OD-jev .....	9
2.3. Pregled metod določanja porabe odjemalcev .....	10
2.3.1. Preostala krivulja .....	10
2.3.2. Sintetični postopek .....	12
2.3.3. Analitični postopek .....	13
2.4. Primerjava postopkov .....	15
2.5. Pregled literature na področju oblikovanja NOD-ov .....	18
2.5.1. Modeliranje porabe gospodinjstev z uporabo verjetnostnih funkcij .....	18
2.5.2. Določanje OD-jev na podlagi meritev porabe odjemalcev .....	19
<b>3. DOLOČANJE NOD-OV</b> .....	<b>25</b>
3.1. Priprava podatkov za nadaljnjo obdelavo .....	27

---

3.1.1.	Valčki.....	28
3.1.1.1.	Valčna multiresolucijska analiza.....	29
3.1.1.2.	Odstranjevanje šuma (Glajenje).....	32
3.1.2.	Normiranje .....	35
3.2.	Razpoznavanje vzorcev .....	38
3.2.1.	Fuzzy C-means rojenje (FCM) .....	39
3.2.1.1.	Mere veljavnosti rojenja .....	42
3.2.1.2.	Ponazoritev uporabe FCM algoritma in mer veljavnosti rojenja.....	44
3.2.2.	Hierarhične metode in mere podobnosti.....	48
3.2.3.	Nevronska omrežja .....	52
3.2.3.1.	Samo-organizirajoča omrežja .....	52
3.2.3.2.	Verjetnostna Nevronska Omrežja.....	55
3.3.	Dodelitev NOD-ov posameznim UO-jem.....	58
3.3.1.	Značilnosti odjemalčeve dejavnost .....	58
3.3.2.	Uporaba tipa odjemalčeve dejavnosti za dodeljevanje NOD-ov .....	61
3.4.	Ugotovitve .....	63
<b>4.</b>	<b>NOVA METODA UGOTAVLJANJA PORABE ELEKTRIČNE ENERGIJE ...</b>	<b>65</b>
4.1.	Izpeljava metode in kriteriji, ki jih mora upoštevati .....	65
4.2.	Optimizacija rekonstruiranih OD-jev.....	68
4.2.1.	Splošno o optimizaciji in linearnem programiranju.....	68
4.2.2.	Določitev meje popravljanja rekonstruiranih OD-jev .....	70
4.3.	Zapis matematičnega modela linearnega programiranja.....	70
<b>5.</b>	<b>REZULTATI.....</b>	<b>75</b>
5.1.	Rezultati oblikovanja NOD-ov .....	75
5.1.1.	Določanje števila skupin OD-jev .....	76
5.1.2.	Rezultati razvrščanja OD-jev s FCM algoritmom .....	82
5.1.3.	Rezultati razvrščanja s hierarhično metodo .....	84
5.1.4.	Rezultati razvrščanja s samoorganizirajočim omrežjem .....	86
5.1.5.	Primerjava metod razvrščanja.....	87
5.1.6.	Rezultati določitve in dodelitev NOD-ov posameznim UO-jem .....	89
5.2.	Primerjava meritev in rekonstruiranih OD-jev .....	90

---

---

5.3.	Rezultati optimizacije obremenitvenih diagramov .....	93
5.4.	Uporaba izsledkov disertacije za nadaljnje delo.....	97
<b>6.</b>	<b>SKLEP</b> .....	<b>101</b>
6.1.	Rezultati dela.....	101
6.2.	Izvirni prispevki disertacije .....	102
<b>7.</b>	<b>LITERATURA</b> .....	<b>103</b>
<b>PRILOGE</b>	.....	<b>107</b>
Priloga A	.....	107
Priloga B	.....	108

# Seznam uporabljenih simbolov in oznak

## OZNAKE

EES	elektroenergetski sistem
RTP	razdelilna transformatorska postaja
OD	obremenitveni diagram
UO	upravičen odjemalec
SO	sistemski operater
DO	distribucijski operater
NOD	nadomestni obremenitveni diagram
TP	transformatorska postaja
FCM	fuzzy c-means algoritem
SOM	samo-organizirajoča omrežja
VNO	verjetnostno nevronska omrežje
sym4	Simlet valček reda 4
NC	noise clustering
NO	nevronska omrežja
SKD	standardni klasifikat dejavnosti
OD	obremenitveni diagram
NO	nevronska omrežja

## SIMBOLI

$P$	delavna moč
$Q$	jalova moč
$U$	napetost
$I$	tok
$P_k$	konična moč
$p(t)$	trenutna moč
$T_0$	perioda vzorčenja
$t$	čas
$P_{\text{izgub}}$	obremenitvena krivulja izgub
$P_{\text{merjeni}}$	obremenitvena krivulja merjenih odjemalcev
$P_{\text{nemerjeni}}$	obremenitvena krivulja nemerjenih odjemalcev
$P_{\text{OD}}$	poraba obremenitvenega diagrama
$P_{\text{NOD}}$	obremenitev nadomestnega obremenitvenega diagrama
$P_{\text{O}}$	rekonstruiran odremenitveni diagram
$P_{\text{D}}$	dobaviteljev odremenitveni diagram
$P_{\text{PK}}$	preostala krivulje

$z_k$	delež k-tega odjemalca ali dobavitelja
$W_{OL}$	letna energija odjemalca
$\Pi$	množica vseh odjemalcev, ki se napajajo iz distribucijskega omrežja,
$\Gamma$	množica ali posamezen odjemalcev, katerim energijo dobavlja k-ti dobavitelj
$\Xi$	število dobaviteljev na območju distribucijskega omrežja
$P_B$	bazna moč, ki predstavlja povprečno dnevno moč odjemalca
$f_{G1}$	koeficient ne-monotonosti
$f_{G2}$	koeficient polnosti
$f_{G3}$	koeficient modulacije med pasovnim in trapeznim delom
$MC_{ph}$	koeficient modulacije vršnega dela
$MC_{oph}$	koeficient modulacije ne-vršnega dela
$MC_{23-6}$	koeficient modulacije pasovnega dela
$P_{min}$	minimalna moč OD-ja
$P_{max}$	maksimalna moč OD-ja
$P_{med}$	povprečna moč OD-ja
$P_{med\ ph}$	povprečna moč koničnega dela OD-ja
$P_{med\ oph}$	povprečna moč nekoničnega dela OD-ja
$P_{med\ 23-6}$	povprečna moč pasovnega dela OD-ja
$P_{med\ 6-23}$	povprečna moč trapeznega dela OD-ja
$I_S$	sestavljani indeks
$x_N$	normiran OD
$\psi(t)$	valčki
$f(t)$	časovna vrsta ali funkcija
$c_{a,b}$	koeficinti valčne transformacije
$a$	faktor skaliranja pri valčkih
$b$	faktor pomika po časovni osi $t$ pri valčkih
$f(n)$	časovno diskreten signal, v našem primeru je to OD
$c_d$	valčni koeficienti detajlov
$c_a$	valčni koeficienti aproksimacije
$h(n)$	visokofrekvenčni filter za dekompozicijo signala
$g(n)$	nizkofrekvenčnega filtra za dekompozicijo signala
$J$	končna resolucija ali skala razcepa signala
$h'(n)$	visokofrekvenčni filter za rekonstrukcijo
$g'(n)$	nizkofrekvenčni filter za rekonstrukcijo
$e_i$	naključne normalno porazdeljene napake s povprečjem nič in varianco $\sigma^2$
$x$	vektor, vzorec ali obremenitveni diagram
$P(c_d, p)$	pragovna funkcija
$p$	prag glajenja z valčno multiresolucijsko analizo
$z$	normirana vrednost OD-ja
$\kappa$	faktor za katerega vrednost zmanjšamo konico
$\xi$	množica konic OD-ja
$o$	parameter konice

<b>X</b>	matrika vzorcev
<b>B</b>	matrika prototipov
$\beta$	prototip
<b>C</b>	število rojev ali skupin
<b>U</b>	matrika pripadnosti
<b>u</b>	vektor pripadnosti
$d(\mathbf{x}, \beta)$	norma (razdalja) med vzorcem in prototipom
<b>m</b>	potenčni faktor mehčanja
$\Delta$	spremenljivka, ki določa pogoj za prekinitev iterativne optimizacije
$f_1, f_2$	meri ali kriterija veljavnosti rojenja
<b>S</b>	množica vzorcev razporejenih v isti roj ali skupino
$D(C_r, C_s)$	razdalja med roji ali skupinami
$d(\bar{\mathbf{x}}_r, \bar{\mathbf{x}}_s)$	razdalja med središčema rojev ali skupin
$O_r(d)$	okolica vzbujenega nevrona
<b>R</b>	radij okolice vzbujenega nevrona
<b>w</b>	uteži nevronov
$dw(t)$	sprememba vrednosti uteži nevronov
<b>a1</b>	izhodni vektor radial basis nivoja VNO-ja
<b>b</b>	sprožilni prag nevrona
<b>A1</b>	matrika izhodov prvega nivoja VNO-ja
<b>T</b>	ciljna matrika drugega nivoja
<b>A2</b>	matrika izhodov drugega nivoja VNO-ja
<b>C</b>	izhodna matrika VNO-ja
$W_{MES}$	odjemalčeva mesečna energija
$H_{MVT}$	mesečne obratovalne ure visoke tarife
$H_{MNT}$	mesečne obratovalne ure nizke tarife
$f(\mathbf{x})$	kriterijska ali ciljna funkcija
$g(\mathbf{x})$	sistem neenačb omejitev
$h(\mathbf{x})$	sistem enačb omejitev
$\sigma_{NOD}$	standardni odklon NOD-a
$\sigma_{OD}$	standardni odklon skupine odjemalcev
$W_{ODk}$	energija k-te skupine odjemalcev
$W_{NODk}$	površina pod k-tim NOD-om, ki predstavlja njegovo energijo
$P^+, P^-$	pozitivni in negativni popravek rekonstruiranega OD-ja
$P_{DPK}$	dejanska preostala krivulja
$P_{OPK}$	ocenjena preostala krivulja
$P_{O_{i,k}}^{\max}$	zgornja meja negotovosti rekonstruiranih OD-jev
$P_{O_{i,k}}^{\min}$	spodnja meja negotovosti rekonstruiranih OD-jev
$P_{MO}$	modificirani ali optimalni OD-ji skupin odjemalcev

## Povzetek

V zadnjem desetletju so elektroenergetski sistemi (EES) po svetu doživeli velike strukturne spremembe, ki so vodile do oblikovanja trgov z električno energijo in s tem večanju konkurence na strani dobave električne energije. Organizacijske spremembe prinašajo podjetjem na področju elektroenergetike nove priložnosti ter izzive. Prinašajo tudi razvoj novih, uporabnejših orodij za izvedbo zastavljenih nalog in doseganje zelenih ciljev.

Z odpiranjem trgov električne energije v Evropi je bil prvi korak v večini držav usmerjen k oblikovanju pravil za delovanje trga na debelo, kjer so sodelovali predvsem proizvajalci kot ponudniki električne energije na eni strani ter veliki odjemalci in dobavitelji končnim odjemalcem na drugi. Za slednje je določitev dinamike odjema električne energije enoumna, saj so opremljeni z ustrezno merilno opremo, ki jim omogoča njeno spremljanje ali pa so uporabljali obstoječe meritve v razdelilnih transformatorskih postajah (RTP-jih). S postopnim večanjem stopnje odprtosti trga ali nižanjem meje, ko odjemalci svobodno izbirajo svojega dobavitelja, je določanje dinamike porabe težja, saj manjši odjemalci nimajo ustreznih merilnih naprav, njihova vgraditev pri vseh odjemalcih pa je ekonomsko nesmotrna. Alternativa vgraditvi merilnih naprav je določanje dinamike porabe na podlagi tipičnih obremenitvenih diagramov (OD-jev). S tem omogočimo tudi manjšim odjemalcem prosto izbiro dobavitelja in dostop do odprtega trga z električno energijo.

Uporaba OD-jev ni novo področje, kakor tudi ne sama ideja, da jih lahko učinkovito uporabimo pri določanju porabe posameznega odjemalca, saj so jih uporabljali že v preteklosti. Nova je ideja uporabe OD-jev kot pripomočkov pri spodbujanju trgovanja z električno energijo. Njihov glavni namen uporabe je določanje in obračun odstopanj dejanske ali dobavljene električne energije od napovedane.

V disertaciji smo zasnovali in predstavili novo metodo določanja OD-jev odjemalcev brez sprotnih meritev. Metoda upošteva analitično določeno preostalo krivuljo, ki predstavlja splošno in dokaj natančno obliko odjema vseh nemerjenih odjemalcev določenega distribucijskega omrežja ter nadomestne obremenitvene diagrame (NOD-e), ki podajajo dinamiko odjema posameznih skupin odjemalcev. Ker temelji na optimizaciji z uporabo linearnega programiranja, ki najde globalni optimum, je metoda hitra in enostavna. Predlagana metoda odpravlja slabost analitičnega postopka, ki predvideva enako obliko OD-jev vseh odjemalcev, kar pa se največkrat ne sklada s stanjem v realnosti, saj imajo odjemalci OD-je različnih oblik in tudi slabost sintetičnega postopka, ker odpravlja morebitna dodatna tveganja distribucijskega operaterja za neravnotežja med porabo določeno s sintetičnim in analitičnim postopkom. V disertaciji smo zato prednosti nove metode utemeljevali na dejanskih meritvah OD-jev. Pri tem smo analizo izvedli na meritvah porabe električne energije odjemalcev za en delavnik, kjer smo predpostavili, da poznamo dnevno porabo električne energije posamezne

skupine odjemalcev. Rezultati kažejo sposobnost predlagane metode, da učinkovito odpravi pomanjkljivost obstoječih. Kljub predstavitvi rezultatov samo na podlagi analize enega dneva, lahko predlagano metodo relativno enostavno razširimo na daljše časovno obdobje, npr. en mesec, kar ustreza tudi sedanjemu obračunskemu obdobju za električno energijo.

Nova metoda določanja porabe odjemalcev brez sprotnih meritev temelji na NOD-ih, ki so izdelani na podlagi meritev OD-jev posameznih odjemalcev. Obstoječe metode določanja NOD-ov temeljijo na dveh načinih. Prvi predvideva uporabo večletnega merjenja porabe odjemalcev, razdeljenih v znane skupine, medtem ko drugi način predlaga uporabo metod razpoznavanja vzorcev, vendar ne podaja postopka, kako NOD-e dodeliti odjemalcem. Disertacija predlaga kompromis med obema načinoma določanja NOD-ov, ki temelji na metodah razpoznavanja vzorcev in dodeljuje NOD-e odjemalcem na podlagi njihove registrirane dejavnosti. V tem okviru smo analizirali OD-je upravičenih odjemalcev (UO-jev) in predlagali postopke, ki so se izkazali za učinkovite, kar smo pokazali s primerjavo rekonstruiranih OD-jev posameznih skupin odjemalcev in njihovih izmerjenih OD-jev. Pred razvrščanjem odjemalcev v podobne skupine smo OD-je gladili in izločili šum z valčno multiresolucijsko analizo ter jih normirali.

Na podlagi izsledkov, ki smo jih pridobili pri oblikovanju optimizacijskega modela, namenjenega določanju dinamike porabe odjemalcev brez sprotnih meritev, pri izdelavi NOD-ov za upravičene odjemalce in na podlagi strokovne literature s tega področja smo predlagali nadaljnje naloge, ki so potrebne za učinkovito uporabo zasnovane metode v praksi. Najpomembnejša in hkrati najzahtevnejša med njimi se zdi izdelava NOD-ov za tarifne odjemalce.

**Ključne besede:** trg z električno energijo, obračun odstopanj, obremenitveni diagrami, valčna analiza, razpoznavanje vzorcev, linearno programiranje



## Abstract

Electric power systems world wide have been witnessing tremendous changes during the last decade. The main objective and effort has been given to the deregulation and introduction of power markets causing competition between the suppliers of electric energy. These organizational changes brought new challenges and opportunities to the companies acting on such markets. Development of new applicable tools has been encouraged to carry out new tasks in order to gain desirable goals.

With the gradual opening of the power markets in Europe, the first step in the majority of countries has been done toward the development of rules for the wholesale power market. At this market, only power producers on one side and big consumers or suppliers on the other could participate. For them, determination of the actual power consumption in each time interval was not a problem, since they have appropriate metering equipment enabling them to keep track and record their load profiles. On the other hand, power suppliers were using metering equipment in the substations in order to determine the load pattern of all consumers supplied from the substation. With the gradual increasing of the level of competition when smaller consumers could participate at the power market, this is no longer an easy task to do. Smaller consumers do not have installed metering equipment, which could be used to determine their load profiles. However, their installation is not an issue since such decision is not an economic one. An alternative to the installation of the metering equipment is the use of the typical load profiles, enabling smaller consumers to participate at the power market.

Use of typical load profiles is not a novel concept of a research, nor it is the idea that they can be used to determine the consumers' load profiles, since they were already used in the past to help formation of retail electricity tariffs. New idea is to use them as a tool in order to motivate the competition at the power market. Their main objective is to calculate and to settle up the deviations between the actual and forecasted consumption.

The dissertation addresses a new methodology for deriving and applying load profiles to consumers without time meters. It is developed from the two well known approaches: the area model and the category model, which have some limitations. The area or analytical model assumes that all consumers supplied from a substation have the same load profile behaviour, i.e. the shape of the so called rest curve. Even if we take into account the simplicity of such a model, the assumption is too superficial since consumers have different LPs' patterns. The category or synthetic model that uses some standard or typical load profiles (TLPs) gives better insight into the consumers' LP. However, it gives some additional risk to the distribution system operators (DSO). The main limitation is in the allocation of the discrepancies between the rest curve and the reconstructed TLP to the DSO as the increase or decrease of the distribution system losses in the respective time intervals causing additional

risk. In order to overcome the limitations of the presented models, the dissertation introduces new model for the "actual" consumers' LPs determination including advantages of the both discussed models. In the new model, the rest curve would give general assumptions regarding LP's pattern considering all consumers while TLPs would give better insight into the individual consumer's LP. Since the proposed model is based on the optimization technique using linear programming algorithm giving global optimum, it is fast and simple. The proposed model has been tested on measured LPs and proved to be able dealing with the assigned task in an efficient way. Even the results were obtained on the analysis of only one day the proposed model can be easily extended to the longer period, e.g. one month, which corresponds to the existing settlement time interval.

The proposed model uses TLPs assigned to the coherent groups of consumers without time-interval meters. In the past, few methods of the TLPs establishment have been presented which are based on field measurements of individual consumer's load curves. Between them, two groups can be identified. The typical feature for the first group is that the TLPs are derived from load-survey systems according to some predefined consumers' groups or groups identified during the process of TLPs determination. The second group is obtained by identifying TLPs depending on the shape of the load curves using various pattern-recognition methods. Limitations of the first TLP-determination approach are that the measurements need to be performed over a long time period, while the second approach doesn't answer the question how to allocate the TLPs to the representative groups of consumers when using the pattern-recognition methods as a load profile modelling tool. Although some authors propose an approach of allocating the TLPs based on the observations of the individual consumers LP for a certain period, we believe that such methodology is still time consuming and expensive.

In the dissertation, we propose a methodology that overcomes limitations of the existing methods and forms clear and representative groups of consumers without long lasting measurements. It is based on the pattern recognition methods and assigns TLPs to eligible consumers according to their economic activity. In order to achieve that, wavelet multi-resolution analyses, together with probabilistic neural networks, were used. The methodology is established on the predefined consumers' classes based on clustering of normalized individual consumer's load curves. The proposed method has been successfully tested on measured LPs of eligible consumers obtained from the Slovenian distribution companies. From various applied pattern-recognition methods, the fuzzy c-means clustering algorithm has been proven to be the most efficient.

Based on the experiences obtained during the research and on the studied references, we also proposed some additional steps which should be done, in order to efficiently use the proposed methodology in practice. Among them, development of the TLPs for tariff consumers seems to be the most important and challenging.

**Keywords:** power market, energy deviations settlement, load profiles, wavelet analysis, pattern recognition methods, linear programming

# 1. Uvod

## 1.1. Predstavitev tematike

V zadnjem desetletju so elektroenergetski sistemi po svetu doživeli velike strukturne spremembe, ki so vodile do oblikovanja trgov z električno energijo in s tem večanju konkurence na strani dobave električne energije. Motivi so bili različni. V nekaterih državah so spremembe posledica želje, da se posamezne vlade znebijo finančnih obveznosti za izgradnjo novih proizvodnih virov. Drugje so bili cilji prestrukturiranje elektroenergetskega sektorja, da bi omogočal odjemalcem svobodno izbiro dobaviteljev električne energije na podlagi njene cene ter storitev, ki jih ponujajo. Razlogov za spremembe je več, vendar jih v večini primerov vodi politika [1], [2]. Organizacijske spremembe prinašajo podjetjem na področju elektroenergetike nove priložnosti ter izzive. Zaželen in vzpodbujen je razvoj novih, uporabnejših orodij za izvedbo zastavljenih nalog in doseganje zelenih ciljev.

Z odpiranjem trgov električne energije v Evropi je prvi korak večina držav usmerila k oblikovanju pravil za delovanje trga na debelo, kjer so sodelovali predvsem proizvajalci kot ponudniki električne energije na eni strani ter veliki odjemalci in dobavitelji končnim odjemalcem na drugi. Za sledenje porabljene energije v izbranem časovnem intervalu se uporablja ustrezna merilna oprema, ki meri dinamiko odjema, ali meritve v razdelilnih transformatorskih postajah. S postopnim nižanjem meje, ko odjemalci svobodno izbirajo svojega dobavitelja, je določanje dinamike porabe težja, saj manjši odjemalci nimajo ustreznih merilnih naprav, njihova vgraditev pri vseh odjemalcih pa je ekonomsko nesmotrna. Alternativa vgraditvi merilnih naprav je določanje dinamike porabe na podlagi tipičnih obremenitvenih diagramov (OD-jev). S tem omogočimo tudi manjšim odjemalcem prosto izbiro dobavitelja in dostop do odprtega trga z električno energijo.

Uporaba OD-jev ni novo področje, kakor tudi ne sama ideja, da jih lahko učinkovito uporabimo pri določanju porabe posameznega odjemalca, saj so jih uporabljali že v preteklosti. Nova je ideja uporabe OD-jev kot pripomočkov pri spodbujanju trgovanja z električno energijo in prav to nas je vodilo k oblikovanju ustreznih modelov in metod, namenjenih določanju dinamike porabe odjemalcev brez sprotnih meritev.

Področje trga z električno energijo v Sloveniji ureja energetska zakon [2] ter ustrezne uredbe in pravilniki. Z njim se na področju elektroenergetike uvajajo subjekti, ki opravljajo nove dejavnosti in so zanje odgovorni, npr. z uvedbo trga z električno energijo mora vsak dobavitelj ali upravičen odjemalec (UO) porabo ali vozni red napovedovati sam.

Energetska zakon določa, da je sistemski operater (SO) odgovoren za vodenje prenosnega sistema. Glede na čas njegove pristojnosti razdelimo na izravnavo proizvodnje in porabe ter

dodeljevanje dostopa do omrežja. Maksimalni časovni interval izravnave kot posledice razlike med napovedano in dejansko porabo obsega čas od zaključka trgovanja na borzi do konca intervala, za katerega je trgovanje potekalo. Pri dnevnem trgu znaša maksimalni časovni interval 36 ur. Odstopanja, ki pri tem nastanejo, SO zaračuna odgovornim nosilcem bilančnih skupin na podlagi pravilnika [5]. Pri tem mu pomaga distribucijski operater (DO), ki določi dejanske vozne rede posameznih bilančnih skupin, dobaviteljev ali odjemalcev. Odjemalcem, ki nimajo ustrezne merilne opreme, določi dinamiko porabe z uporabo OD-jev. Poudariti moramo, da je določanje dinamike pomembno predvsem za obračun odstopanj napovedane porabe od dejanske in ne toliko za sam obračun energije, ki ga izvedemo na podlagi števnih meritev.

V literaturi [7], ki obravnava pregled področja merjenja in oblikovanja OD-jev v svetu, zasledimo predvsem dva načina oblikovanja OD-jev. Prvi je območni model z analitični postopkom določanje odjemalčevih OD-jev, ki uporablja tako imenovano preostalo krivuljo odjema. Model je sporen, ker predpostavlja, da imajo vsi odjemalci brez sprotnih meritev porabe znotraj določenega območja distribucijskega omrežja enako obliko OD-jev. Gre za obliko preostale krivulje. Drugi model oblikovanja OD-jev je kategorijski s sintetičnim postopkom, ki uporablja nadomestne obremenitvene diagrame (NOD-e). Z vidika upoštevanja različnih oblik odjema odjemalcev je primernejši drugi model, ki pa ima tudi določene omejitve ali pomanjkljivosti. Največja pomanjkljivost je ta, da pri tem postopku razliko med preostalo krivuljo ter seštevkom rekonstruiranih OD-jev enostavno pripišemo DO-ju kot povečanje ali zmanjšanje izgub v distribucijskem omrežju. Omenjena razlika predstavlja za DO-ja določeno tveganje. Pomanjkljivost slednjega modela je tudi v dolgotrajnih postopkih določanja NOD-ov.

## 1.2. Namen disertacije

V disertaciji smo želeli celovito osvetliti področje določanja OD-jev odjemalcev brez sprotnih meritev porabe ali njihovih dobaviteljev. Zaradi pomanjkljivosti obstoječih metod predstavljamo nov model in postopke določanja OD-jev odjemalcev, ki obeta odpravo njihovih slabosti. Od predlaganega modela pričakujemo, da upošteva prednosti obstoječih modelov. Upoštevati mora končno obliko preostale krivulje ter NOD-e posameznih skupin odjemalcev. Z upoštevanjem preostale krivulje zmanjšamo ali celo odpravimo neugodne rizike, ki bi jih imel DO. Uporaba NOD-ov odpravlja pomanjkljivost, ki predvideva enako obliko OD-jev odjemalcev znotraj določenega območja distribucijskega omrežja.

V disertaciji smo zasledovali cilj, ki predstavlja nov način določanja OD-jev odjemalcev temelječ na optimalni kombinaciji preostale krivulje obremenitev in NOD-ov. Prva podaja natančnejšo informacijo o dejanski obremenitvi vseh odjemalcev brez sprotnih meritev napajanih iz distribucijskega omrežja, medtem ko NOD-i vsebujejo informacije o splošnih značilnostih obremenitev posameznih skupin odjemalcev. Pri tem moramo upoštevati to, da ohranimo ustrezne podatke o mesečni porabi električne energije ter da je razlika med seštevkom rekonstruiranih OD-jev in preostalo krivuljo minimalna ali je sploh ni.



NOD-i so bistvena podlaga predlaganega modela določanja OD-jev, zato je glavnina disertacije osredotočena temu področju. Zasedovali smo dva cilja: (i) odprava dolgotrajnih postopkov izdelave NOD-ov z uporabo metod in tehnik razpoznavanja vzorcev, ki predstavljajo novo metodo dela v elektroenergetiki, (ii) določitev jasnih in enostavnih postopkov dodeljevanja NOD-ov odjemalcem.

Namen disertacije je s predstavljenimi gradniki, preostalo krivuljo in NOD-i ter primernim optimizacijskim postopkom zasnovati model, ki naj bi bil sposoben določiti OD-je posameznih odjemalcev brez sprotnih meritev. Na podlagi tako določenih OD-jev, ki bi predstavljali dejansko porabo, bi bilo mogoče izvesti ustrežnejši obračun odstopanj dejanske porabe od napovedane za tiste odjemalce ali njihove dobavitelje, ki nimajo ustreznih meritev dinamike odjema.

### 1.3. Potek dela

Po uvodnem prvem poglavju podaja drugo poglavje pregled ožjega znanstvenega področja. Predstavlja dejavnike, ki vplivajo na porabo električne energije in načine zajema podatkov o porabi električne energije, kjer na kratko opišemo razporeditev in vrsto merilne opreme ter način merjenja, podajmo pregled obstoječih metod določanja porabe odjemalcev. Sledi primerjava obstoječih metod, kjer analiziramo njihove prednosti in pomanjkljivost, zaradi katerih smo iskali dodatne rešitve, ki bi omogočale boljšo razdelitev analitično določene preostale krivulje. V nadaljevanju obravnavamo pregled literature na področju oblikovanja NOD-ov, kjer izpostavimo modeliranje porabe z uporabo verjetnostnih funkcij ter določanje NOD-ov na podlagi meritev porabe odjemalcev.

Sledita tretje in četrto poglavje, ki sta osrednji področji raziskovanja in dajeta glavno težo disertaciji. Tretje poglavje obravnava določanje NOD-ov, ki so podlaga četrtega poglavja, v katerem predlagamo novo metodo ugotavljanja porabe električne energije odjemalcev brez sprotnih meritev.

Tretje poglavje prikazuje postopek oblikovanja in določanja NOD-ov upravičenih odjemalcev brez sprotnih meritev. Zanje smo se odločili, ker so v začetnem obdobju nastajanja tega dela edini imeli prost dostop do trga z električno energijo. V tem poglavju predlagamo enostaven ter eksakten postopek, kako dobljene NOD-e pripisati posameznemu upravičenemu odjemalcu. Postopki, ki smo jih obravnavali v disertaciji, temeljijo na uporabi metod in tehnik razpoznavanja vzorcev. Principi izdelave in dodeljevanja NOD-ov uporablja razvrščanje meritev porabe električne energije, ki smo jih pridobili od slovenskih distribucijskih podjetij, v podobne skupine. Poglavje obsega tri večje razdelke. V prvem obravnavamo pripravo podatkov za nadaljnjo uporabo, kjer želimo z glajenjem OD-jev ob uporabi valčne multiresolucijske analize in z normiranjem izločiti čimveč dejavnikov, ki vplivajo na njihovo obliko in nadaljnjo razvrščanje. V drugem razdelku obravnavamo metode razpoznavanja vzorcev, ki predstavljajo novo metodo dela v elektroenergetiki. Tretji razdelek obravnava dodeljevanje NOD-ov posameznemu upravičenemu odjemalcu.

Četrto poglavje je drugi ključni del disertacije, kjer predstavljamo novo metodo določanja porabe, ki temelji na optimalni uporabi NOD-ov in preostale krivulje. Razdelili smo ga na tri večje razdelke. V prvem obravnavamo izpeljavo predlagane nove metode in kriterije ter omejitve, ki jih moramo upoštevati pri oblikovanju njenega matematičnega modela. Glede na njene lastnosti predlagamo uporabo ustrezne optimizacijske metode. V drugem razdelku podajamo teoretične podlage optimizacije s poudarkom na linearnem programiranju, ki je hitro in daje globalni optimum. Tretji razdelek obravnava zapis kriterijske funkcije in omejitve iz prvega razdelka v matematični obliki, primerni za uporabo predlagane optimizacijske metode. V tem razdelku smo podali tudi nekatere rezultate nove metode in ugotovitve.

Peto poglavje, ki smo ga razdelili na štiri večje razdelke, prikaže rezultate in preverja novo metodo na dejanskih podatkih slovenskih distribucijskih podjetji. Prvi razdelek podaja rezultate oblikovanja NOD-ov, kjer smo z uporabo različnih metod razvrščanja in ustrezne mere veljavnosti rojenja najprej določili število skupin, v katere smo OD-je razvrstili. V nadaljevanju sledijo rezultati razvrščanja OD-jev s posamezno metodo v znano končno število skupin. Primerjavo rezultatov razvrščanja OD-jev, kjer smo izbrali najprimernejšo metodo, podaja drugi razdelek. Na podlagi njenih rezultatov smo določili končne NOD-e in jih dodelili odjemalcem. Sledi razdelek, kjer primerjamo rekonstruirane OD-je posameznih skupin odjemalcev z dejanskimi meritvami in nato še razdelek, kjer podajamo rezultate predlagane metode določanja OD-jev odjemalcev brez sprotnih meritev. V zadnjem razdelku podajamo smernice za nadaljnje delo na podlagi izsledkov iz disertacije.

Na koncu disertacije sledi sklepno poglavje, ki povzema celotno delo in navaja izvirne prispevke znanosti.

## 2. Pregled stanja

Poraba električne energije v distribucijskem omrežju se spreminja s časom in krajem v omrežju. Tem spremembam mora slediti proizvodnja, ki mora skupaj s prenosnim in distribucijskim omrežjem zagotavljati neprekinjeno oskrbo odjemalcev s kakovostno električno energijo v vsakem trenutku. Zaradi naštetih dejstev sodobni distribucijski sistemi potrebujejo čim natančnejše podatke o porabi električne energije na različnih nivojih distribucijskega omrežja, ki jih lahko uporabijo za različne namene. Tako lahko podatke o porabi električne energije poleg določanja ustrezne tarifne politike in obračuna odstopanj uporabijo tudi za načrtovanje obratovanja in izgradnje distribucijskega omrežja, vodenje porabe, ocenjevanje višine izgub v distribucijskem omrežju, napoved obremenitvenega diagrama – voznih redov, svetovanje odjemalcem ter zagotavljanje ustreznih informacij odjemalcem in drugim javnim institucijam.

Najkoristnejša informacija o porabi električne energije v distribucijskem sistemu je podatek, kako posamezen porabnik ali skupina porabnikov porablja električno energijo ob različnih urah dneva, različnih dnevih v tednu in različnih sezonah v letu. Prav tako je pomemben podatek, kolikšen del celotne porabe posameznega dobavitelja na območju izbranega distribucijskega omrežja predstavlja poraba opazovanih odjemalcev ter kako se obremenitev opazovanih odjemalcev združuje na različnih nivojih ali območjih opazovanega omrežja.

### 2.1. Dejavniki, ki vplivajo na porabo električne energije

Na porabo električne energije in s tem posledično na obliko obremenitvenega diagrama (OD-ja) vpliva več dejavnikov. Najpomembnejši med njimi so:

- dejavniki odjemalca: tip odjema, način ogrevanja, velikost stavbe, vrsta in število naprav, dejavnost odjemalca, število stanovalcev, uslužbencev, itd.;
- časovni dejavniki: ura v dnevu, dnevi v tednu (tudi posebni dnevi), letni čas – sezona;
- klimatski dejavniki: zunanja temperatura, veter, sončna obsevanost zemeljskega površja, itd.;
- ostali dejavniki: korelacije med električnimi aparati, predhodne vrednosti bremena ali obremenitvenega diagrama.

Obremenitveni diagram opisuje časovni potek porabe električne energije znotraj opazovane periode, ki jo ponavadi predstavlja en dan. Povezavo med dejavniki in obremenitvijo lahko modeliramo z različnimi metodami, vendar se večina metod največkrat ne izkažejo za uspešne. Razlog je predvsem v pomanjkanju zadostnega števila razpoložljivih podatkov ter sistematskega spremljanja in analize medsebojnih vplivov, na podlagi katerih bi lahko

zanesljivo modelirali medsebojne odvisnosti obremenitve in izbranega vplivnega dejavnika. Ti so predvsem nekateri klimatski dejavniki, dejavniki odjemalca in korelacije z drugimi spremenljivkami.

### 2.1.1. Dejavniki odjemalca

Dejavniki odjemalca, ki v prvi vrsti vplivajo na obliko odjemalčevega OD-ja, so število, tip in velikost električnih naprav in aparatov, ki jih odjemalci uporabljajo. Vendar jih lahko ne glede na spreminjajoče se število naprav, ki jih odjemalci pri svojem delu uporabljajo, združimo v skupine s podobnimi karakteristikami. Osnovna delitev odjemalcev, ki se pogosto uporablja, je delitev na gospodinjstva, industrijo ter trgovine in storitve. Predstavljena delitev je zelo groba, saj so včasih oblike OD-jev znotraj ene skupine zelo različne.

### 2.1.2. Časovni dejavniki

Poraba električne energije se časovno spreminja v odvisnosti od človekove aktivnosti in ekonomske dejavnosti odjemalcev. Prvi dejavnik ima odločujoč vpliv na obliko obremenitvenega diagrama predvsem pri gospodinjstvih, medtem ko ima ekonomska dejavnost odločujoč vpliv npr. pri industrijskih odjemalcih. Poraba je čez dan večja, medtem ko je ponoči nižja. Poleg tega se poraba opazno razlikuje pri različnih tipih dni; poraba je večja med delavniki (od ponedeljka do petka), kakor med vikendi (sobota in nedelja) in prazničnimi ali predprazničnimi dnevi. Na porabo vplivajo letni časi ali sezona. Vpliv časovnih dejavnikov na porabo zahteva analizo OD-jev na urni, dnevni in letni podlagi.

### 2.1.3. Klimatski dejavniki

Klimatski dejavniki, kakor so npr. zunanja temperatura, hitrost vetra, sončna obsevanost zemeljskega površja, vplivajo na porabo električne energije in s tem neposredno na obliko OD-ja. Vprašanje na katerega je treba odgovoriti je, pri katerih odjemalcih je vpliv najbolj opazen. Zunanja temperatura občutno vpliva na porabo odjemalcev s temperaturno odvisnimi bremenji, kot so grelne ali klimatske naprave. Čeprav zunanja temperatura ni edini dejavnik, ki vpliva na potrebe po toploti, se med ostale dejavnike uvrščajo med drugim še hitrost vetra in sončna obsevanost zemlje, ki se največkrat uporabljajo pri analizah odvisnosti porabe od zunanjih dejavnikov. Razlogi za to so predvsem v omejenem poznavanju vpliva ostalih dejavnikov. Skorajda nemogoče je npr. zadovoljivo izmeriti in modelirati vpliv spremembe sončnega obsevanja zemlje na spremembo porabe.

Čeprav je korelacija med temperaturo in porabo predvsem pri gospodinjstvih očitna, jo je včasih težko določiti zaradi vpliva termostatsko vodenih ogrevalnih naprav. Takšno je segrevanje vode v nočnem času, ko električni bojlerji akumulirajo toploto, ki je namenjena dnevni uporabi.

V praksi se večinoma kot vplivni klimatski dejavnik uporablja samo zunanja temperatura, vendar njenega vpliva ni mogoče vedno zaznati in določiti. Tako je podrobnejša analiza



korelacije med temperaturo in porabo električne energije nekaterih odjemalcev pokazala, da temperatura nima odločujočega vpliva na odjem. Včasih je bil odločujoči dejavnik tip dejavnosti, katero je odjemalec opravljal, saj je bila poraba odvisna predvsem od proizvodnega procesa in naprav, ki jih je odjemalec uporabljal.

#### 2.1.4. Ostali dejavniki

Med ostale dejavnike, ki vplivajo na porabo električne energije, štejemo medsebojne vplive električnih naprav na porabo in vrednosti obremenitev v predhodnih časovnih obdobjih. Medsebojni vpliv električnih naprav na porabo predstavlja primer, ko uporaba nekaterih električnih naprav povzroči zmanjšanje porabe električne energije, namenjene ogrevanju. Nadalje lahko uporaba ene električne naprave povzroči potrebo po uporabi druge električne naprave. Vpliv teh dejavnikov je praktično nemogoče zadovoljivo analizirati in modelirati.

Za razliko od medsebojnih vplivov električnih naprav na porabo imajo predhodne vrednosti porabe pomemben vpliv na porabo. Razlog je v periodičnosti vzorca obremenitve. Spreminjanje porabe med zaporednima dnevnoma istega tipa, npr. med dvema delavnikoma ali urama, ima močno medsebojno korelacijo. Analiza koreliranosti obremenitvenih diagramov na podlagi predhodnih podatkov omogoča predvidevanje bodoče porabe z zadovoljivo natančnostjo. To dejstvo uspešno izkoriščajo pri napovedovanju porabe na sistemskem nivoju. Za posamezne odjemalce ali skupine odjemalcev predhodnih podatkov praviloma ni, ker jih nihče ni meril, zato je napovedovanje zanje praktično nemogoče. Bilo bi tudi nesmotno, saj bi bila natančnost takšne napovedi močno vprašljiva.

## 2.2. Zajem podatkov o porabi električne energije

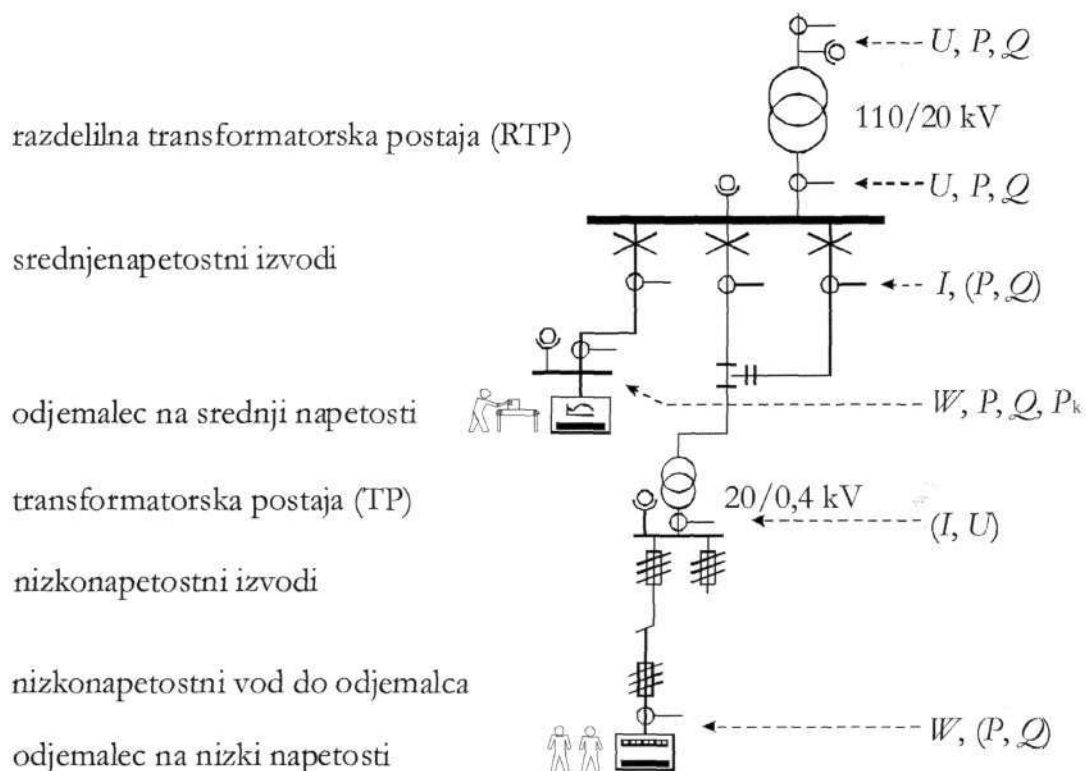
### 2.2.1. Razporeditev in vrsta merilne opreme

Poleg v prejšnjem razdelku predstavljenih dejavnikov ima na OD-je odločujoč vpliv mesto merjenja porabe v distribucijskem sistemu. Zaradi tega posvečamo v nadaljevanju nekaj pozornosti tej tematiki.

Sl. 2.1 prikazuje del distribucijskega omrežja, ki obsega nizko- in srednjenapetostno omrežje. V slovenskem EES-u poteka razmejitev med prenosnim in distribucijskim omrežjem v RTP-jih pri transformaciji napetostnega nivoja iz 110 kV na 20 kV. Takšnih razmejitvenih mest med prenosnimi in distribucijskimi podjetji je v EES-u več, vendar smo se pri prikazu omejili samo na predstavitev enega mesta, saj je konfiguracija srednje- in nizkonapetostnega omrežja načeloma enaka. Na sliki smo označili možna merilna mesta, kjer v splošnem merimo porabo električne energije ali OD-je. Merimo ju lahko neposredno z ustreznimi števci, ali ju izračunamo posredno na podlagi merjenja tokov in napetosti. Na sliki smo v oklepaju podali veličine, ki jih izračunamo, ali jih redkeje merimo.

V distribucijskem omrežju merimo porabljeno energijo,  $W$ , odjemalcev na nizki in srednji napetosti. Pri slednjih v splošnem merimo tudi njihove OD-je,  $P$  in  $Q$ , medtem ko je bila pri

prvih takšna merilna oprema praviloma izjema. Z odprtjem trga z električno energijo se število merilne opreme namenjene merjenju OD-jev povečuje, in sicer predvsem pri odjemalcih z višjo inštalirano močjo. Poleg odjemalcev merimo OD-je tudi v RTP-jih, medtem ko moč posameznega izvoda iz RTP-ja ali njegov OD ponavadi določimo računsko na podlagi merjenja toka izvoda  $I$  in napetosti zbiralk  $U$ . Odjemalcem na srednji napetosti merimo tudi konično moč  $P_k$ .



Sl. 2.1: Shematski prikaz distribucijskega omrežja. Na sliki smo označili možna merilna mesta, kjer se v splošnem meri poraba električne energije. Porabo merimo direktno s števci ali jo izračunavamo posredno na podlagi merjenja toka in napetosti.

Poleg dejavnikov, ki vplivajo na porabo električne energije in s tem na obliko OD-jev, opisanih v razdelku 2.1, vpliva na njihovo obliko tudi mesto merjenja. Primarno nas zanimajo OD-ji posameznih odjemalcev, vendar imajo ti OD-ji včasih zelo raznolike oblike z različno velikimi kratkotrajnimi porasti moči, ki so posledica naključnih vklopov in izklopov električnih naprav. S tem je njihova ponovljivost zelo majhna, kar onemogoča določanje splošnejših lastnosti ali oblike OD-ja, ki bi predstavljal celotno skupino odjemalcev. Z večanjem števila hkrati merjenih odjemalcev – meritve izvajamo na višjem nivoju, npr. v TP-ju ali na izvodu iz RTP-ja – postajajo takšni OD-ji vedno bolj gladki in s tem vedno bolj reprezentativni za celotno skupino odjemalcev. Tipičen takšen primer so gospodinjstvi porabniki. Pri tem moramo paziti, da merimo samo odjemalce iz iste skupine. Pogoji je zelo zahteven in ga je v praksi zelo težko doseči. Zavedati se moramo, da z večanjem števila hkrati merjenih odjemalcev lahko določimo najverjetnejši prispevek posameznega odjemalca k celotnemu OD-ju in ne dejanskega prispevka, ki nas niti ne zanima.

Iz Sl. 2.1 vidimo, da odjemalcem na nizki napetosti praviloma ne merimo OD-ja. Edini podatek, ki ga imajo na razpolago distribucijska podjetja, je poraba električne energije, izmerjene s števcji. Ta podatek je v času pred odprtjem trga z električno energijo popolnoma zadostoval. Z nastopom novih pogojev vodenja obratovanja distribucijskih podjetji, ki so nastali zaradi deregulacije EES-a, predstavljeni podatek ne zadostuje več. Zato uporabimo OD-je posameznega RTP-ja ali celotnega distribucijskega podjetja ali vnaprej pripravljene tipične OD-je, ki smo jih v nadaljevanju poimenovali nadomestni obremenitveni diagrami (NOD-i). Metode, ki opisujejo uporabo OD-jev za določanje porabe posameznega odjemalca, bomo predstavili v razdelku 2.3, kjer bomo opisali že obstoječi metodi in v razdelku 4, kjer predlagamo novo metodo.

## 2.2.2. Načini merjena porabe in njihov vpliv na obliko OD-jev

Podobno kot lokacija merjenja vpliva na obliko OD-je tudi način merjenja. Pri tem imamo v mislih predvsem lastnosti merilne naprave, s katero meritve OD-jev izvajamo. Lastnosti, ki nas zanimata, sta metoda in perioda vzorčenja. Obe pomembno vplivata na obliko OD-ja [3]. Metodi vzorčenja sta dve:

- Diskretno vzorčenje, ki ga uporabljamo predvsem pri procesiranju signalov in podaja vrednost obremenitve v izbranem trenutku opazovanja. Opisuje ga naslednja enačba za moč  $p$ , kjer je  $t_i = i \cdot T_0$  in  $T_0$  perioda merjenja:

$$p(t_i) = p(i \cdot T_0); \quad i = 1, \dots, n. \quad (2.1)$$

- Vzorčenje z uporabo metode integracije periode, ki daje srednjo vrednost obremenitve v izbrani periodi vzorčenja. Tipične vrednosti periode  $T_0$  so: 5 min, 10 min, 15 min, 30 min ali 60 min. Obremenitev  $p(t_i)$  izračunamo po naslednji enačbi:

$$p(t_i) = \frac{1}{T_0} \int_i^{i+T_0} p(\tau) d\tau; \quad i = 1, \dots, n. \quad (2.2)$$

Pomanjkljivost prve metode je, da za verno rekonstrukcijo signala rabimo visoko frekvenco vzorčenja ali kratko periodo vzorčenja in s tem veliko število podatkov, ki jih težko obvladujemo. Pomanjkljivost druge metode je v izgubi dinamike, ki je hitrejša od periode vzorčenja, vendar daje ta metoda bolj gladke OD-je, ker izračunava srednjo vrednost obremenitve v izbranem sorazmerno kratkem časovnem intervalu.

Za zadovoljivo obvladovanje problematike merjenja in nadaljnje obravnave porabe električne energije ne potrebujemo podatkov o njeni trenutni vrednosti in podatkov o dinamiki, hitrejši od periode vzorčenja. Zaradi tega se pri merjenju OD-jev uporablja druga metoda vzorčenja, ki je, podobno kot prva metoda, tudi občutljiva na dolžino periode vzorčenja. Razlike med izmerjeno porabo pri 15 minutnem intervalu vzorčenja in 60 minutnim intervalu vzorčenja lahko dosežejo tudi do 10 % obremenitve, ki jo izmerimo pri daljšem časovnem intervalu vzorčenja [6].

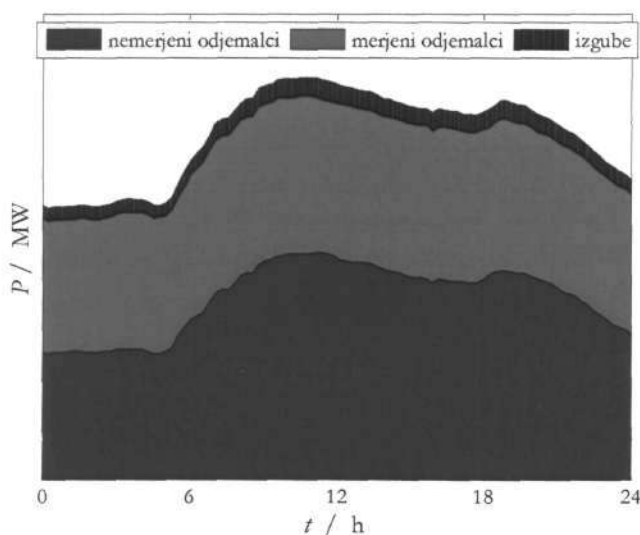
## 2.3. Pregled metod določanja porabe odjemalcev

Vpliv različnih dejavnikov na porabo električne energije in obliko OD-ja smo predstavili v predhodnih razdelkih. Sedaj prehajamo na pregled obstoječih metod namenjenih določanju porabe odjemalcev brez sprotnih meritev. Pri tem nimamo v mislih napovedovanja bodoče porabe, temveč "razdelitev" že dobavljene električne energije med posamezne odjemalce ali dobavitelje brez sprotnih meritev. Zanima nas, kolikšen del OD-ja, npr. izmerjenega v enem od RTP-jev ali na celotnem območju distribucijskega omrežja, predstavlja poraba opazovanega odjemalca ali dobavitelja.

V razdelku 2.2 smo videli, da s števci merimo porabo električne energije vsem odjemalcem, vendar ta podatek ne zadostuje pogojem trga z električno energijo. V novih pogojih obračuna energije moramo poznati tudi njihov OD. Opreme, ki bi omogočala spremljanje sprotne porabe in beleženje njene vrednosti za vsak, v splošnem, poljubno dolg interval, nimajo vsi odjemalci. Da bi lahko omogočili neoviran nastop na trgu in s tem prosto izbiro dobavitelja električne energije vsem odjemalcem, sta se oblikovali dve metodi določanja porabe odjemalcev. Obe imata za podlago s števci izmerjeno porabo električne energije odjemalcev. Prva metoda temelji na OD-jih, izmerjenih v RTP-jih, medtem ko druga metoda uporablja vnaprej oblikovane tipične OD-je, ki opredeljujejo spreminjanje porabe skupin odjemalcev znotraj obravnavanega časovnega obdobja. Metodi uporabljata analitični ali sintetični postopek [4], [7].

### 2.3.1. Preostala krivulja

Da bi si lažje predstavljali omenjena postopka in s tem razumeli vprašanje "razdeljevanja" že dobavljene energije, si oglejmo primer, ki ga prikazuje Sl. 2.2. Na sliki vidimo OD, ki opisuje celotno porabo poljubno izbranega območja. Predpostavimo, da slika predstavlja OD celotnega distribucijskega omrežja.



Sl. 2.2: Sestava OD-ja.

V splošnem mesto merjenja OD-ja ne predstavlja ovire, ki bi onemogočala pojasnitev "razdeljevanja" že dobavljene energije. Podobno razdelitev OD-ja zasledimo na več mestih v distribucijskem omrežju, npr. v RTP-jih ali celo v posameznih izvodih iz RTP-jev. Težave nastanejo, ko zahtevamo podatka o porabi in številu odjemalcev, ki sestavlja opazovan OD in kolikšen del njihove celotne porabe je vključen v OD-ju, če odjemalci niso bili celotno opazovano obdobje napajani iz istega mesta v omrežju, npr. RTP-ja. Različna prenapajanja med izvodi ali RTP-ji v omrežju so posledica npr. vzdrževalnih del. Zaradi tega je hkratna obravnava celotnega distribucijskega omrežja preprostejša.

Celotni OD distribucijskega omrežja v splošnem sestavljajo tri obremenitvene krivulje, glej Sl. 2.2:

- obremenitvena krivulja izgub,  $P_{izgub}$ ,
- obremenitvena krivulja merjenih odjemalcev,  $P_{merjeni}$ ,
- obremenitvena krivulja nemerjenih odjemalcev,  $P_{nemerjeni}$ .

Pri opredelitvi različnih tipov obremenitve smo določili krivuljo nemerjenih odjemalcev. Pri tem smo imeli v mislih podatek, da njim ne merimo OD-ja, medtem ko energijo merimo vsem odjemalcem. V nadaljevanju bomo s izrazom "nemerjeni odjemalci" označevali odjemalce brez sprotnih meritev, t.j. odjemalce, ki niso opremljeni z ustrežno merilno opremo, ki omogoča spremljanje in shranjevanje njihovih OD-jev.

Izmed predstavljenih krivulj je mogoče enoumno določiti, poleg celotnega OD-ja, samo obremenitveno krivuljo merjenih odjemalcev. Največjo neznanko predstavlja določanje obremenitvene krivulje izgub v distribucijskem omrežju. Če poznamo njen potek, določimo obremenitveno krivuljo nemerjenih odjemalcev enoumno s preprostim odštevanjem, ki ga za  $i$ -ti interval zapišemo z naslednjo enačbo:

$$P_{nemerjeni,i} = P_{OD,i} - P_{izgub,i} - P_{merjeni,i}, \quad (2.3)$$

kjer  $P_{OD}$  predstavlja celotno porabo OD-ja. Za obremenitveno krivuljo nemerjenih odjemalcev se je uveljavilo ime "preostala krivulja", ker predstavlja porabo vseh preostalih odjemalcev, katerim OD-ja ne merimo. V nadaljevanju bomo obremenitveno krivuljo nemerjenih odjemalcev (2.3) označevali s izrazom preostala krivulja.

Določanje izgub je zahtevna naloga, ki po našem poznavanju področja, praktično še ni zadovoljivo rešena v svetu, čeprav so teoretične podlage zanjo že dolgo znane. Razlog je v pomanjkanju podatkov. Kljub temu izgube lahko ocenimo na podlagi razlike med energijo, ki jo distribucijska podjetja prevzamejo od prenosnega podjetja in energijo, ki jo prodajo končnim porabnikom. Pomanjkljivost ocene je, da predpostavlja enakomerno porazdelitev izgub preko celotnega dneva ali leta kot konstantni delež celotne obremenitve [5]. To je v neskladju z dejanskim stanjem, ker se izgube spreminjajo nelinearno v odvisnosti od toka  $I$  in posredno tudi od moči  $P$ . Zaradi tega bi se veljalo tej tematiki temeljiteje posvetiti, vendar je v disertaciji ne bomo posebej obravnavali. Predpostavili bomo, da izgube poznamo.



Kot ilustracijo razmerji med posameznimi obremenitvenimi krivuljami na Sl. 2.2 podajamo okvirne podatke enega izmed slovenskih distribucijskih podjetji, kjer 40,5%-ni delež celotne dobavljene energije odpade na merjene odjemalce, 54,1%-ni delež na nemerjene odjemalce, medtem ko izgube predstavljajo 5,4%.

Preostala krivulja predstavlja podlago, ki služi za določanje OD-jev nemerjenih odjemalcev in zajema informacijo o spreminjanju porabe vseh nemerjenih odjemalcev, ki se napajajo iz opazovane točke. Preostala krivulja podaja splošno informacijo o dnevni dinamiki porabe.

Poleg razdelitve na merjene in nemerjena odjemalce poznamo tudi delitev, kot jo v Sloveniji podaja energetski zakon [2] in je odvisna od odjemalčeve priključne moči. Ločimo upravičene odjemalce (UO) in tarifne odjemalce. Meja med njimi je odvisna od stopnje odprtosti trga z električno energijo. V prvem koraku je mejo določala priključna moč 41 kW; UO-ji so bili odjemalci s priključno močjo nad 41 kW, medtem ko so bili ostali odjemalci tarifni odjemalci. V drugem koraku so s povečanjem stopnje odprtosti trga postali UO-ji tudi tisti odjemalci, ki so predstavljali preostali odjem. Tarifni odjemalci ostajajo samo gospodinjstva. Razlika med tipoma odjemalcev je, da UO-ji lahko prosto nastopajo na trgu in si sami izbirajo svojega dobavitelja električne energije, medtem ko si tarifni odjemalci dobavitelja ne morejo izbirati.

### 2.3.2. Sintetični postopek

Podlaga za sintetični postopek so tipični ali nadomestni obremenitveni diagrami (NOD-i), ki predstavljajo odjemalce s podobno obliko OD-ja. Distribucijski operater v skladu z vnaprej določenim postopkom dodeli NOD vsakemu odjemalcu ali skupini odjemalcev brez sprotnih meritev. Postopki dodeljevanja so lahko različni. Enega izmed njih predlagamo v nadaljevanju disertacije in sicer v razdelku 3.3. NOD-i, ki so normirani, npr. v Nemčiji so vsi normirani na letno porabo električne energije 1000 kWh, predstavljajo spreminjanje obremenitve določene skupine odjemalcev. Postopek za enega od dobaviteljev, ki dobavlja električno energijo odjemalcem, katerim je bil dodeljen različen NOD, npr. trgovinam in odjemalcem s področja zdravstva, prikazuje Sl. 2.3. Vsak  $i$ -ti interval  $k$ -tega NOD-a skaliramo, pomnožimo ga s faktorjem  $c_k$ , odvisnem od porabljene energije odjemalca, po (2.4).

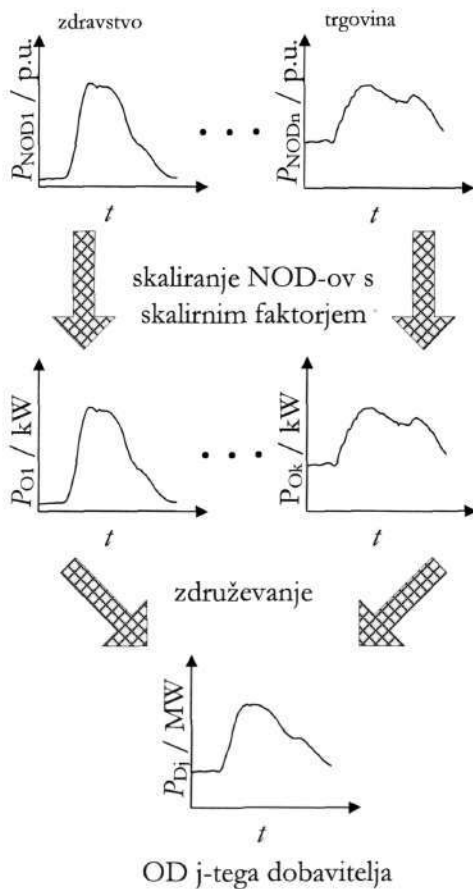
$$P_{Ok,i} = P_{NODk,i} \cdot c_k (W) \quad (2.4)$$

Rezultat skaliranja je ocenjen OD,  $P_{Ok}$ , odjemalca ali skupine odjemalcev, če NOD skaliramo s skupno energijo odjemalcev, ki jim je dodeljen isti NOD. V naslednjem trenutku distribucijski operater za  $i$ -ti interval sešteje ocenjene OD-je odjemalcev in določi dobaviteljev  $P_D$  OD, kot sledi:

$$P_{Dj,i} = \sum_{k=1}^N P_{Ok,i}, \quad (2.5)$$

kjer je  $N$  število različnih NOD-ov dodeljenih odjemalcem, ki pogodbeno sodelujejo z  $j$ -tim dobaviteljem. Vsak odjemalec dobi samo en NOD. Tako dobljen OD distribucijski operater

izroči dobavitelju in predstavlja njegov dejanski obremenitveni diagram. Če odjemalec nastopa na trgu samostojno, distribucijski operater določi njegov OD po enačbi (2.4). S seštevanjem OD-jev vseh dobaviteljev določi operater ocenjeno preostalo krivuljo.



Sl. 2.3: Sintetični postopek.

### 2.3.3. Analitični postopek

Analitični postopek ima za podlago celotni OD distribucijskega omrežja, ki ga vodi distribucijski operater. Operater v prvem koraku določi preostalo krivuljo v skladu z enačbo (2.3), t.j. odšteje ocenjene izgube nastale na distribucijskem omrežju in moči vseh izmerjenih OD-jev. V naslednjem koraku razdeli preostalo krivuljo na posamezne odjemalce ali njihove dobavitelje s tem, da jo pomnoži z ustreznim deležem. Razdelitev za  $i$ -ti interval podaja naslednja enačba:

$$P_{Dk,i} = P_{PK,i} \cdot z_k, \quad (2.6)$$

kjer je:

$P_{PK}$  – preostala krivulja,

$z_k$  – delež  $k$ -tega odjemalca ali dobavitelja.

Delež izračunamo na dva načina, odvisno od tega ali želimo uporabiti preprost analitični postopek ali razširjen analitični postopek. Delež  $z_k$  za preprost analitični postopek zapišemo kot sledi:

$$z_k = \frac{\sum_{i \in \Gamma} W_{OLi}}{\sum_{j \in \Pi} W_{OLj}}, \quad (2.7)$$

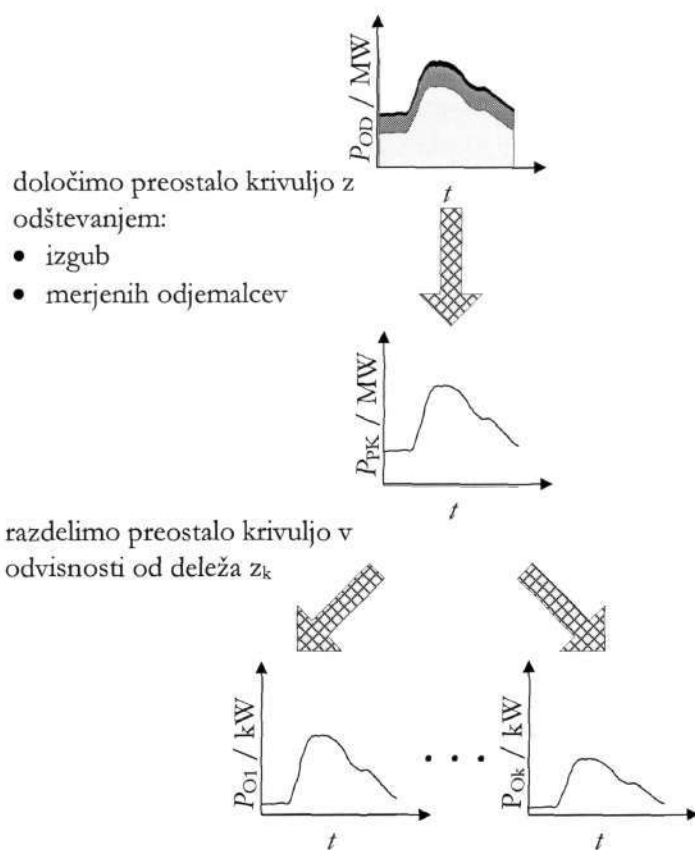
kjer je:

$W_{OL}$  – letna energija odjemalca,

$\Pi$  – množica vseh odjemalcev, ki se napajajo iz distribucijskega omrežja,

$\Gamma$  – množica odjemalcev, katerim energijo dobavlja k-ti dobavitelj, lahko je to tudi posamezen odjemalec.

Distribucijski operater določi delež dobavitelja v odvisnosti od razmerja med seštevkom energij vseh odjemalcev, katerim dobavitelj energijo dobavlja in celotno energijo dobavljeno odjemalcem na področju distribucijskega omrežja. Za dobavitelje je to razmerje nespremenjeno, dokler en ali več odjemalcev ne zamenjajo dobavitelja. Zato imajo OD-ji vseh dobaviteljev in odjemalcev enako obliko, ki je enaka obliki preostale krivulje. To se vidi na Sl. 2.4, ki prikazuje analitični postopek določanje OD-jev odjemalcev brez sprotnih meritev.



Sl. 2.4: Analitični postopek.



Preprost analitični postopek ne upošteva različnih dobaviteljevih skupin odjemalcev, saj v splošnem množico odjemalcev  $\Gamma$  sestavljajo različni tipi odjemalcev, ki imajo različne OD-je. Da bi se izognil ali ublažil glavno pomanjkljivost analitičnega postopka, ki predvideva enako obliko OD-jev, distribucijski operater uporabi razširjen analitični postopek. Postopek predstavlja kombinacijo obeh; analitičnega in sintetičnega postopka. V prvem koraku distribucijski operater izvede sintetični postopek ter določi OD-je vsem odjemalcem, dobaviteljem in skupni OD vseh nemerjenih odjemalcev. V naslednjem koraku določi za  $i$ -ti interval delež  $k$ -tega dobavitelja za  $m$ -ti tip odjemalca, kot sledi

$$z_{k,m,i} = \frac{P_{Okm,i}}{\sum_{k \in \Xi} P_{Dk,j}}, \quad (2.8)$$

kjer je:

$P_{Okm}$  – ocenjen OD  $m$ -tega tipa odjemalcev, ki pogodbeno sodelujejo s  $k$ -tim dobaviteljem,

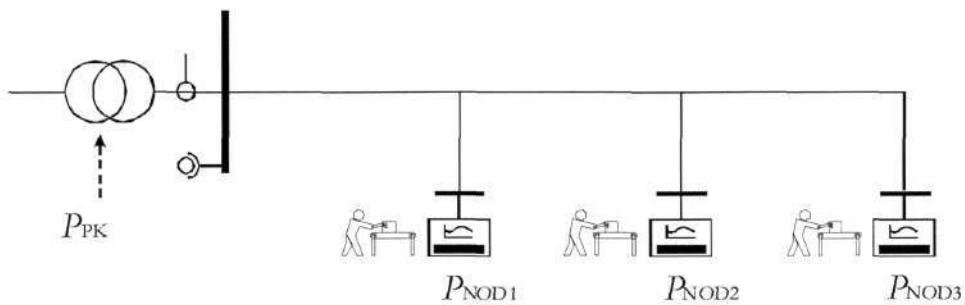
$P_{Dj}$  – ocenjen OD  $j$ -tega dobavitelja,

$\Xi$  – število dobaviteljev na območju distribucijskega omrežja.

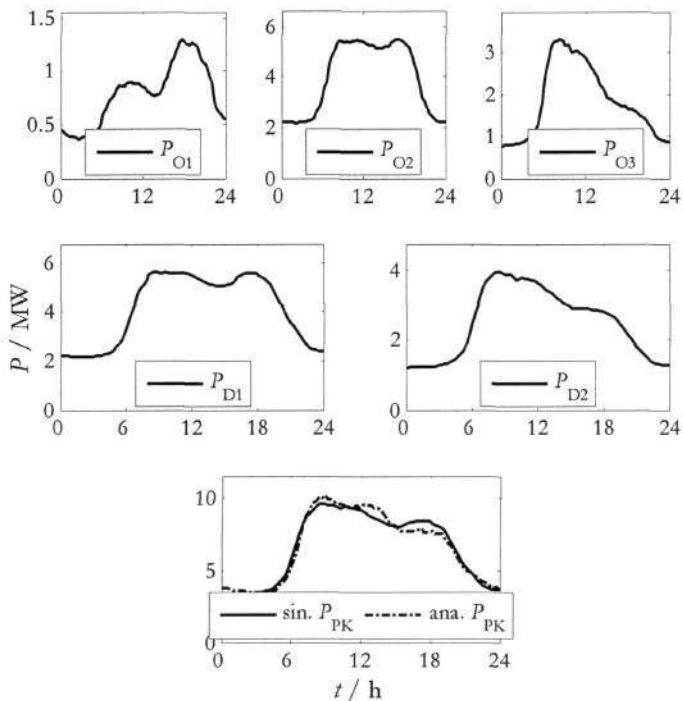
Delež podaja razmerje med sintetičnim OD-jem odjemalcev istega tipa in sintetičnim OD-jem celotnega distribucijskega omrežja. V zadnjem koraku distribucijski operater razdeli preostalo krivuljo na dobaviteljeve OD-je v skladu z njihovimi deleži (2.8), ki so funkcije časa in se spreminjajo neodvisno od sestave odjemalcev, katerim dobavitelj dobavlja energijo. Zaradi tega je oblika dobaviteljevega OD-ja različna od oblike preostale krivulje, kar je posledica različnih oblik NOD-ov, dodeljenih različnim tipom odjemalcev.

## 2.4. Primerjava postopkov

Vsak izmed predstavljenih postopkov ima svoje prednosti in slabosti. Nekatere med njimi smo omenili že prej. Podrobneje si jih bomo ogledali na preprostem fiktivnem primeru dela distribucijskega omrežja, ki ga prikazuje Sl. 2.5. Ne glede na omejenost primera na del omrežja ostajajo lastnosti postopkov nespremenjeni. Z njimi na enak način obravnavamo celotno distribucijsko omrežje ali več distribucijskih omrežji skupaj. Na sliki vidimo RTP, kjer poznamo preostalo krivuljo  $P_{PK}$  in tri skupine odjemalcev, katerim dodelimo različni NOD. Odjemalcem dobavljata energijo dva dobavitelja. Prvi dobavitelj dobavlja energijo odjemalcem, ki imajo NOD1, 70 % odjemalcem (energijsko gledano in ne številčno), ki imajo NOD2 in 30 % odjemalcem, ki imajo NOD3. Ostalim odjemalcem dobavlja energijo drugi dobavitelj. Njuna OD-ja smo določili v skladu s postopkoma predstavljenima v predhodnima razdelkoma. Rezultate sintetičnega postopka podaja Sl. 2.6, kjer smo prikazali rekonstruirane OD-je  $P_O$  vseh treh skupin odjemalcev, dobaviteljeva OD-ja  $P_D$  ter sintetično preostalo krivuljo  $P_{PKS}$  in analitično preostalo krivuljo  $P_{PKA}$ . Iz slike vidimo, da se obliki dobaviteljevih OD-jev razlikujeta zaradi različnih oblik NOD-ov in različnega deleža odjemalcev, katerim dobavljata energijo. Njuna vsota predstavlja sintetično preostalo krivuljo, katere oblika se prav tako razlikuje od oblike dobaviteljevih OD-jev.



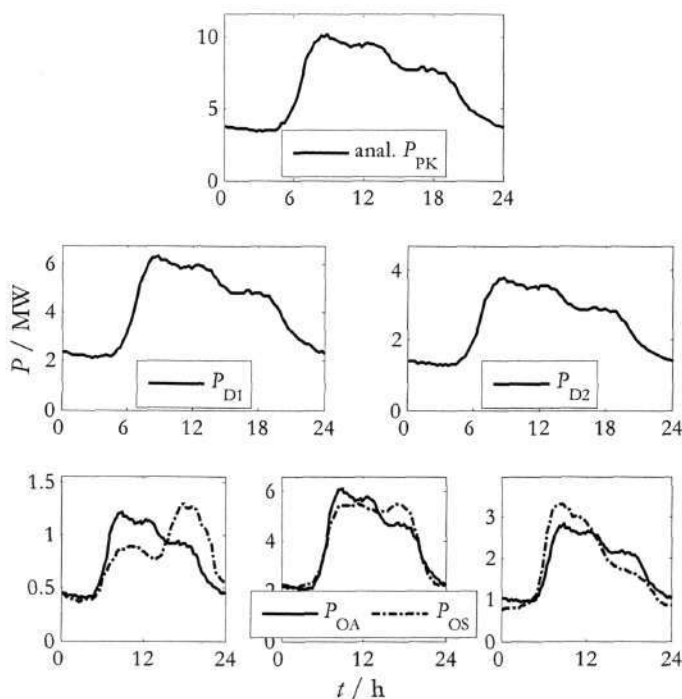
Sl. 2.5: Preprost model distribucijskega omrežja.



Sl. 2.6: Rezultati sintetičnega postopka.

Sl. 2.6 prikazuje tudi primerjavo med sintetično in analitično preostalo krivuljo. Vidimo, da se obliki krivulj razlikujeta, medtem ko je površina pod njima enaka; dobavljena energija odjemalcem je enaka ne glede na izbrano metodo določanja OD-jev. Ta pogoj je univerzalen in mora biti izpolnjen ne glede na izbran model ali postopek določanja OD-jev dobaviteljev ali odjemalcev.

Rezultate analitičnega postopka prikazuje Sl. 2.7, kjer smo prikazali preostalo krivuljo  $P_{PK}$ , OD-ja obeh dobaviteljev  $P_{D1}$  in  $P_{D2}$  ter primerjavo OD-jev posamezne skupine odjemalcev, izračunanih z analitičnim  $P_{OA}$  in s sintetičnim postopkom  $P_{OS}$ . Iz slike vidimo, da imata OD-ja obeh dobaviteljev enako obliko, tj. obliko preostale krivulje. Razlikujeta se samo po površini pod krivuljo, ki predstavlja porabljeno ali dobavljeno energijo. Primerjava analitičnih in sintetičnih OD-jev posamezne skupine odjemalcev kaže, da se v nekaterih primerih prvi močno razlikujejo od slednjih. Sintetični OD-ji bolje opisujejo potek porabe električne energije posamezne skupine odjemalcev.

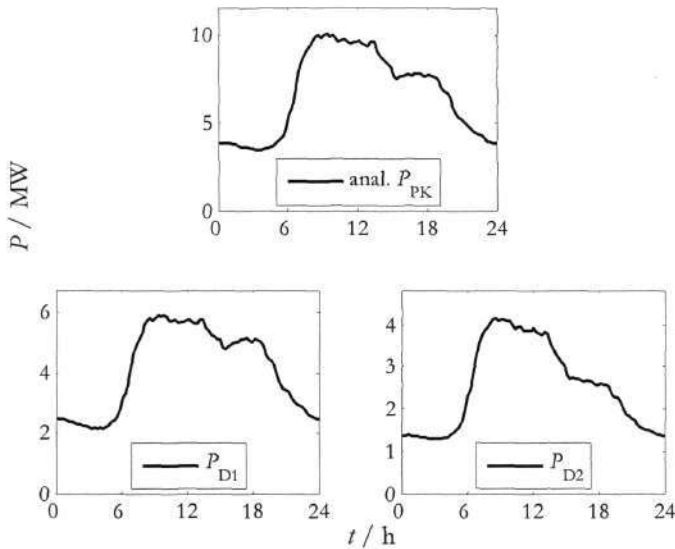


Sl. 2.7: Rezultati analitičnega postopka.

Najopaznejša razlika med OD-ji odjemalcev je pri prvi skupini, medtem ko je najmanjša pri drugi skupini. Razlog je v količini energije, ki jo posamezna skupina odjemalcev porabi. Več kot je porabi, večji je vpliv skupine odjemalcev na obliko preostale krivulje in s tem se njen OD manj razlikuje od oblike preostale krivulje. Da bi omilil ali zmanjšal pomanjkljivost analitičnega postopka, distribucijski operater uporabi razširjen analitični postopek. Rezultate prikazuje Sl. 2.8. Iz slike vidimo, da sta obliki dobaviteljevih OD-jev različni od oblike preostale krivulje zaradi spreminjajočega se deleža posameznega dobavitelja.

Na podlagi prikazanih primerov lahko zaključimo:

- Sintetični postopek je lažje uporabljati, ker individualno lahko obravnavamo odjemalce, ki zamenjajo dobavitelja. Odjemalčevi OD-ji so "podobni" dejanskim OD-jem, kar se vidi v [8]. Vendar ima postopek pomanjkljivost v tem, da mora distribucijski operater obvladovati tveganje za neravnotežje med porabo celotnega distribucijskega omrežja, določeno s sintetičnim in analitičnim postopkom, glej Sl. 2.6. Z analitičnim postopkom določimo preostalo krivuljo natančneje.
- Rezultat analitičnega postopka so OD-ji, ki imajo enako obliko, kar v realnosti ne drži. Odjemalci imajo različne oblike OD-jev. V tem primeru prevzamejo tveganje za neravnotežje porabe dobavitelji ali odjemalci. Pri skupini odjemalcev z NOD1 na Sl. 2.5, katerim dobavlja energijo samo prvi dobavitelj, močno spremeni način in količino porabe energije, kar občutno sprememni preostale krivulje. Drugi dobavitelj nosi dodatno tveganje za plačilo odstopanj, ker se je posledično spremenil tudi njegov OD.



Sl. 2.8: Rezultati razširjenega analitičnega postopka.

Glavno pomanjkljivost analitičnega postopka omili razširjen analitični postopek, ki predvideva spreminjajoči delež (2.8) dobavitelja. Delež se določi na podlagi sintetičnega postopka, ki v nekaterih primerih vodi k relativno velikim odstopanjem sintetične preostale krivulje od analitične. Odstopanja so odvisna od natančnosti NOD-ov. Zato smo iskali dodatne rešitve, ki bi omogočale boljšo razdelitev analitično določene preostale krivulje. Pri tem upoštevamo, da imajo odjemalci različne oblike OD-jev, ki jih določajo NOD-i. Ti dve lastnosti sta ključnega pomena za oblikovanje novega modela in se bomo k njima vrnil v poglavju 4, kjer bomo podrobneje predstavili predlagan model določanja porabe odjemalcev ali dobaviteljev.

## 2.5. Pregled literature na področju oblikovanja NOD-ov

Eno od osnovnih podlag predlaganega modela določanja porabe v 4. poglavju, so NOD-i, ki smo jih morali določiti. Postopek njihovega določanja prav tako predstavlja pomemben del disertacije, saj smo v tem okviru predlagali nekatere nove ideje, ki predstavljajo izvirne prispevke [8], [9]. Za boljšo osvetlitev problematike določanja NOD-ov podajamo v tem razdelku pregled literature na tem področju, kjer bomo osvetlili nekatere pomanjkljivosti dosedanjih metod.

### 2.5.1. Modeliranje porabe gospodinjstev z uporabo verjetnostnih funkcij

Poznavanje obremenitve je zelo pomembno za vrednotenja ukrepov vodenja porabe, ki jih izvajajo posamezna distribucijska podjetja. Pri tem je bilo dosti pozornosti posvečeno gospodinjstvom, ki predstavljajo občuten delež v celotni porabi električne energije. Zaradi tega so raziskovalci na tem področju poizkušali razviti ustrezne modele, ki bi omogočali analizo in ovrednotenje vpliva ukrepov vodenja porabe na porabo gospodinjstev. Enega izmed modelov podaja literatura [10], kjer avtorji modeliranja porabe odjemalcev gospodinjstevskega tipa na način "od končnih porabnikov do OD-ja posameznega območja ali

izvoda" (ang. bottom-up). Takšna metodologija je zelo kompleksna, ker temelji na socialnih dejavnikih, povezanih z načinom življenja, ki so zelo subjektivni in jih je težko določiti z zadovoljivo natančnostjo. Kompleksnost naj bi odtehtala njena natančnost. Model deluje na principu združevanja porabe v dveh korakih. V prvem koraku določijo OD-je posameznih gospodinjstev, ki jih v naslednjem koraku združijo in s tem določijo OD celotnega območja.

Enote modela predstavljajo električne naprave, ki jih gospodinjstva uporabljajo in število gospodinjstev članov. Povezavo med njimi so modelirali z uporabo verjetnostnih funkcij, ki odražajo povprečno težnjo članov gospodinjstva k uporabi določene naprave. Tipične funkcije obnašanja gospodinjstev so: število članov družine in njihova prisotnost v stanovanju v določenem delu dneva, domače aktivnosti, uporaba gospodinjstev naprav tekom dneva, nagnjenost članov gospodinjstva k opravljanju določene aktivnosti, število in vrsta električnih naprav, ki jih gospodinjstvo ima, ipd.

Omenjeni način določanja OD-jev zahteva ogromno podatkov, ki jih je včasih zelo težko ali celo nemogoče pridobiti. Če jih uspemo pridobiti, je njihova natančnost zelo vprašljiva. Avtorji so izvedli poglobljeno analizo uporabe električnih naprav in analizo obnašanja samo gospodinjstev odjemalcev, zato bi jo zelo težko posplošili na ostale tipe odjemalcev, kot so industrijski odjemalci. Model je primeren za analizo vpliva ukrepov vodenja porabe in napovedovanje porabe gospodinjstev in ne za uporabo v postopkih namenjenih "razdeljevanju" že dobavljene električne energije med posamezne odjemalce ali dobavitelje brez sprotnih meritev. Zaradi tega tem metodam ne bomo posvečali dodatne pozornosti. Omenili smo jih kot zanimiv in podroben, vendar zelo kompleksen način modeliranja porabe gospodinjstev.

### 2.5.2. Določanje OD-jev na podlagi meritev porabe odjemalcev

Primernejši način od modeliranja porabe z uporabo verjetnostnih funkcij je določanje OD-jev in njihovih karakteristik z uporabo meritev porabe posameznih odjemalcev. V tem okviru je mnogo držav oblikovalo in izpeljalo projekte na področju "raziskovanja" porabe (ang. load research projects). Projekti so se v tujini izvajali že v 70-tih in 80-tih letih prejšnjega stoletja. Cilj je bil zbrati in analizirati podatke o obremenitvah iz različnih delov distribucijskega omrežja, ki bi omogočali distribucijskim podjetjem učinkoviteje voditi omrežje, npr. določati izgube, sekcionirati omrežje, ocenjevati stanje v distribucijskem omrežju, načrtovati razvoj omrežja, izvajati ukrepe vodenja porabe. Literatura [11] navaja države, ki so takšne projekte izpeljale, npr. Nemčija, Danska, Španija, Francija, Norveška, Švedska, Finska. V preteklosti je shranjevanje in obvladovanje velikih količin podatkov predstavljalo glavno oviro pri izvajanju omenjenih projektov. S prihodom sodobne merilne in komunikacijske opreme, ki omogoča daljinsko odčitavanje in prenos podatkov ter zmogljivih računalnikov, ki omogočajo enostavno shranjevanje in obdelavo velikih količin podatkov, so "load research" programi postali enostavneje obvladljivi. Njihov cilj je analiza in uporaba pridobljenih podatkov. Ne glede na razpoložljivost sodobne opreme so takšni programi še vedno časovno zahtevni. Vzrok je v podatkih, ki jih ni mogoče pridobiti čez noč. Takšni programi lahko trajajo več let



ali celo desetletje, npr. na Finskem so z "load research" programom, ki je trajal več kot deset let, začeli leta 1983 [11]. Rezultat je 42 skupin odjemalcev, katerim so določili OD-je.

V strokovnih revijah težko najdemo reference iz začetnega obdobja izvajanja "load research" programov, ki opisujejo rezultate in postopke uporabljene pri določanju OD-jev in njihovem razvrščanju v podobne skupine. Avtorji so pretežno raziskovali načine in postopke modeliranja predvsem gospodinjskih odjemalcev [14]. Razlog je bil v velikem številu zahtev, pridobljenih z izkušnjami, katere je morala takšna razvrstitev upoštevati. Zaradi tega je veljala teza, da je razvrščanje odjemalcev bolj umetnost kakor računanje in ga najbolje opravijo izkušeni analitiki [11]. Glavne lastnosti, katerim mora zadostiti razvrstitev, so:

- varianca znotraj posamezne skupine odjemalcev mora biti čim manjša,
- število skupin ne sme biti preveliko, da lažje obvladujemo veliko število podatkov,
- skupine morajo biti reprezentativne. Lastnosti in oblike OD-jev se ne smejo prekrivati,
- skupine moramo preprosto povezati s podatki, ki jih imajo distribucijska podjetja,
- vse skupine naj bi predstavljale čim večjo populacijo odjemalcev; v idealnem primeru vse odjemalce.

Za razliko od začetnega obdobja so avtorji v drugi polovici 90-tih letih prejšnjega stoletja začeli intenzivneje objavljati članke na temo določanja OD-jev odjemalcev. Literatura [15]-[18] podaja rezultate določanja in uporabe OD-jev na Tajvanu, kjer so avtorji odjemalce razdelili v devet skupin glede na obstoječe tarifne skupine. V analizo so zajeli vse tipe odjemalcev; od gospodinjstev z enofaznim priključkom, ki predstavljajo 85 % vseh odjemalcev, do industrijskih odjemalcev na visoki napetosti (69 kV ali 161 kV), ki porabijo več kot polovico vse električne energije v sistemu, čeprav jih je samo 40. Avtorji so nazorno pokazali način izbire ustreznega števila merjenih odjemalcev znotraj posamezne skupine, ki temelji na standardni deviaciji porabe električne energije. Večja kot je deviacija posamezne skupine, več odjemalcem morajo meriti OD-je, npr. merili so samo majhen delež gospodinjstev, ker imajo zelo skladne OD-je, medtem ko so zaradi raznolikosti OD-je merili vse odjemalce, priključene na najvišjo napetost. V ta namen so uporabili teorijo plastnega vzorčenja (ang. Stratified Sampling Theory), ki daje dobre rezultate. Na podlagi pridobljenih podatkov so s statističnimi metodami določili najverjetnejše OD-je za posamezno skupino, ki jih v nadaljevanju združijo, da dobijo OD-je določenega območja. Primerjava rekonstruiranih OD-jev z merjenimi kaže ustrezno natančnost predlagane metode.

Literatura [19], [20] opisuje določanje in primer uporabe OD-jev v Braziliji. Avtorji so odjemalce, priključene na nizko napetost, razdelili na tri najpomembnejše glavne skupine, ki predstavljajo gospodinjstva, komercialo in industrijske odjemalce. Vsako izmed skupin so razdelili na podskupine glede na odjemalčevo dejavnost in porabljeno mesečno energijo. Na primer, gospodinjske odjemalce so razdelili na deset skupin glede na porabljeno mesečno energijo v razponu, kakor ga prikazuje Tab. 2.1. Poleg razdelitve v skupine prikazuje tabela tudi frekvenčno porazdelitev odjemalcev v vsaki skupini, tj. procentualni delež skupine

odjemalcev v odvisnosti od vseh gospodinjstkih odjemalcev. Podatki veljajo za eno izmed brazilskih mest.

Tab. 2.1: Razdelitev gospodinjstkih odjemalcev glede na porabljeno mesečno energijo v Braziliji.

skupina	1	2	3	4	5	6	7	8	9	10
razpon (kWh/mesec)	0-50	51- 100	101- 150	151- 200	201- 250	251- 300	301- 400	401- 500	501- 1000	nad 1001
delež (%)	10	12	17	18	12	12	11	5	2	1

Meritve so zajele vse skupine gospodinjstkih odjemalcev, 47 glavnih komercialnih dejavnosti, 26 glavnih industrijskih dejavnosti ter nekatere transformatorske postaje in izvode. Rezultat statistične analize so bili povprečni OD-ji, ki so bi predhodno normirani z bazno močjo  $P_B$  in krivulja, ki podaja standardno deviacijo OD-jev. Bazno moč so določili kot sledi:

$$P_B = \frac{1}{24} \int_{24 \text{ ur}} \mathbf{M}(t) dt = \frac{\text{mesečna energija (kWh)}}{24 \cdot \text{št. dni v mesecu}}, \quad (2.9)$$

kjer  $P_B$  predstavlja povprečno dnevno moč odjemalca.

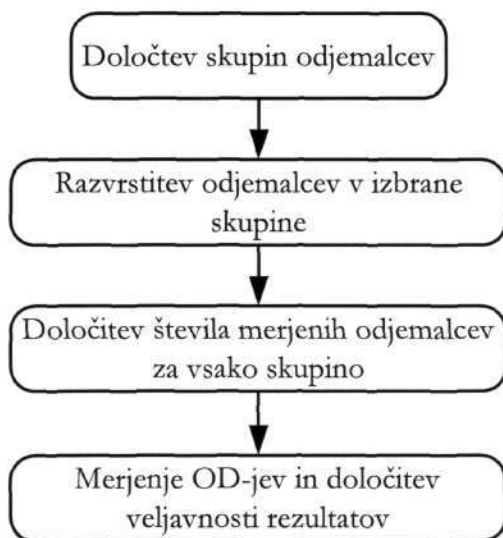
Začetno razdelitev gospodinjstkih odjemalcev so združili in določili štiri končne skupine v razponu 0-50, 51-200, 201-300, 301-400 kWh/mesec. Dobljeni OD-ji izkazujejo veliko standardno deviacijo, ki je istega reda kot povprečen OD. Začetnih 47 podskupin komercialnih odjemalcev so združili v štiri reprezentativne skupine glede na podobnost OD-jev podskupin. Industrijskih odjemalcev niso dodatno združevali. Rezultat je 26 različnih OD-jev; za vsako analizirano podskupino en OD.

Dobljene OD-je so uporabili za oceno OD-jev v različnih točkah omrežja in jih primerjali z merjenimi OD-ji transformatorjev, ki napajajo gospodinjstke in komercialne odjemalce. Rezultati kažejo primerljivost ocenjenih in izmerjenih OD-jev. Poleg ocenjevanja obremenitve v točkah omrežja predlagajo avtorji uporabo dobljenih OD-jev za ocenjevanje življenjske dobe transformatorjev v distribucijskem omrežju [20]. Najprej določijo testno množico transformatorjev, katerim natančno izračunajo predvideno življenjsko dobo na podlagi izmerjenih OD-jev in modelirane zunanje temperature. V naslednjem koraku ocenijo obremenitve ostalih transformatorjev in jih primerjajo z obremenitvijo testnih transformatorjev. Rezultat primerjave je določitev predvidene življenjske dobe transformatorja, tj. transformatorja, ki imata najbolj podobno obremenitev, imata enako predvideno življenjsko dobo.

Predlagan postopek je hiter in dobrodošel za ocenjevanje življenjske dobe velikemu številu transformatorjev, vendar smo do njega vseeno zadržani. Primerjava zgolj obremenitev ni zadostna. Primerjati je treba tudi zunanje temperature, kjer transformatorji obratujejo, ker ima zunanja temperatura pomemben vpliv na njihovo življenjsko dobo. Dodaten vpliv imajo

okvare v omrežju. Kratkostični tokovi povzročajo velike sile v transformatorju, ki prav tako negativno vplivajo na življenjsko dobo transformatorja. Zato menimo, da je natančnost določanja predvidene življenjske dobe samo na podlagi primerjave obremenitev transformatorja vprašljivo.

Postopek določanja OD-jev na podlagi vnaprej določenih skupin odjemalcev predlagajo avtorji v [21]. Sestavljajo ga štiri koraki, kot prikazuje Sl. 2.9:



Sl. 2.9: Predlagan postopek določanja OD-jev.

Avtorji predlagajo določanje skupin odjemalcev na podlagi informacij, ki jih posedujejo distribucijska podjetja. Ker so posamezne skupine zelo obsežne, ni ekonomično vsem odjemalcem znotraj skupine meriti porabo. Stroški izvedbe meritev bi bili preveliki. Zato avtorji uporabijo ustrezno statistično metodo, ki omogoča določitev primernejšega, manjšega vzorca odjemalcev znotraj posamezne skupine. V ta namen predlagajo, podobno kot avtorji v [16], teorijo "plastnega vzorčenja". Z uporabo predlagane metode in razpoložljivih sredstev, namenjenih merjenju porabe določijo razdelitev merilnih garnitur za merjenje obremenitve med posamezne skupine odjemalcev.

V [22] najdemo opis določanja OD-jev, kakor jih uporabljajo v Veliki Britaniji, kjer so odjemalce pod 100 kW priključne moči razdelili v osem tipičnih skupin. Odjemalci na 100 kW priključne moči imajo sprotne meritve, ki beležijo porabo v pol urnih intervalih. Za vsako izmed vnaprej formiranih skupin določijo OD-je na podlagi meritev izpred dveh let, kar predstavlja po našem mnenju zelo dolgo vmesno dobo. Izdelava OD-jev temelji na dolgoletni bazi podatkov, pridobljeni v okviru "load research" programa, katero so razširili na približno 2500 merjenih odjemalcev. Pri tem so posvetili posebno pozornost negospodinjiskim odjemalcem pod 100 kW.

Literatura [23]-[26] obravnava določanje OD-jev posameznih odjemalcev s pomočjo razvrščanja – uporabe metod in tehnik razpoznavanja vzorcev, medtem ko so metode razvrščanja uporabljene za primerjavo OD-jev posameznih izvodov iz RTP-jev v [27]. Slednji



primer za nas ni toliko zanimiv, ker obravnava OD-je izvodov in ne posameznih odjemalcev. Zaradi tega ga ne bomo podrobneje predstavljali. Avtorji v [23] obravnavajo razvrščanje OD-jev z uporabo indeksov, izračunanih iz izmerjenega OD-ja. V ta namen predlagajo šest enostavnih koeficientov (2.10), tj. koeficient ne-monotonosti  $f_{G1}$ , koeficient "polnosti" (ang. fill-up)  $f_{G2}$ , koeficient modulacije vršnega, ne-vršnega in pasovnega dela OD-ja ( $MC_{ph}$ ,  $MC_{oph}$  in  $MC_{23-6}$ ) ter koeficient modulacije med pasovnim in trapeznim delom OD-ja  $f_{G3}$  in sestavljen (2.11) indeks  $I_S$ .

$$\begin{aligned}
 f_{G1} &= \frac{P_{\min}}{P_{\max}} \\
 f_{G2} &= \frac{P_{\text{med}}}{P_{\max}} \\
 MC_{ph} &= \frac{P_{\text{med ph}}}{P_{\text{med}}} \\
 MC_{oph} &= \frac{P_{\text{med oph}}}{P_{\text{med}}} \\
 MC_{23-6} &= \frac{P_{\text{med 23-6}}}{P_{\text{med}}} \\
 f_{G3} &= \frac{P_{\text{med 23-6}}}{P_{\text{med 6-23}}}
 \end{aligned} \tag{2.10}$$

$$I_S = 0.4 \cdot f_{G1} + 0.4 \cdot f_{G2} + 0.2 \cdot MC_{23-6}, \tag{2.11}$$

kjer so:

$P_{\min}$ ,  $P_{\max}$ ,  $P_{\text{med}}$  – minimalna, maksimalna in povprečna moč določena z razmerjem med mesečno porabo in mesečnim številom ur,  $P_{\text{med ph}}$ ,  $P_{\text{med oph}}$  povprečna moč koničnega in nekoničnega dela OD-ja,  $P_{\text{med 23-6}}$ ,  $P_{\text{med 6-23}}$  povprečna moč pasovnega in trapeznega dela OD-ja posameznega odjemalca in  $I_S$  sestavljeni indeks.

Primernost indeksov so testirali na 427 odjemalcih razdeljenih v 14 skupin s kriterijem oblikovanim na podlagi geometrične razdalje, ki ga podaja enačba (2.12), kjer ind predstavlja posamezen indeks in  $x_N$  posamezen normiran obremenitveni diagram.

$$GD(\text{ind}) = \sum_{i=1}^{24} (\text{ind} - x_{Ni})^2 \tag{2.12}$$

Rezultati kažejo, da je za razvrščanje odjemalcev v nove skupine na podlagi predlaganih indeksov najprimernejši sestavljeni indeks. Takšen rezultat je do neke mere pričakovan, saj sestavljen indeks "nosi" več informacij o obravnavanem OD-ju kakor enostavni indeksi. Tako je razvrščanje OD-jev na njegovi podlagi primernejše. V [24] in [25] so avtorji razvrščali normirane OD-je, opisane z nekaterimi izmed indeksov, ki jih podaja enačba (2.10). Pri tem so uporabljali metodo avtomatskega hierarhičnega rojenja, ki temelji na prestavljanju OD-jev iz ene skupine v drugo, v odvisnost od razdalje med indeksi in ne zahteva vnaprejšnjega določanja števila skupin ali rojev. Za preverjanje primernosti indeksov, namenjenih

razvrščanju, so predlagali dva nova kriterija, ki sta v bistvu izpeljanki ali poenostavitvi že obstoječih kriterijev. Omenjeni kriteriji so na področju razpoznavanja vzorcev poimenovani kot mere veljavnosti rojenja. O njih bomo več govorili v razdelku 3.2.1.1. Na podlagi oblikovanih skupin in tipičnih OD-jev avtorji želijo določiti novo tarifno politiko v želji, da bi dobavitelji električne energije maksimizirali prihodke. V tem okviru tudi predlagajo, da bi odjemalcem brez sprotnih meritev določali OD-je na podlagi predhodnega merjenja njihovih OD-jev, katero bi trajalo od 3 do 6 mesecev. Tako bi določili, kateri skupini odjemalec pripada in s tem kakšen OD ima. Takšen postopek je pregleden in po našem mnenju zelo učinkovit, vendar je še vedno časovno zelo potraten in drag, če imamo veliko število odjemalcev. V literaturi [26] najdemo primerjavo hierarhičnih metod in samo-organizirajočih omrežij za razvrščanje OD-jev. Avtorji so, za razliko od prejšnjih dveh primeru, razvrščali celotne normirane OD-je in ne indekse izračunane iz njih.

V Sloveniji v preteklosti ni bilo dosti narejenega na področju določanja OD-jev, saj "load research" programov ni bilo. Nekatere izmed prvih začetkov na področju določanja in analiziranja OD-je podaja literatura [6], [28] - [30]. Prvi dve referenci obravnavata lastnosti OD-jev, medtem ko v zadnjih dveh najdemo prvi poizkus oblikovanja NOD-ov v Sloveniji. Avtorji so za razvrščanje OD-jev uporabili hierarhično metodo, katero bomo obravnavali tudi v tem delu. Omenjenim začetkom je sledilo intenzivnejše delo na tem področju, katero je tudi pripeljalo do oblikovanja pričujoče disertacije.

Pregleda problematike določanja OD-jev na podlagi meritev porabe posameznih odjemalcev lahko sklenemo z naslednjimi ugotovitvami:

- Prvi način temelji na znanih skupinah odjemalcev ter večletnih meritvah in analizah odjema, npr. avtorji v [15] poročajo o štiriletnem obdobju merjenja in analiziranja porabe. Na podlagi izpeljanih "load research" programih so oblikovali ustrezne skupine odjemalcev, katerim so kasneje določili OD-je. Avtorji v večini ne podajajo kriterijev, kako so posamezne skupine oblikovali. Enostavno jih privzamejo iz tarifne razdelitve odjemalcev.
- Drugi način obravnava določanje OD-jev z uporabo metod razpoznavanja vzorcev. V večini primerov avtorji poročajo o primerjavi uporabljenih metod in oblikovanju diagramov, vendar nikjer ne podajo postopka, kako OD-je dodeliti posameznim skupinam odjemalcev ter katerim odjemalcem je treba meriti porabo. Omenjena naloga je ena izmed pomembnejših pri izbiri načina izdelave NOD-ov.

V disertaciji smo poizkušali poiskati kompromis med obema načinoma določanja NOD-ov, kar bomo podrobneje predstavili v nadaljevanju.

V tem razdelku smo z "določanjem obremenitvenih diagramov" vedno imeli v mislih določanje nadomestnih obremenitvenih diagramov, NOD-ov. Omenjeno formulacijo smo povzeli iz literature, kje avtorji uporabljajo naslednje izraze za NOD-e: obremenitvene krivulje (ang. load curves) ali obremenitveni profili (ang. load profiles).

### 3. Določanje NOD-ov

Uvodoma smo zapisali, da ideja OD-jev ni nova, vendar je nova njihova uporaba pri spodbujanju trgovanja z električno energijo. Odjemalcem brez sprotnih meritev in njihovim dobaviteljem, ki nastopajo na trgu, omogočajo ocenjevanje obremenitve. Ker se je v Sloveniji trg električne energije odprl najprej samo za upravičene odjemalce (UO-je), bomo v tem poglavju predstavili samo postopek določanja NOD-ov UO-jev. Podlago predlaganega in v nadaljevanju podrobneje predstavljenega postopka predstavljajo OD-ji posameznih odjemalcev, sestavljenih iz meritev moči kot povprečne vrednosti energije v posameznih časovnih intervalih. V slovenskem EES-u je za distribucijska podjetja predpisan 15 minutni časovni interval merjenja porabe. Tako OD  $\mathbf{x}_k$  sestavlja 96 četrt-urnih vrednosti:

$$\mathbf{x}_k = [x_{k,i}] , \quad i = 1, \dots, 96 . \quad (3.1)$$

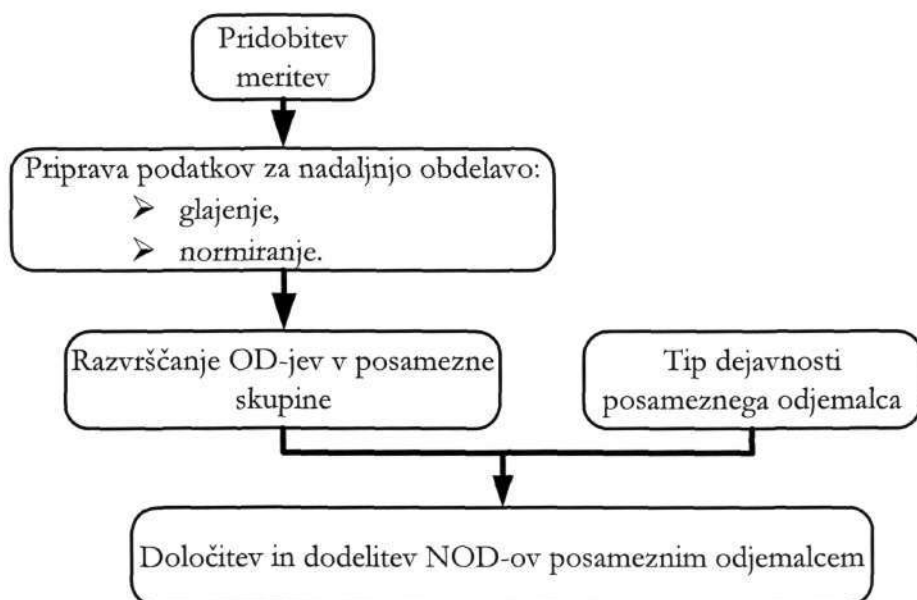
Način odjema električne energije odjemalcev različnih dejavnosti je lahko podoben, zato njihove OD-je lahko združimo tako, da jih predstavlja samo en NOD, ki je povprečna vrednost vseh OD-jev v izbrani skupini odjemalcev.

Pri pregledu literature s področja določanja NOD-ov smo spoznali dva različna postopka. Prvi temelji na večletnih raziskavah obremenitev posameznih skupin odjemalcev. Tako so raziskovalci določili število skupin ter karakteristike odjemalcev, ki sestavljajo omenjene skupine. Drugi postopek temelji na razvrščanju izmerjenih OD-jev posameznih odjemalcev. Pri njem je treba določiti način, s katerim bomo NOD-e dodelili posameznim skupinam UO-jev. V disertaciji predlagamo uporabo novega postopka določanja ali oblikovanja NOD-ov, ki temelji na metodah razpoznavanja vzorcev oz. razvrščanju OD-jev.

Predlagani postopek oblikovanja NOD-ov in njihove dodelitve posameznim UO-jem prikazuje Sl. 3.1. Razdelimo ga na štiri večje korake:

- pridobitev meritev od upravljavcev distribucijskih omrežji,
- priprava meritev za nadaljno obdelavo,
- razvrščanje OD-jev,
- določitev in dodelitev NOD-ov posameznim UO-jem na podlagi odjemalčevega tipa dejavnosti.

Meritve porabe posameznih UO-jev za vsak 15 minutni interval vzorčenja smo dobili od slovenskih distribucijskih podjetij. Vpliva, katere odjemalce naj distribucijska podjetja opremijo z ustreznimi merilnimi napravami, nismo imeli.



Sl. 3.1: Postopek oblikovanja in določanja NOD-ov posameznim UO-jem.

Pripravo meritev, drugi korak v postopku določanja NOD-ov, sestavljata dva podkoraka:

- pridobljene OD-je najprej razdelimo glede na tip dneva in jih. V ta name smo uporabili valčno analizo,
- normiranje.

Podrobnejši opis izvedenih nalog v drugem koraku podajamo v razdelku 3.1 – Priprava podatkov za nadaljnjo obdelavo.

Razvrščanje OD-jev v posamezne NOD-e je ena izmed ključnih nalog postopka oblikovanja NOD-ov. V ta namen smo uporabili tehnike razpoznavanja vzorcev. V pričujočem delu bomo razvrščanje OD-jev imenovali rojenje (ang. clustering). Terminologijo smo povzeli po literaturi [36]. Zaradi zahtevnosti rojenja je za doseganje "optimalnih" rezultatov razvrščanja treba preizkusiti več metod. V nalogi smo preizkusili naslednje metode rojenja:

- rojenje s pomočjo mehke logike, Fuzzy C-means (FCM) algoritem,
- hierarhične metode, rojenje z izračunom Evklidske razdalje med posameznim OD-ji, vzorci, ter povprečne razdalje med roji,
- nevronska omrežja, samo-organizirajoča omrežja (SOM).

Posvetili smo se tehnikam razpoznavanja vzorcev in jih podrobneje opisujemo v razdelku 3.2 Razpoznavanje vzorcev. V njem so predstavljene posamezne metode rojenja ter način določanja končnega števila rojev. Posamezni roji določajo NOD-e.

Zadnji, četrti korak postopka je dodelitev NOD-ov posameznim skupinam UO-jev. Za dodelitev NOD-ov odjemalcem smo uporabili tip dejavnosti posameznega UO-ja. Podrobnejši opis dodeljevanja NOD-ov podajamo v razdelku 3.3 Dodelitev NOD-ov

posameznim UO-jem. Pri tem smo uporabili metodo nadziranega razpoznavanja vzorcev, in sicer verjetnostna nevronska omrežja (VNO), ki smo jih podrobneje opisali v razdelku o razpoznavanju vzorcev.

Do sedaj smo na kratko opredelili posamezne korake pri določanju NOD-ov. V nadaljevanju sledijo podrobnejši opisi posameznik korakov ter opis uporabljenih metod v posameznih korakih.

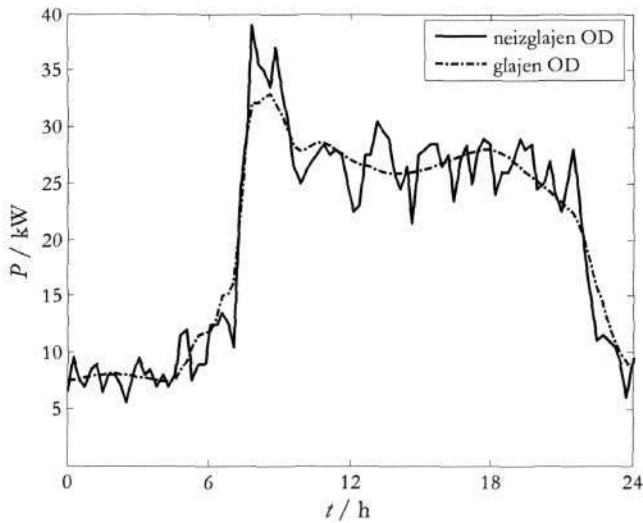
### 3.1. Priprava podatkov za nadaljnjo obdelavo

Ustrezna priprava meritev je pogoj za njihovo uspešno razvrščanje. V tem koraku postopka oblikovanja NOD-ov je treba meritve organizirati na način, kjer predhodno izločimo največ dejavnikov, ki vplivajo na obliko OD-jev ter s tem na njihovo nadaljnjo razvrstitev. Zato smo OD-je razdelili glede na tip dneva, delavnik od ponedeljka do petka, sobota in nedelja ter glede na sezono ali mesec v letu. Vpliv na porabo imajo tudi posebni dnevi, kot so prazniki, vendar se z njimi ne bomo posebej ukvarjali. V razdelkih 2.1.1 in 2.2.2 smo zapisali, da na obliko izmerjenega OD-ja v veliki meri vplivata tudi naslednja faktorja:

- število električnih naprav, katere posamezni UO-ji v določenem trenutku uporabljajo,
- način merjenja in čas vzorčenja.

Drugi faktor je za distribucijska podjetja določen in nanj ne moremo vplivati. Prav tako ne moremo vplivati na število naprav, ki jih UO-ji uporabljajo. Njihov vpliv poizkusimo izločiti tako, da za delavnik posameznega UO-ja določimo povprečje več delavnikov. V delu predstavlja delavnik povprečje štirih sredinih delavnikov v mesecu. Za sredo smo se odločili, ker je najbolj tipičen delovni dan. Povprečje štirih delavnikov posploši obliko OD-ja, vendar nekateri OD-ji še vedno vsebujejo veliko mero šuma, ki ga je treba odstraniti ter s tem zgladiti OD. V ta namen smo uporabili valčno multiresolucijsko analizo, ki je del teorije valčkov. Zanje smo se odločili, ker valčki predstavljajo bazne funkcije, katerih oblika je bolj podobna OD-jem, kakor sta trigonometrični funkciji sinus in kosinus. Z izbiro primerne valčka lahko natančneje opišemo posamezen OD.

Za izločanje šuma bi lahko uporabili različne klasične metode, kot so nizkopropustni filtri ali Fourierova transformacija, vendar imajo te metode določene omejitve. Njihova pomanjkljivost je bodisi v zamegljevanju ostrih lastnosti analiziranih signalov bodisi v tem, da niso primerna za analizo nestacionarnih signalov, ker poizkušajo razširiti informacijo o lastnostih signala v izbranem trenutku na široko območje frekvenc. Primer merjenega ter glajenega OD-ja prikazuje Sl. 3.2.



Sl. 3.2: Primer zglajenega in neizglajenega OD-ja.

### 3.1.1. Valčki

Valčki so funkcije  $\psi(t)$ , ki jih uporabimo kot osnovne gradnike, s katerimi opišemo in analiziramo druge funkcije, podobno kot uporabimo sinuse in kosinuse pri Fourierovi analizi. Da tovrstne funkcije imenujemo valčki, morajo zadostovati naslednjim pogojem:

- morajo biti oscilatorne narave,
- imeti morajo efektivno omejene amplitude, ki hitro padajo proti nič,
- imeti morajo povprečno vrednost nič,
- biti morajo dobro časovno ter frekvenčno lokalizirani.

Omenjene lastnosti omogočajo dobro določanje frekvenc, ki so se v časovni vrsti pojavile le v nekem obdobju. V nalogi bomo za časovno vrsto privzeli OD-je, ter jih za potrebe predstavitve valčne analize označevali kot časovno vrsto ali funkcijo  $f(t)$ .

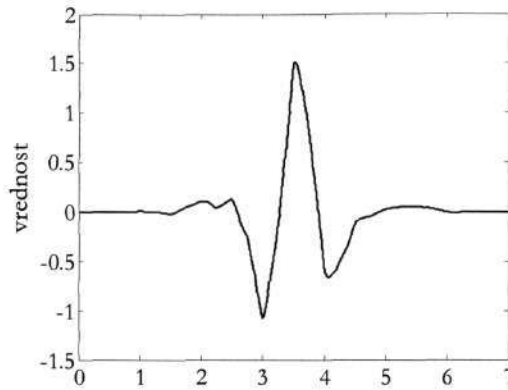
Poljubno funkcijo  $f(t)$  zapišemo kot linearno kombinacijo izpeljanih valčkov

$$f(t) = \sum_{a,b} c_{a,b} \cdot \frac{|a|^{-1/2}}{b} \cdot \psi(t-b), \quad a, b \in \mathbb{Z}, \quad (3.2)$$

kjer so  $c_{a,b}$  koeficinti valčne transformacije.

Nove oblike valčkov dobimo s pomočjo skaliranja in premikov osnovnega valčka. Skaliranje pomeni, da je osnovni valček ali raztegnjen ali stisnjen, medtem ko translacija pomeni premikanje osnovnega valčka v nekem časovnem območju. Pri tem se energija novih valčkov ohranja. Aproximacijo osnovnega valčka družine Simlet reda 4 prikazuje Sl. 3.3.



Sl. 3.3: Aproximacija osnovnega valčka družine Simlet reda 4  $\psi(t)$ .

Določitev novih oblik valčkov iz osnovnega valčka podaja naslednja enačba:

$$\psi_{a,b} = |a|^{-1/2} \cdot \psi\left(\frac{t-b}{a}\right), \quad (3.3)$$

kjer je:  $a$  faktor skaliranja in  $b$  faktor, s katerim se pomikamo po časovni osi  $t$ .

Valčno transformacijo zapišemo kot sledi:

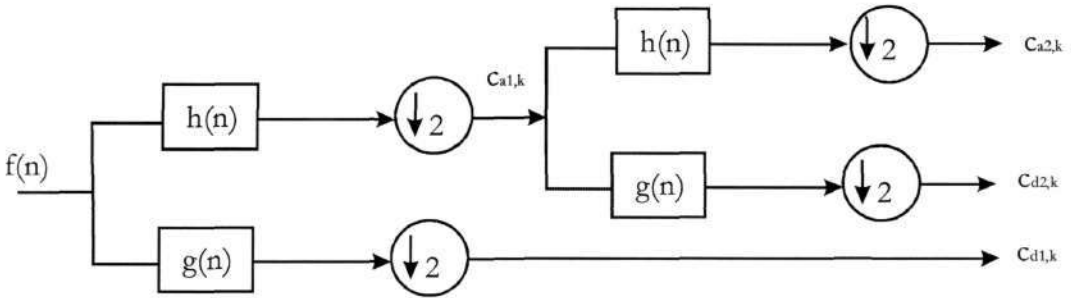
$$c_{a,b} = \int_{-\infty}^{\infty} f(t) \cdot \frac{1}{\sqrt{a}} \cdot \psi\left(\frac{t-b}{a}\right) dt. \quad (3.4)$$

Transformacijo dosežemo s skaliranjem in časovnim premikom po časovni osi osnovnega valčka ter izračunom podobnosti med signalom in valčkom na ustrezni skali in časovni osi. V praksi imamo dve obliki transformacije, zvezno in diskretno transformacijo. Razlikujeta se glede na zalogo vrednosti parametrov  $a$  in  $b$ . Zaloga vrednosti parametrov zvezne transformacije je:  $a \in \mathbb{R}^+ - \{0\}$ ,  $b \in \mathbb{R}$ . Z zveznim premikanjem in skaliranjem osnovnega valčka pridobimo veliko redundantne informacije, zato se velikokrat uporablja diskretna valčna transformacija, kjer zaloga vrednosti parametrov temelji na potencah števila dva:  $a = 2^j$ ,  $b = k \cdot 2^j$ ,  $j, k \in \mathbb{Z}$ . Takšne vrednosti imenujemo diadične vrednosti. V delu smo uporabljali diskretno valčno transformacijo.

### 3.1.1.1. Valčna multiresolucijska analiza

Valčna multiresolucijska analiza je orodje, ki omogoča analizo poljubnih signalov na različnih časovnih skalah in pri različnih frekvencah. Njegova osnovna ideja je, da signal razcepimo na mehko komponento, tj. približek signala in na detajlno komponento signala. Približek signala predstavljajo nizkofrekvenčne komponente signala na večjih skalah, medtem ko predstavljajo detajli komponente signala na nižjih skalah pri višjih frekvencah.

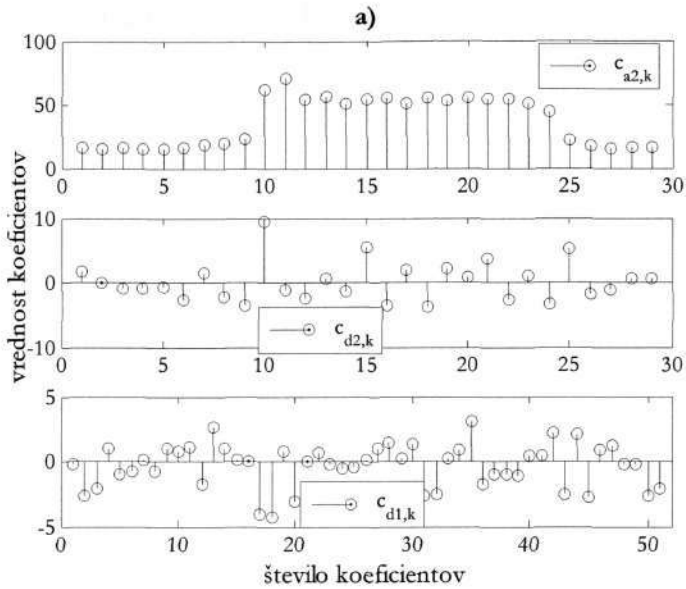
Multiresolucijsko analizo izvedemo s konvolucijo časovno diskretnega signala  $f(n)$ ,  $n \in \mathbb{Z}$ , ali časovne vrste z ustreznim filtrom. Rezultat so valčni koeficienti, ki predstavljajo detajle  $c_{dj,k}$  in aproksimacijo  $c_{aj,k}$ . Signal  $f(n)$ , v našem primeru je to OD, na prvem nivoju razdelimo na koeficiente  $c_{dj,k}$  s konvolucijo z visokofrekvenčnim filtrom, ki ga opisujejo njegovi koeficienti  $h(n)$  ter na koeficiente  $c_{aj,k}$  s konvolucijo z nizkofrekvenčnim filtrom  $g(n)$ . Proces razcepa signala nadaljujemo do želene resolucije ali skale  $J$  tako, da v vsakem nadaljnjem koraku izvedemo enak proces in sicer z razdelitvijo aproksimacije. Simbolični, rekurzivni potek razcepa imenujemo valčno dekompozicijsko drevo, ki ga za razcep na dveh skalah prikazuje Sl. 3.4. Takoj za filtriranjem sledi diadično zmanjševanje resolucije za faktor 2.



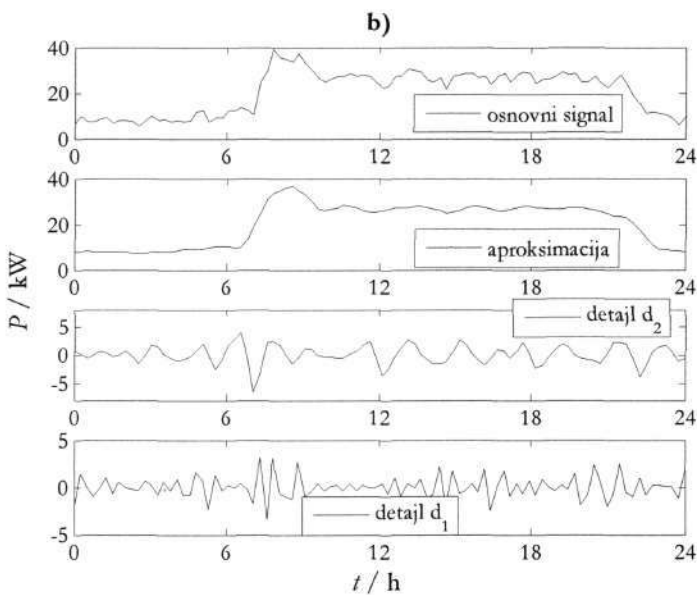
Sl. 3.4: Valčno dekompozicijsko drevo za razčlenitev na dveh skalah.

Podobno kakor razcep signala na posamezne detajle in na aproksimacijo si predstavljamo sintezo posameznih prispevkov v izvorni signal v nasprotni smeri, kakor ga prikazuje zgornja slika. Pri tem namesto visoko in nizko frekvenčnih filtrov za dekompozicijo uporabljamo visoko  $h'(n)$  in nizko frekvenčne  $g'(n)$  filtre za rekonstrukcijo.

Primer razčlenitve vhodnega signala, tj. obremenitvenega diagrama, s Simlet valčki četrtega reda (sym4) prikazuje Sl. 3.5. Red valčka določa število filtrskih koeficientov. Razčlenitev je narejena na dveh skalah  $2^j$ ,  $j = 1, 2$ . Iz Sl. 3.5 se vidi, da se število valčnih koeficientov pri prehodu iz nižje na višjo skalo močno zmanjša. Teoretično se bi moralo število koeficientov zmanjšati za polovico, vendar je zaradi procesa filtriranja, ki poteka s konvolucijo signala in filtra, signal nekoliko "zamazan" zaradi dodajanja dodatnih vzorcev v signal [31]. Zaradi tega imamo nekoliko večje število koeficientov. Poleg tega se tudi vidi, da so valčni koeficienti detajlov  $c_{dj,k}$  na posamezni skali veliki le pri velikih spremembah originalnega signala. Tako se pomembne informacije o signalu zberejo v majhnem številu koeficientov z veliko absolutno vrednostjo. Z večanjem skale se število koeficientov povečuje. Sl. 3.6 prikazuje prispevke koeficientov na posameznih skalah k vhodnemu signalu. Vidimo, da s koeficienti aproksimacije opišemo vhodni signal grobo, z manjšo ločljivostjo, in s koeficienti detajlov večje in manjše podrobnosti pri različnih ločljivostih. Vsota prispevkov valčnih koeficientov k vhodnemu signalu da vrednosti vhodnega signala, kar je v skladu s tem, kar smo zapisali v uvodnem delu pričujočega razdelka. Poljuben signal ali funkcijo  $f(t)$  zapišemo kot vsoto prispevkov valčnih koeficientov na posameznih skalah.



Sl. 3.5: Razčlenitev vhodnega signala s Simlet valčki četrtega reda: a) valčni koeficienti aproksimacije  $c_{a2,k}$  in dveh detajlov  $c_{d2,k}$  in  $c_{d1,k}$ .



Sl. 3.6: Razčlenitev vhodnega signala s Simlet valčki četrtega reda: b) prispevek valčnih koeficientov na vsakem nivoju k vhodnemu signalu.

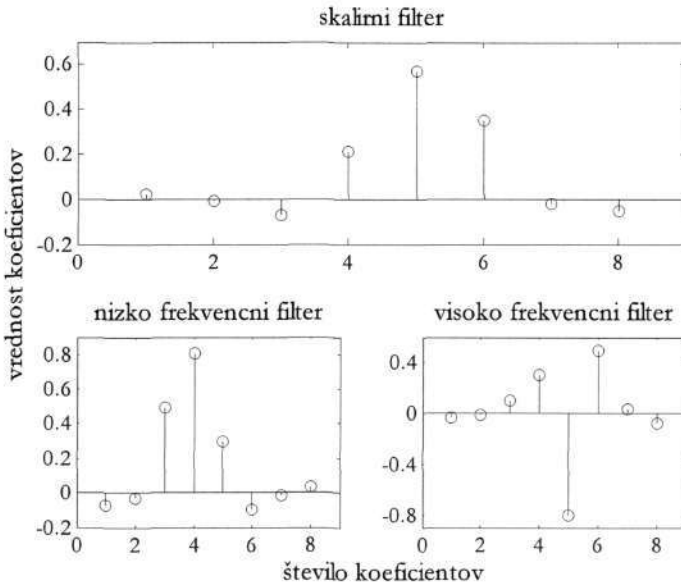
Povedano se vidi tudi iz enačbe (3.5), ki podaja inverzno diskretno valčno transformacijo:

$$f(t) = \sum_{i,k \in \mathbb{Z}} c_{i,k} \cdot \psi_{i,k}(t), \quad (3.5)$$

kjer so  $c_{i,k}$  valčni koeficienti.

Obstaja neposredna povezava med obliko valčka ter uporabo posameznega nizko in visoko frekvenčnega filtra. Izbrani valčki določajo filtre, ki jih uporabimo za multiresolucijsko analizo

ter obratno, izbrani filtri določajo valček. Ko izberemo določen valček, določimo njegov skalirni filter imenovan na podlagi katerega izračunamo nizko in visoko frekvenčna filtra [31]. Primer skalirnega filtra in dekompozicijskih filtrov za valček Simlet reda štiri prikazuje Sl. 3.7.



Sl. 3.7: Primer filtrov za valček Simlet reda 4.

### 3.1.1.2. Odstranjevanje šuma (Glajenje)

Vzemimo, da imamo časovno vrsto, ki ima vzorce v časovnih korakih  $t_k = t_0 + i\Delta t$ ,  $i = 0, \dots, N-1$  in jo zapišemo z naslednjo enačbo:

$$x_i = f(t_i) + e_i, \quad i = 0, \dots, N-1 \quad (3.6)$$

kjer je:  $f(t)$  ustrežna gladka neznana funkcija,  $e_i$  naključne normalno porazdeljene napake s povprečjem nič in varianco  $\sigma^2$ .

Poiskati želimo takšno funkcijo  $g(t)$ , pri kateri je povprečna kvadratna napaka

$$\epsilon = \frac{1}{N} \sum_{i=0}^{N-1} (f(t_i) - g(t_i))^2 \quad (3.7)$$

minimalna, ter vsaj toliko gladka kot funkcija  $f(t)$ .

Aprksimacijo in detajle razčlenjene časovne vrste, tj. OD-ja, prikazujeta Sl. 3.5, kjer smo podali posamezne valčne koeficiente, in Sl. 3.6, kjer podajamo prispevke k časovni vrsti. Iz slik se vidi, da so prispevki nekateri detajlov zelo majhni, kar ustreza tudi majhnim absolutnim vrednostim valčnih koeficientov. Če te koeficiente odstranimo ne spremenimo dosti oblike časovne vrste ali signala. Za uspešno glajenje časovne vrste je tako treba najti primeren način,

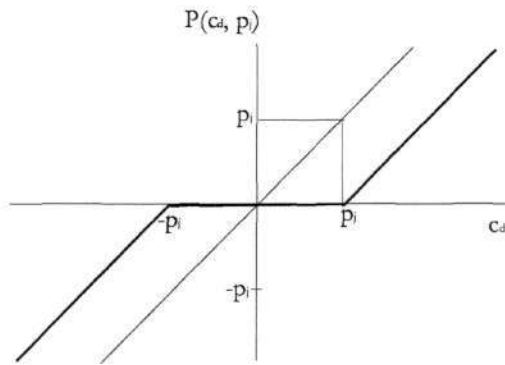
kako omejiti koeficiente detajlov  $c_{d,j,k}$ . Potek glajenja OD-jev s pomočjo valčne multiresolucijske analize razdelimo na tri korake:

- po rekurzivni shemi za razcep časovne vrste izračunamo valčne koeficiente,
- Koeficiente detajlov  $c_{d,j,k}$  na vseh skalah  $2^j$  popravimo s pragovno funkcijo  $P(c_d, p_j)$ , kjer je  $p_j$  ustrezen prag,
- po rekurzivni shemi za sestavljenje izračunamo zglajeno časovno vrsto.

Primer mehke pragovne funkcije, katero smo uporabljali v naši nalogi in je izvedena v Matlabovi orodjarni o valčkih ter jo opisuje [31], podaja enačba:

$$P(c_d, p_j) = \begin{cases} c_d + p_j, & c_d < -p_j \\ 0, & |c_d| < p_j \\ c_d - p_j, & c_d > p_j \end{cases} \quad (3.8)$$

ter jo prikazuje Sl. 3.8.



Sl. 3.8: Mehka pragovna funkcija  $P(c_d, p)$ .

Z mehko pragovno funkcijo vse koeficiente detajlov  $c_d$  manjše od praga  $p_j$  postavimo na nič ter upoštevamo, da koeficienti zunaj praga tudi vsebujejo šum, zato jih za vrednost praga premaknemo proti nič. Za uspešno glajenje je ključnega pomena določitev ustreznega praga  $p_j$ , ki ga določimo z naslednjo enačbo:

$$p_j = \hat{\sigma}_j \cdot \sqrt{2 \cdot \ln N_j}, \quad (3.9)$$

kjer je:  $N_j$  število koeficientov na  $j$ -ti skali,  $\hat{\sigma}_j$  ocena standardnega odklona.

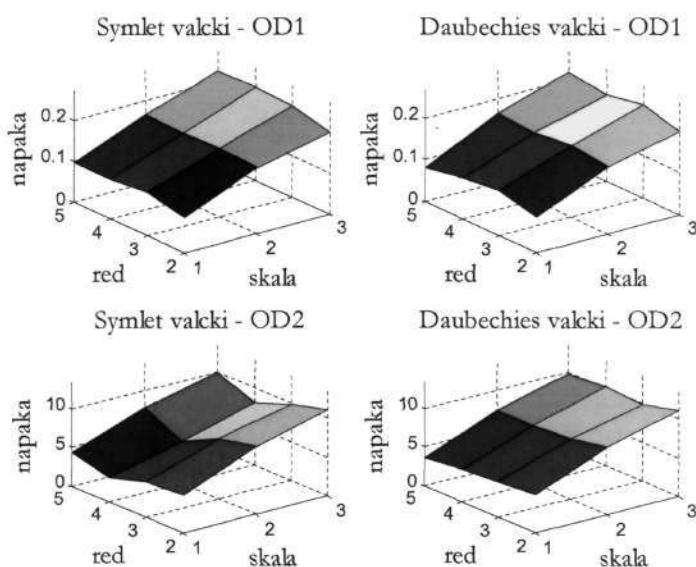
Oceno standardnega odklona naredimo na podlagi mediane absolutnih odstopanj (MAD) vrednosti valčnih koeficientov ter količnika 0.6745 [31] kot prikazuje naslednja enačba:

$$\hat{\sigma}_j = \frac{\text{MAD}(cd)}{0.6745}. \quad (3.10)$$

Verjetnost, da je zglijena funkcija  $g(t)$  vsaj tako gladka kakor funkcija  $f(t)$ , je za predstavljeno izbiro pragovne funkcije ter vrednosti pragov zelo velika. Predstavitev glajenja z valčno multiresolucijsko analizo smo povzeli po literaturi [31], [32].

Poleg izbire pragovne funkcije ter velikosti praga je pri glajenju z valčki treba določiti tudi družino valčka, red valčka ter število skal, na katerih bomo gladili. Omenjena naloga ni vedno enostavno rešljiva. To velja še posebej za izbiro skale razcepa, ker moramo pri njeni izbiri poiskati kompromis med stopnjo zglijenosti OD-ja ter izgubo informacije o OD-ju.

Primerjali smo glajenje z valčki Symlet in Daubechies. Spreminjali smo red valčka od dva do pet ter skalo razcepa  $j$  od ena do tri. Skala številka tri predstavlja maksimalno stopnjo razčlenitve OD-ja za analizirana valčka. Odvisna je od izbranega valčka in dolžine signala, v našem primeru OD-ja, ki ga sestavlja 96 vrednosti v enem dnevu. Določa jo pogoj, da je pri razčlenitvi OD-ja na valčne koeficiente vsaj en koeficient pravičen [31]. Pri tem smo opazovali odvisnost srednje kvadratne napake med zglijenim OD-jem in dejanskim OD-jem od reda valčka ter skale razčlenitve OD-ja. Analizirali smo 100 poljubno izbranih OD-jev iz baze podatkov, kar predstavlja približno 10 % vseh odjemalcev. Rezultate glajenja dveh OD-jev z Daubechies in Symlet valčki prikazuje Sl. 3.9



Sl. 3.9: Srednja kvadratna napaka pri glajenju s Symlet in Daubechies valčki.

Iz zgornje slike se vidi, da obe družini valčkov podajata primerljive rezultate. Za predstavljen primer glajenja red valčka ne vpliva poglavitno na velikost napake pri valčkih iz Daubechies-ove družine, medtem ko je omenjena odvisnost pri Symlet valčkih večja, glej OD2. Na napako močno vpliva velikost skale razčlenitve  $J$ , z večanjem skale se napaka povečuje. Poleg reda valčka in skale razčlenitve na velikost napake najbolj vpliva oblika odjemalčevega OD-ja, saj se razlikujejo tudi za nekaj velikostnih razredov, kar se vidi iz primerjave vrednosti napak OD-jev na zgornji sliki. V prvem primeru dosega napaka vrednosti do 0,22, medtem ko pri drugem OD-ju dosega napaka vrednosti do 10. Ti OD-ji, ki vsebujejo več šuma, so bili v procesu



glajenja "bolj" zglajeni kot preostali, zato bolj odstopajo od neizglajenih OD-jev vendar jim vseeno "sledijo". Povprečne vrednosti in mediane srednjih kvadratnih napak za analizirane vrednosti skal N in redov valčkov prikazuje Tab. 3.1.

Tab. 3.1: Povprečne vrednosti in mediane srednjih kvadratnih napak.

	Symlet valčki			Daubechies valčki		
	povprečna vrednost napake			povprečna vrednost napake		
red valčka	N=1	N=2	N=3	N=1	N=2	N=3
2	2,079	4,196	6,050	2,079	4,198	6,050
3	2,056	4,205	5,823	2,056	4,205	5,823
4	2,094	4,281	5,806	2,047	4,263	5,934
5	1,945	4,104	5,965	2,090	4,190	6,045
	mediana napake			mediana napake		
red valčka	N=1	N=2	N=3	N=1	N=2	N=3
2	0,540	1,094	1,461	0,540	1,094	1,461
3	0,562	1,143	1,530	0,562	1,143	1,530
4	0,546	0,986	1,390	0,588	1,069	1,514
5	0,506	1,068	1,467	0,600	1,240	1,629

Iz tabele se vidi, da glajenje z obema družinama valčkov daje primerljive rezultate za vse skale razčlenitve in rede valčkov. Cilj glajenja je odstranitev čim več šuma, zato smo izbrali maksimalno dovoljeno skalo razčlenitve  $N = 3$ , višja kot je skala razčlenitve, bolj je OD zglajen. Pri dani skali smo določili tisti red valčka, ki ima minimalno povprečno kvadratno napako, tj. red 4. Primerjava povprečnih vrednosti in mediane napak kaže, da je prva relativno majhna v primerjavi z drugo. Ugotovimo, da večina glajenih OD-jev ne odstopa dosti od neizglajenih ali merjenih OD-jev ter da morajo obstajati nekateri primeri, ko so takšna odstopanja občutnejša, kar sklepamo na podlagi razmerij med medianami in povprečnimi vrednostmi srednjih kvadratnih napak.

### 3.1.2. Normiranje

Ker se lahko konična poraba posameznih UO-jev razlikuje za več velikostnih razredov, OD-je normiramo. Normiranje izvedemo zato, ker metode, ki smo jih uporabljali pri rojenju, temeljijo na razdalji ali distanci med posameznima OD-jema. Pri normiranju je treba paziti, da ostanejo medsebojna razmerja posameznih vrednosti OD-ja po normiranju nespremenjena.

Dve obliki normiranja, ki zadostujeta zgornjemu pogoju, podajata enačbi (3.11) in (3.12):

$$z_{ij} = \frac{x_{ij}}{\max x_j}, \quad i = 1, \dots, 96 \quad (3.11)$$

kjer je:

$z_{ij}$  – normirana vrednost OD-ja,

$x_{ij}$  – dejanska vrednost OD-ja,

$x_j$  – OD j-tega odjemalca.

V tem primeru so vse vrednosti meritve  $x_j$  deljene z njihovo maksimalno vrednostjo – konično močjo. Literatura [19] uporablja normiranje, ki ga podaja naslednja enačba:

$$z_{ij} = \frac{x_{ij}}{P_B} \quad (3.12)$$

$$P_B = \frac{1}{24} \int_{24 \text{ ur}} \mathbf{M}(t) dt = \frac{\text{mesečna energija (kWh)}}{24 \cdot \text{št. dni v mesecu}}$$

kjer  $P_B$  predstavlja povprečno mesečno porabo posameznega odjemalca. Z normiranjem po enačbi (3.11) transformiramo vse meritve na interval  $[0,1]$ , medtem ko normiranje po enačbi (3.12) ne daje enotnega intervala za vse meritve. Zaradi tega smo v nalogi uporabili normiranje po enačbi (3.11).

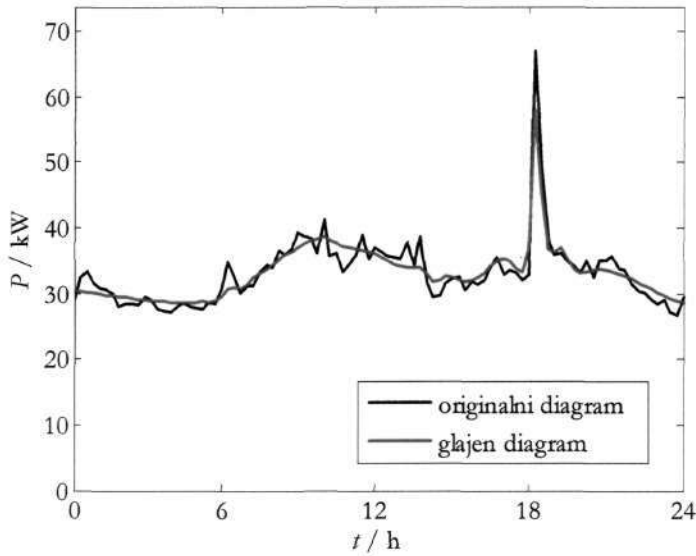
Poleg omenjenega načina normiranja predlagajo avtorji v literaturi s področja razpoznavanja vzorcev še dodatne načine normiranja podatkov, ki niso primerni za uporabo pri razvrščanju ali rojenju OD-jev. Enega izmed predlogov je uporabil avtor v [29]. Podaja ga naslednja enačba:

$$z_{ij} = \frac{x_{ij} - \min x_j}{\max x_j - \min x_j}, \quad i = 1, \dots, 96. \quad (3.13)$$

Zaloga vrednosti normiranega diagrama po enačbi (3.13) je na intervalu  $[0, 1]$  in je enaka, kot pri normiranju z enačbo (3.11). Njena pomanjkljivost je, da "odreže" celotno pasovno obremenitev odjemalca in s tem poruši razmerje med točkami OD-ja. Uporaba takšnega normiranja bi transformirala OD s pretežno pasovnim odjemom v diagram s trapeznim odjemom. Takšen način normiranja za nas ni primeren ne glede na njegovo siceršnjo uporabnost pri razvrščanju vzorcev.

Uporaba izbranega načina normiranja zahteva dodatno previdnost, saj v določenih primerih maksimalna vrednost OD-ja lahko podaja zavajajoče rezultate. Enega izmed takšnih primerov prikazuje Sl. 3.10. Na sliki vidimo primer OD-ja, ko je poraba relativno konstantna preko celotnega dneva z izjemo enega 15 minutnega intervala, ko nastopi konica. Razloga za nastop opisane konice sta dva. Prvi razlog je v dejanskem povečanju porabe v opazovanem časovnem intervalu. Drugi razlog predstavlja motnja v delovanju merilnega inštrumenta, ki zabeleži malo verjetno vrednost porabe. V postopku analize OD-jev ne moremo vedeti, kaj je vzrok omenjenemu skoku porabe, zato je najprimerneje, da opisano konico predhodno zmanjšamo. Zahteva po začetnem zmanjšanju konic se pokaže tudi v postopku glajenja, ko algoritem namenjen glajenju ne odpravi ali zmanjša konice dovolj, glej Sl. 3.10. Iz slike se vidi, da glajenje ne odpravi ali zmanjša vrednosti konice. Če predstavljeni OD normiramo z uporabo enačbe (3.11), naredimo napako, ki vpliva na rojenje, saj se OD ne razvrsti v skupino, ki predstavlja

odjemalce s pasovnim odjemom električne energije. Pomanjkljivost odpravimo tako, da omenjeno konico pred postopkom glajenja zmanjšamo.



Sl. 3.10: Primer zavajajočih rezultatov pri uporabi konične moči kot faktorja normiranja.

Konice, kot je na Sl. 3.10, lahko nastopijo na različnih delih OD-ja in ne samo kot konična obremenitev, zato smo jih pred postopkom glajenja zmanjšali z uporabo enačbe:

$$x_{ij} = x_{i-1,j} + \kappa \cdot |x_{i-1,j} - x_{ij}| \quad (3.14)$$

kjer je:

$\kappa$  – faktor za katerega vrednost zmanjšamo konico,  
 $x_{ij}$  – dejanska vrednost OD-ja.

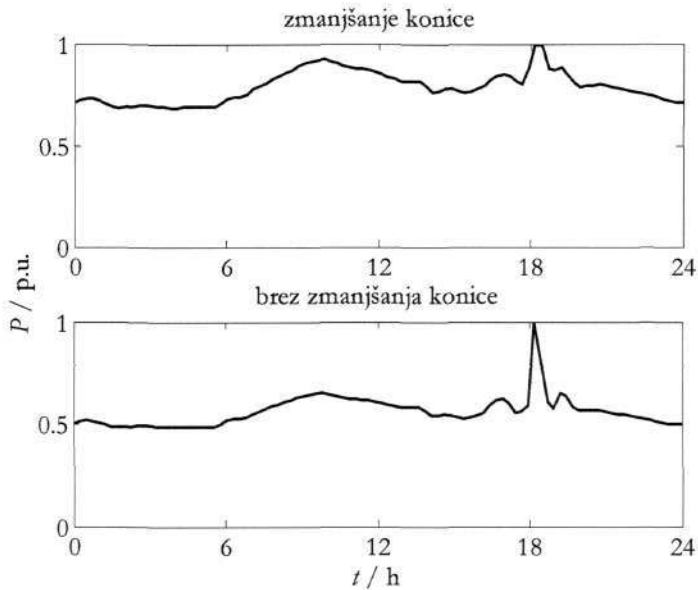
Enačba (3.14) podaja vrednost znižanja konice. Predhodno je treba določiti, katere vrednosti OD-ja predstavljajo konico. V nalogi smo poimenovali konice vse  $i$ -te,  $i \in \xi$ , vrednosti  $j$ -tega OD-ja  $x_{ij}$ , kjer je sprememba porabe pred in po nastopu konice večja od parametra  $\sigma$ , kar zapišemo z enačbo:

$$i \in \xi \quad \left| \begin{array}{l} \frac{x_{ij} - x_{i-1,j}}{\max x_j} > \sigma \\ \frac{x_{i+1,j} - x_{ij}}{\max x_j} < -\sigma \end{array} \right. , \quad i = 1, \dots, 96, \quad (3.15)$$

kjer je:

$\xi$  – množico vseh  $i$ -jev, kjer nastopajo konice,  
 $\sigma = 0.1$ ,  $\kappa = 1/3$  – empirično določena parametra na podlagi vizualnega pregleda rezultatov odstranjevanja konic.

Primerjavo rezultatov normiranja s predhodnim odstranjevanjem konic in brez njihove odstranitve za OD na Sl. 3.10 prikazuje Sl. 3.11. Iz slike se vidi, da zmanjšanje konice pripomore k izboljšanju normiranja.



Sl. 3.11: Primerjava normiranih OD-jev pri zmanjšanju konice ter brez zmanjšanja.

V nadaljevanju si bomo podrobneje ogledali metode, namenjene razpoznavanju vzorcev, katere smo uporabljali v nalogi, saj si posamezen OD  $x_i$  predstavljamo kot vzorec, opisan s 96 parametri, vrednostmi.

### 3.2. Razpoznavanje vzorcev

Razpoznavanje vzorcev je področje, ki opisuje, kako "stroji" opazujejo okolico, se učijo razlikovati zelene vzorce od ozadja, jih medsebojno grupirati ter podajati jasne in razumne odločitve o njihovih kategorijah in lastnostih. Avtomatsko razpoznavanje, opis, klasifikacija in grupiranje vzorcev ima velik pomen pri različnih disciplinah, kot so: medicina, biologija, psihologija, računalniški vid, umetna inteligenca. Vendar, kaj sploh je vzorec? Vzorec je definiran kot nasprotje kaosa. Je celota, ki jo je možno poimenovati. Tako je vzorec lahko prstni odtis ali obremenitveni diagram.

Če imamo množico vzorcev, katerih lastnosti niso podane vnaprej, se njihovo razpoznavanje ali razvrščanje izvede s pomočjo ene izmed naslednjih tehnik:

- nadzirajoče razvrščanje,
- nenadzirajoče razvrščanje.

Pri nadzirajočem razvrščanju se vhodni vzorec spozna kot pripadnik v naprej določenega razreda, ki ga je določi načrtovalec. Načrtovalec pozna obliko ali lastnosti vzorca, ki ga išče v vhodni množici. Pri nenadzirajočem razvrščanju vzorce dodelimo vnaprej neznanemu razredu

na podlagi podobnosti med vzorci. Načrtovalec išče možne podobne oblike vzorcev v vhodni množici. Pri tem načrtovalec pozna najmanj informacij o izbrani množici vzorcev, kar otežuje razvrščanje. Načrtovalec mora biti sposoben naučiti se posameznih lastnosti iz vhodne množice vzorcev. Nenadzirajoče razvrščanje je najzahtevnejša oblika razvrščanja, saj načrtovalec orodja za razvrščanje ne pozna lastnosti vzorcev ter končnega števila podobnih skupin, v katere mora podano množico vzorcev razvrstiti. V to kategorijo spadajo metode rojenja (ang. cluster analysis), ki jih nadalje delimo na hierarhične, nehierarhične, geometrijske ter ostale.

V nalogi se bom omejil na obravnavo in uporabo metod iz prvih dveh skupin. Njihova bistvena medsebojna razlika je, da je treba pri nehierarhičnih metodah vnaprej podati končno število podobnih skupin. Glavni cilj metod rojenja je, da imajo vzorci znotraj določenega roja podobne karakteristike, medtem ko se vzorci, ki pripadajo različnim skupinam, očitno razlikujejo med seboj.

Zaradi zahtevnosti rojenja je za doseganje "optimalnih" rezultatov razvrščanja treba preizkusiti več metod. Metode, primerne za reševanje problema rojenja OD-jev, ki smo jih preizkusili, temeljijo na mehki logiki, statistiki in nevronskega omrežju. Njihove predstavitve podajamo v nadaljevanju.

### 3.2.1. Fuzzy C-means rojenje (FCM)

Metodo je predlagal Bazdek [33] in temelji na teoriji mehke logike, kjer stopnjo pripadnosti vzorca določenemu roju podaja pripadnost. Fuzzy C-means (FCM) algoritem je en izmed najpogosteje uporabljenih algoritmov, ki temeljijo na teoriji mehkih množic in na podlagi prototipov. Prototip je vzorec, ki predstavlja celotno populacijo vzorcev v posamezni skupini ali roju. Algoritem definiramo na naslednji način.

Predpostavimo, da imamo množico  $N$  vzorcev  $\mathbf{X} = \{\mathbf{x}_k \mid k = 1, \dots, N\}$ , kjer posamezen vzorec  $\mathbf{x}_k = [x_{k,1}, x_{k,2}, \dots, x_{k,M}]$  opišemo z  $M$  parametri. Vzemimo, da  $\mathbf{B} = [\beta_1, \beta_2, \dots, \beta_C]$  predstavlja  $C$  prototipov ( $1 < C < N$ ), vsak opisuje posamezno skupino ali roj. Vsak prototip  $\beta_i$  opišemo z  $M$  parametri. V nalogi bomo z  $\beta_i$  označevali prototipe in roje. Algoritem predpostavlja, da je število podobnih skupin znano. Stopnjo pripadnosti vzorca  $\mathbf{x}_k$  v roju  $\beta_i$  označimo z  $u_{i,k}$ . Definiramo pripadnostno matriko  $\mathbf{U} = [u_{i,k}]$  dimenzij  $N \times C$ , ki mora zadostovati naslednjima pogojema:

$$u_{i,k} \in [0,1] \text{ za vsak } i, \quad 0 < \sum_{k=1}^N u_{i,k} < N \text{ za vse } i, \quad (3.16)$$

$$\sum_{i=1}^C u_{i,k} = 1, \text{ za vsak } k. \quad (3.17)$$

Problem rojenja  $N$  vzorcev v  $C$  skupin določimo kot optimizacijo kriterijske funkcije  $J(\mathbf{B}, \mathbf{U}, \mathbf{X})$  (3.18), ki jo moramo minimizirati:

$$J(\mathbf{U}, \mathbf{B}, \mathbf{X}) = \sum_{i=1}^C \sum_{k=1}^N u_{i,k}^m \cdot d^2(\mathbf{x}_k, \beta_i) \quad \min \quad (3.18)$$

kjer je:

$u_{k,i}$  – stopnja pripadnost vzorca k roju  $i$ ,

$d(\mathbf{x}_k, \beta_i)$  – norma (razdalja) med vzorcem  $\mathbf{x}_k$  in prototipom  $\beta_i$ ,

$m$  – potenčni faktor mehčanja,  $m \in (1, \infty)$ .

Razdaljo  $d(\mathbf{x}_k, \beta_i)$  definiramo na različne načine. V odvisnosti od njene definicije dobimo različne algoritme, katerim je skupna lastnost, da temeljijo na prototipih. V nalogi se bomo omejili na FCM algoritem, ki uporabljajo Evklidsko razdaljo za izračun razdalj med posameznim vzorcem in izbranim prototipom:

$$d(\mathbf{x}_k, \beta_i) = \|\mathbf{x}_k - \beta_i\|. \quad (3.19)$$

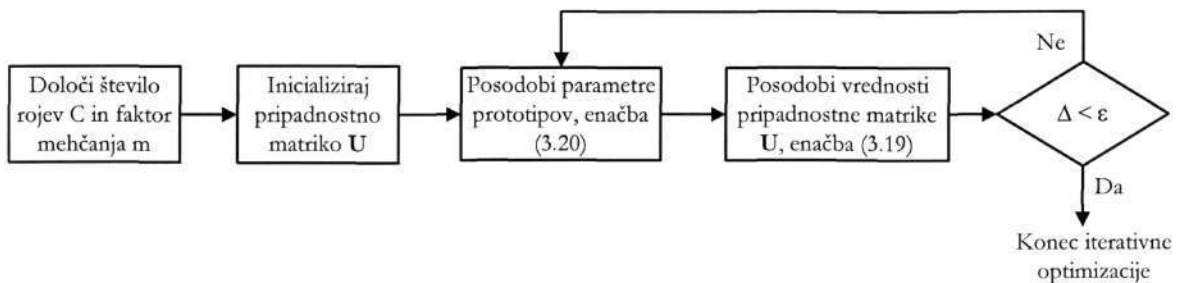
Minimizacija kriterijske funkcije (3.18) pri upoštevanju pogojev (3.16) in (3.17) poda analitični pogoj, kateremu morajo zadostovati pripadnosti  $u_{i,k}$ ,

$$u_{i,k} = \frac{1}{\sum_{j=1}^C \left( \frac{d^2(\mathbf{x}_k, \beta_j)}{d^2(\mathbf{x}_k, \beta_i)} \right)^{\frac{1}{m-1}}}. \quad (3.20)$$

Optimizacija poteka iterativno, zato je treba v vsaki iteraciji posodobiti vrednosti prototipov, ki jih podaja naslednja enačba:

$$\beta_i = \frac{\sum_{k=1}^N u_{k,i} \cdot \mathbf{x}_k}{\sum_{k=1}^N u_{k,i}}. \quad (3.21)$$

V splošnem zapišemo FCM algoritem v štirih korakih na način, kakor ga prikazuje Sl. 3.12:



Sl. 3.12: Prikaz postopka delovanja FCM algoritma.



Pogoj za prekinitev (stabilna razvrstitev) iterativne optimizacije je, da je spremenljivka  $\Delta$  manjša od poljubno majhnega števila  $\varepsilon$ . Spremenljivko  $\Delta$  definiramo kot razliko vrednosti kriterijske funkcije  $J(\mathbf{B}, \mathbf{U}, \mathbf{X})$  v dveh zaporednih iteracijah, enačba (3.22), ali kot normo razlike pripadnostnih matrik  $\mathbf{U}$  v dveh zaporednih iteracijah, enačba (3.23). V disertaciji smo uporabili enačbo (3.22) za izračun pogoja prekinitve iterativne optimizacije.

$$\Delta = J_i(\mathbf{U}, \mathbf{B}, \mathbf{X}) - J_{i-1}(\mathbf{U}, \mathbf{B}, \mathbf{X}) \quad (3.22)$$

$$\Delta = \|\mathbf{U}_i - \mathbf{U}_{i-1}\| \quad (3.23)$$

FCM algoritem uporablja potenčni faktor mehčanja  $m$ , katerega vrednost je poljubno pozitivno število večje od ena. Natančna vloga ali pomen faktorja mehčanja ni znan. Prav tako ni znano, kako izbrati njegovo optimalno vrednost. V nalogi smo s preizkušanjem različnih vrednosti izbrali faktor mehčanja  $m = 1,2$ .

Rezultat mehkega rojenja so prototipi  $\mathbf{B}$ , ki predstavljajo določeno populacijo vzorcev v izbrani skupini, ter pripadnostna matrika  $\mathbf{U}$ . Pripadnost vzorca  $u_{i,k}$ , zajema vrednosti med nič in ena, enačba (3.16), zato omenjeni algoritem ne podaja jasne (crisp) ločitve, kateri skupini ali roju pripada določen vzorec. Omenjeno dejstvo razumemo kot pomanjkljivost, če potrebujemo jasno razločitev pripadnosti (nič – ne pripada, ena – pripada).

V nasprotnem primeru se pokaže, da ima mehko rojenje prednost pred trdim rojenjem, saj povezujemo koncept pripadnosti pri teoriji mehke logike s konceptom utežnostnih funkcij pri statistiki. S pomočjo tega dejstva si lahko razlagamo trditve strokovnjakov, ki zagovarjajo teorijo mehke logike, češ, da je "mehak" način tolerantnejši do šuma in posebnežev (outliers) kakor trdi. Posebneži so vzorci, ki se močno razlikujejo od ostalih vzorcev v vhodni množici  $\mathbf{X}$ . Zaradi tega bi bilo smiselno, da imajo majhno pripadnost v vsakem izmed rojev  $C$ , vendar pogoj (3.17) ne dovoljuje, da bi imele vse pripadnosti vrednosti manjše od  $1/C$ . Posebnež  $\mathbf{x}_k$  ima pogosto vsa razmerja  $d(\mathbf{x}_k, \beta_i) / d(\mathbf{x}_k, \beta_j)$  enaka približno ena, kar ima za posledico, da so vrednosti pripadnosti blizu razmerja  $1/C$ . Za izračun vrednosti prototipov FCM algoritem uporablja vrednosti pripadnosti, enačba (3.20), zato je njihov položaj še vedno premaknjen iz njegove dejanske lege. Če je posebnež močno oddaljen od centra roja, je položaj prototipa močno premaknjen.

Problematiko posebnežev označuje občutljivost algoritma na šum. Zato so razvili algoritme, ki v določeni meri rešujejo omenjeno težavo. Glavno vodilo oblikovanja algoritmov je:

- posebneži naj bi imeli majhno vrednost pripadnosti v vseh rojih  $C$ ,
- prototipi, dobljeni z algoritmom na podlagi podatkov, ki vsebujejo posebneže, naj se ne bi dosti razlikovali od prototipov, izračunanih na podlagi podatkov brez posebnežev.

Možen način reševanja omenjene naloge je, da se posebneže najprej izloči iz vhodnih podatkov ter šele potem izvede rojenje, vendar je takšna procedura v veliko primerih

praktično nemogoča ali zelo zahtevna. Zaradi tega je Dave [34] predlagal nov algoritem, ki ga je poimenoval "Noise Clustering", NC.

NC algoritem daje zadovoljive rezultate, če so v množici vhodnih vzorcu posebneži. Vendar se v okviru tega dela z omenjenim algoritmom ne bomo posebej ukvarjali. Problematiko rojenja, če so v vhodni množici prisotni posebneži, zadovoljivo rešimo s povečanjem števila rojev. Primer, ki nazorno ilustrira opisano problematiko, bomo obravnavali v nadaljevanju pričujočega dela, glej Sl. 3.14.

Poleg posebnežev je dodaten problem, ki ga je treba rešiti pri rojenju, namreč določitev ustreznega števila rojev ter odvisnost rezultatov rojenja od izbire začetnih prototipov. Oba problema, kakor tudi problem posebnežev, sta povezana s pojmom robustnosti. Navkljub nekaterim predlogom je splošno rešitev problema robustnosti pri neznanem številu rojev iluzorno pričakovati, saj ima vsaka metoda svoje prednosti ter pomanjkljivosti [35].

Ne glede na zgornjo trditev predlaga literatura [35] algoritem, ki obravnava koncept robustnosti mehkega rojenja. Da dosežemo robustno rešitve rojenja moramo izvesti ustrezno število ponovitev rojenja pri različnem številu rojev  $C$  ter uvesti mere veljavnosti (validity measure) rojenja. Mero veljavnosti v splošnem podaja kriterijska funkcija  $f(C, \mathbf{X})$ , katere vrednost je odvisna od števila rojev  $C$  ter vzorcev  $\mathbf{X}$ . Optimalno je tisto rojenje, katerega rezultat je npr. najmanjša vrednost mere veljavnosti. Vrednost mere veljavnosti rojenja pri optimalnem rojenju je odvisna od izbrane mere ali kriterijske funkcije.

### 3.2.1.1. Mere veljavnosti rojenja

Določitev ali izbira ustrezne kriterijske funkcije literatura obsežno obravnava [35], [36]. Nekateri metode merijo stopnjo "mehkosti" v posameznem roju ter predpostavijo, da je najboljša tista delitev, pri kateri je stopnja "mehkosti" najmanjša. Omenjena predpostavka je tudi v skladu z idejo mehkega razvrščanja, saj imajo OD z "najtršo" pripadnostjo najmanjšo negotovost, povezano z njihovo klasifikacijo. Iz tega sledi, da je cilj uporabe mer veljavnosti najti razdelitev, ki minimizira razdelitveno negotovost.

V uvodu razdelka 3.2 smo podali trditev, da je zaradi zahtevnosti rojenja treba razvrščanje za doseganje "optimalnih" rezultatov preizkusiti z več metodami ter primerjati rezultate. Zaradi tega smo izbrali mero veljavnosti rojenja, ki ne temelji na "mehkih" pripadnostih z namenom, da jih uporabimo tudi pri ostalih metodah rojenja. Dodaten razlog, da nismo izbrali mer veljavnosti rojenja, temelječih na "mehkih" pripadnostih, je v zahtevi po "trdi" pripadnosti OD-jev v posameznem roju. Takšna pripadnost je pomembna, ker moramo vedeti, kateri odjemalci sestavljajo končne NOD-e. Podobno, kakor je treba preizkusiti več metod rojenja, je treba za doseganje optimalnih rezultatov uporabiti tudi več mer veljavnosti rojenja. Še zlasti zato, ker niso vse mere enako primerne. Njihova uporabnost je odvisna od podatkov, ki jih moramo rojiti. V nalogi smo preizkusili dve kriterijski funkciji, ki smo ju označili z  $f_1$  in  $f_2$ , za mere veljavnosti rojenja.

Literatura [36] podaja kot mero za določitev ustreznega števila rojev naslednjo kriterijsko funkcijo:

$$f_1(C, \mathbf{X}) = \sum_{i=1}^C \sum_{\mathbf{x} \in S_i} \|\mathbf{x} - \boldsymbol{\beta}_i\|^2, \quad (3.24)$$

kjer je  $S_i$  množica vseh vzorcev razporejenih v  $i$ -ti roj in  $\boldsymbol{\beta}$  prototip roja ali povprečna vrednost vzorcev razporejenih v  $i$ -ti roj.

Če število rojev narašča, kriterijska funkcija monotono pada zato, ker se vsota kvadratov odstopanja zmanjša vsakokrat, ko število rojev povečamo tako, da en vzorec uvrstimo v nov roj. Če  $N$  vzorcev roji v  $C$  strnjениh, dobro razpoznavnih rojih, pričakujemo, da bo vrednost kriterijske funkcije (3.24) hitro padala do naravnega števila rojev  $C$  (funkcija doseže koleno), pozneje pada veliko počasneje vse do vrednosti nič pri  $C = N$ . Optimalno število rojev določimo, ko kriterijska funkcija doseže koleno. Določitev, kdaj je funkcija dosegla koleno, je v določenih primerih subjektivna, saj koleno ni jasno izraženo, zato smo za primerjavo uporabili dodatno kriterijsko funkcijo, ki sta jo predlagala Davies in Bouldin [41]. Kriterijsko funkcijo kot povprečno podobnost rojev določa enačba:

$$f_2(C, \mathbf{X}) = \frac{1}{C} \sum_{i=1}^C \max_{\substack{j=1, \dots, C \\ i \neq j}} \left\{ \frac{\sigma_i + \sigma_j}{D_{ij}} \right\}, \quad C = 2, 3, \dots, N-1, \quad (3.25)$$

kjer je:

$\sigma_i$  – standardni odklon vzorcev v  $i$ -tem in  $j$ -tem roju. Podrobneje ga opisuje enačba (3.26),

$D_{ij}$  – razdalja med prototipoma rojev  $\boldsymbol{\beta}_i$  in  $\boldsymbol{\beta}_j$ . Podrobneje jo opisuje enačba (3.27),

$S_i$  – množica vseh vzorcev v posameznem roju.

$$\sigma_i = \left( \frac{1}{S_i} \sum_{v=1}^{S_i} |\mathbf{x}_v - \boldsymbol{\beta}_i|^2 \right)^{\frac{1}{2}}, \quad i = i, j \quad (3.26)$$

$$D_{ij} = \|\boldsymbol{\beta}_i - \boldsymbol{\beta}_j\| \quad (3.27)$$

Število rojev  $C$ , pri katerem doseže kriterijska funkcija (3.25) najmanjšo vrednost (globalni minimum), imamo za naravno število rojev v dani množici vzorcev.

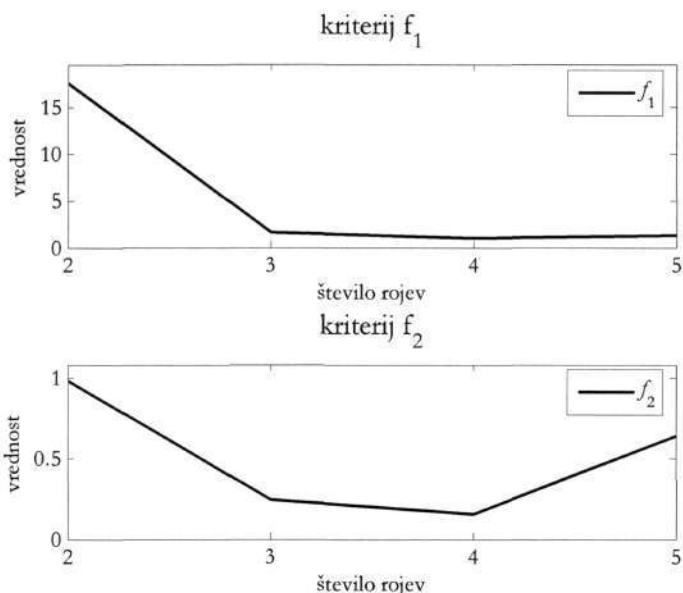
Zgoraj predstavljeni kriterijski funkciji kot meri veljavnosti rojenja zahtevata, da imamo "trdo" pripadnost (ena – pripada ali nič – ne pripada) vzorcev v posameznem roju. FCM algoritem podaja "mehko" pripadnost roju, zato rezultatov ne moremo direktno uporabiti. Pred uporabo mer veljavnosti je treba določiti "trdo" pripadnost. V nalogi smo jo določili na podlagi maksimalne pripadnosti vzorca v določenem roju. Vzorec  $\mathbf{x}_k$  pripada roju  $\boldsymbol{\beta}_i$ , če ima največjo pripadnost  $u_{k,i}$  v množici vzorcev  $S_i$ , ki sestavlja roj  $\boldsymbol{\beta}_i$ :

$$\mathbf{x}_k \in \beta_i \mid \max u_{k,i} \in S_i \quad (3.28)$$

Za jasnejšo ponazoritev uporabe FCM algoritma, vpliva posebnežev na rojenje in zgoraj predstavljenih mer veljavnosti rojenja, bomo v nadaljevanju podali preprost primer, s katerim nazorno predstavimo do sedaj opisane teoretične podlage.

### 3.2.1.2. Ponazoritev uporabe FCM algoritma in mer veljavnosti rojenja

Za potrebe praktične in jasne ponazoritve uporabe FCM algoritma, predstavljenih mer veljavnosti rojenja in vpliva posebnežev na rojenje smo oblikovali naslednji preprost primer. Predpostavimo, da imamo na voljo dva jasno izražena roja, ki vsebujeta vsak po devet vzorcev, ter dva dodatna vzorca, ki predstavljata posebneža. Vse vzorce opisujeta dva parametra  $M = 2$ , v x-y ravnini, sliki Sl. 3.14 in Sl. 3.15. Vzorce smo razvrstili s FCM algoritmom in upoštevali obe meri veljavnosti rojenja, kriterija  $f_1$  in  $f_2$ . V postopku izračuna vrednosti kriterijev smo upošteval "trdo" pripadnost vzorcev v posameznem roju na način, kot ga opisuje enačba (3.28). Število zelenih rojev  $C$  smo spreminjali od dva do pet. Potek vrednosti kriterijskih funkcij  $f_1$  in  $f_2$  v odvisnosti od  $v$  naprej izbranega zelenega števila rojev kaže Sl. 3.13. Iz slike se vidi, da funkciji podajata različno število optimalnih rojev saj, kriterij  $f_1$  doseže koleno pri treh rojih, medtem ko kriterij  $f_2$  doseže minimalno vrednost pri rojenju vzorcev v štiri roje.

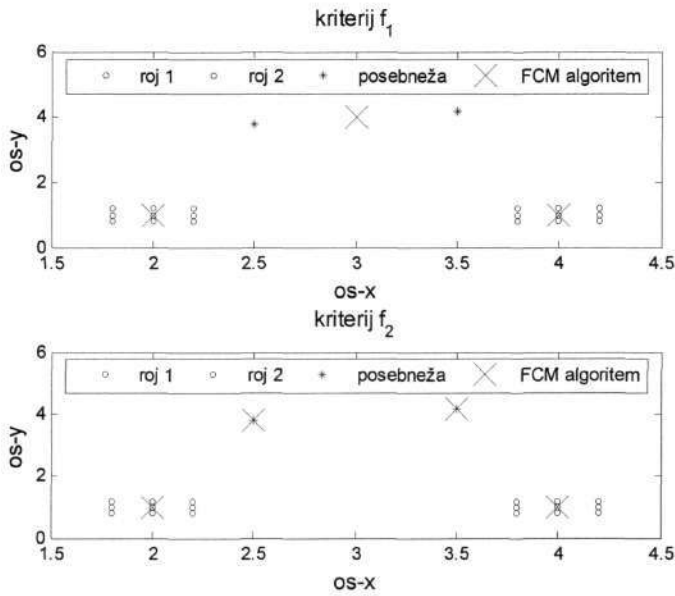


Sl. 3.13: Potek vrednosti kriterijskih funkcij  $f_1$  in  $f_2$  v odvisnosti od števila izbranih rojev.

Na podlagi zgoraj predlaganega števila rojev, določenega s posameznim kriterijem, smo vzorce rojili v tri in štiri roje. Rezultate optimalnega rojenja skupaj z vhodnimi podatki pri uporabi obeh mer veljavnosti rojenja, kriterij  $f_1$  in kriterij  $f_2$  prikazuje Sl. 3.14. Iz Sl. 3.14 se vidi, da rojenje z uporabo mere veljavnosti rojenja dodeli posebnežema ločena roja, medtem ko oba prototipa ležita v centru rojev. Na njuno lego posebneža torej ne vplivata. Primerjava rezultatov rojenja z uporabo dveh različno definiranih mer veljavnosti rojenja kaže, da kriterij

$f_2$  določi vsakemu od posebnježev en roj, medtem ko kriterij  $f_1$  posebnježa združi v en sam roj. Rojenje z uporabo mer veljavnosti rojenja kaže, da je vzorce primerneje razvrstiti v več kot dva roja. Na ta način se izognemo vplivu posebnježev na rojenje.

Na podlagi predstavljenega preprostega primera sestavljenega iz dveh jasno izraženih rojev in dveh posebnježev smo na kratko predstavili uporabo FCM algoritma in mer veljavnosti rojenja. Predpostavili smo, da imamo v naprej podano število rojev,  $C = 2$ . Če bi vzorce razvrstili v predpostavljeno število rojev ter določili "trdo" pripadnost vzorcev posameznemu roju, dobimo nenatančno lego prototipov. Opisano pomanjkljivost odpravimo z uporabo mer veljavnosti rojenja. Na ta način izboljšamo robustnost rojenja. Rezultati rojenja, dobljeni na podlagi preprostega primera, kažejo, da je rojenje optimalno, če posebnježema dodelimo dodaten roj. Rezultat povečanja števila rojev je natančnejša lega prototipov.



Sl. 3.14: Vhodni podatki in lega prototipov pri rojenju s FCM algoritmom in uporabo obeh mer veljavnosti rojenja, kriterij  $f_1$  in  $f_2$ . Optimalno število rojev je tri ali štiri.

V poglavju 3.2.1 smo zapisali, da moramo, pri rojenju vzorcev v skupine ali roje poleg reševanja problematike ustreznega števila rojev in posebnježev analizirati tudi vpliv poljubne izbire začetnih prototipov. V skladu s postopkom delovanja FCM algoritma, ki ga prikazuje Sl. 3.12, naključno generirana pripadnostna matrika  $\mathbf{U}$  ne zagotavlja optimalnega rojenja za vse možne izbire začetnih prototipov  $\mathbf{B}$ , ki jih določimo posredno z uporabo pripadnostne matrike  $\mathbf{U}$  in vzorcev  $\mathbf{X}$ , enačba (3.21). Razlog je v iterativni optimizaciji kriterijske funkcije (3.18), ki je občutljiva na lokalne minimume. Če je število zelenih rojev  $C$  majhno in so "meje" med njimi jasne, naključna inicializacija začetih prototipov  $\mathbf{B}$  vodi k oblikovanju ene same rešitve rojenja, ki je ob enem tudi optimalna. Iterativna optimizacija ne "zapade" v lokalni minimum. S povečevanjem števila rojev  $C$  sam FCM algoritem izgubi na robustnosti. Posledica je množica rešitev rojenja, izmed katerih so optimalne samo nekatere. Robusnost algoritma povečamo z uporabo mer veljavnosti rojenja. Kot smo zapisali že ob koncu razdelka 3.2.1, je optimalno tisto rojenje, ki daje minimalno vrednost izbranega kriterija.



Za doseg optimalne rešitve rojenja je teoretično treba preiskati celotni možni prostor rešitev rojenja. To pomeni, da moramo poiskati vse obstoječe različne kombinacije izbire začetnih prototipov **B** za vse vrednosti zelenih rojev *C*. Omeniti moramo, da z večanjem števila vzorcev, ki jih želimo rojiti, in števila zelenih rojev obsežnost analize zelo hitro narašča [35]. Ne glede na obsežnost teoretične rešitve se v praksi izkaže, da za optimalno rešitev ni treba preiskati celotnega polja mogočih rešitev. To dejstvo potrjuje tudi preprost ilustrativni primer, ki ga podajamo v nadaljevanju.

Prikaza vpliva naključne inicializacije začetnih prototipov **B** na rezultate rojenja bomo izvedli na nekoliko dopolnjenem primeru, ki ga kaže Sl. 3.14. Obravnavanemu primeru smo dodali še en roj in spremenili lego posebnežev. Analizo robustnosti smo izvedli s spreminjanjem vrednosti začetnih prototipov **B** in uporabo mer veljavnosti rojenja za različno število zelenih rojev  $C \in \{3, 4, 5, 6\}$ . Pri tem smo preizkusili deset različnih izbir prototipov. Rezultate analize prikazuje Tab. 3.2. V tabeli smo krepko poudarili tiste vrednosti, kjer imamo optimalno rojenje, minimalna vrednost, za izbrano število zelenih rojev in izbrani kriterij. Rezultate, kjer nastopa globalni optimum, smo uokvirili.

Tab. 3.2: Vrednosti kriterijev  $f_1$  in  $f_2$  pri različnem številu zelenih rojev *C* za deset naključno izbranih inicializacij začetnih prototipov **B**.

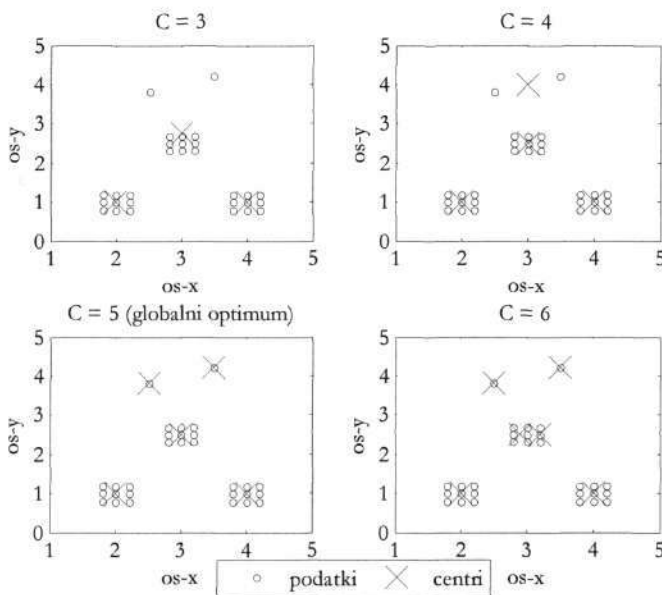
številka izbire	kriterij $f_1$				kriterij $f_2$			
	<i>C</i> = 3	<i>C</i> = 4	<i>C</i> = 5	<i>C</i> = 6	<i>C</i> = 3	<i>C</i> = 4	<i>C</i> = 5	<i>C</i> = 6
1	5,7018	<b>2,02</b>	1,84	1,66	0,43605	<b>0,38459</b>	0,72948	0,98077
2	5,7018	5,5218	1,84	<b>1,26</b>	0,43605	0,8103	0,77078	<b>0,51427</b>
3	5,7018	<b>2,02</b>	1,84	1,66	0,43605	<b>0,38459</b>	0,71781	0,95291
4	5,7018	5,5218	1,84	<b>1,26</b>	0,43605	0,86107	0,67923	0,56338
5	5,7018	<b>2,02</b>	1,84	<b>1,26</b>	0,43605	<b>0,38459</b>	0,71123	0,56338
6	5,7018	<b>2,02</b>	1,84	<b>1,26</b>	0,43605	<b>0,38459</b>	0,77078	0,55257
7	5,7018	5,5218	<b>1,44</b>	<b>1,26</b>	0,43605	0,86107	<b>0,21295</b>	0,54971
8	5,7018	<b>2,02</b>	1,84	1,72	0,43605	<b>0,38459</b>	0,72948	0,61805
9	5,7018	<b>2,02</b>	1,84	1,72	0,43605	<b>0,38459</b>	0,77078	0,66902
10	5,7018	5,5218	<b>1,44</b>	1,72	0,43605	0,8103	<b>0,21295</b>	0,66902

Hitra analiza zgornje tabele kaže, da rojenje vzorcev v *C* = 3 roje ni odvisno od začetne izbire prototipov **B**; vrednost kriterijev je za vse izbire prototipov enaka. Z večanjem števila zelenih rojev se število optimalnih rešitev zmanjšuje. To potrjuje že zapisano dejstvo, da FCM zapade v lokalni minimum in da je algoritem odvisen od izbire prototipov. Da dobimo isto število optimalnih rešitev, moramo povečati število izbir in s tem prostor iskanja optimalnih rešitev. To ni smiselno, saj nas zanima samo ena optimalna rešitev in ne njihova množica. Poleg tega je razlika med optimalnimi rešitvami samo v vrstnem redu rojev, npr. v prvem primeru določene vzorce opisuje ali predstavlja prototip, ki leži na drugem mestu v matriki **B**, medtem ko v drugem primeru iste vzorce opisuje prototip, ki leži na četrtem mestu v matriki **B**.



Izjema je rojenje vzorcev v šest rojev in uporaba kriterija  $f_1$  za mero veljavnosti rojenja. Vidimo, da se je število optimalnih rojenj zopet povečalo, in sicer na pet. Razlog je v izbranih podatkih oz. v njihovi legi. Medsebojna razdalja podatkov je tolikšna, da dve ali več različnih razporeditev rojev dajeta enako vrednost kriterija  $f_1$ ; za enako vrednost kolikor se je povečala ena delna vsota v enačbi (3.24), toliko se je druga zmanjšala. Na podlagi rezultatov v zgornji tabeli ugotovimo, da oba kriterija določite isto začetno izbiro za optimalno.

Opazimo lahko še nekaj. Predstavljena analiza daje tudi odgovor na vprašanje, kolikšno je optimalno število rojev. Iz tabele vidimo, da pet rojev predstavlja globalni optimum. Pri tem imamo minimalno vrednost kriterija  $f_2$ , kriterij  $f_1$  doseže koleno. Osnovne podatke in centre posameznih rojev pri optimalnem rojenju za različne vrednosti  $C = \{3, 4, 5, 6\}$  prikazuje Sl. 3.15.



Sl. 3.15: Lega osnovnih podatkov in centrov optimalnega rojenja za različne vrednosti  $C = \{3, 4, 5, 6\}$  v x-y ravnini. Posebej smo izpostavili globalni optimum rojenja.

Pri postopku izračuna do sedaj prikazanih rezultatov preprostega primera smo predpostavili "trdo" pripadnost vzorca v pripadajočem roju. Pri tem smo naredilo določeno poenostavitev, kateri je treba nameniti nekoliko besed. Analiza pripadnostne matrike  $\mathbf{B}$  kaže, da so vrednosti pripadnosti vzorca  $k$  v roju  $i$ ,  $u_{k,i}$  praktično enake ena, če je vzorec element roja in enake nič, če vzorec ni element roja. Pri tem ne naredimo velike napake, ko določimo "trdo" pripadnost posameznega vzorca izbranemu roju. Rojenje vzorcev v optimalno število rojev podaja relativno jasno pripadnost vzorcev posameznim rojem ter v veliki meri odpravlja probleme s posebnosti.

Zgornje ugotovitve veljajo za predstavljen primer, vendar jih ne moremo v celoti posplošiti, saj je rojenje zelo specifična naloga, odvisna od konkretnih vzorcev, ki jih je treba rojiti. Opozoriti moramo, da koleno ni vedno lahko jasno in enoumno določiti. Ne glede na to, da

ugotovitev ne moremo posplošiti, FCM algoritem z uporabljeno mero veljavnosti daje dobre rezultate za oblikovanje NOD-ov, kar bomo videli tudi v razdelku 5.1.

### 3.2.2. Hierarhične metode in mere podobnosti

Rojenja s hierarhični metodami definiramo podobno kot rojenje z mehкими množicami, kjer predpostavimo, da imamo množico  $N$  vzorcev  $\mathbf{X} = \{\mathbf{x}_k \mid k = 1, \dots, N\}$ , kjer posamezen vzorec  $\mathbf{x}_k = [x_{k,1}, x_{k,2}, \dots, x_{k,M}]$  podaja  $M$  parametrov. Razlika med metodama je, da pri hierarhičnih metodah ni treba vnaprej podati števila rojev  $C$ .

Bistvo aglomerativnih hierarhičnih postopkov rojenja je opis množice vzorcev  $\mathbf{X}$  z zaporedjem  $N$  rojev, ki začne z  $N$  roji (v vsakem roju je natanko en vzorec) in se konča z enim rojem. Stopnje podobnosti med roji v zaporedju rojenj ponazorimo z  $N$  nivojskimi drevesi, ki jih imenujemo dendrogrami [36].

Pri rojenju je veliko možnosti, kako meriti podobnost med vzorci. Pokazalo se je, da se v večini primerov uporablja kot mera podobnosti razdaljo med vzorci [39]. Razdaljo  $d(\mathbf{x}_k, \mathbf{x}_l)$  med  $M$ -dimenzionalnima vzorcema  $\mathbf{x}_k$  in  $\mathbf{x}_l$  najpogosteje merimo z eno izmed oblik razdalje Minkovskega, ki jo podaja naslednja enačba:

$$d(\mathbf{x}_k, \mathbf{x}_l) = \left( \sum_{i=1}^M |x_{k,i} - x_{l,i}|^n \right)^{\frac{1}{n}} \quad (3.29)$$

kjer je  $n \in (0, \infty)$  poljuben parameter.

V odvisnosti od izbire vrednosti parametra  $n$  ločimo naslednje posebne primere razdalje Minkovskega:

- Manhattan ali "City block" ( $n = 1$ ),
- Evklidska ( $n = 2$ ),
- Čebiševa ( $n \rightarrow \infty$ ),
- poleg omenjenih izpeljank se lahko uporabljata tudi razdalja "Canberra" ali koeficient Czekanowskega.

V splošnem lahko uporabljamo vse razdalje, ki zadostujejo pogojem metrike [39]:

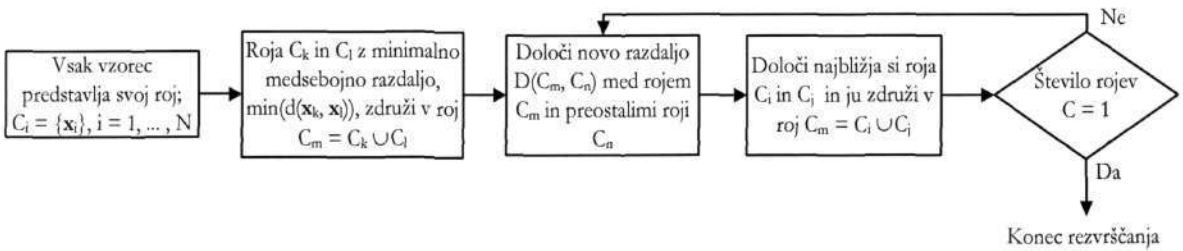
$$\begin{aligned} d(\mathbf{x}_k, \mathbf{x}_l) &= d(\mathbf{x}_l, \mathbf{x}_k) && \text{simetričnost,} \\ d(\mathbf{x}_k, \mathbf{x}_l) &> 0, \text{ če je } \mathbf{x}_k \neq \mathbf{x}_l && \text{nenegativnost,} \\ d(\mathbf{x}_k, \mathbf{x}_l) &= 0, \text{ če je } \mathbf{x}_k = \mathbf{x}_l && \text{razločljivost,} \\ d(\mathbf{x}_k, \mathbf{x}_l) &\leq d(\mathbf{x}_k, \mathbf{x}_z) + d(\mathbf{x}_z, \mathbf{x}_l) && \text{trikotniška neenakost.} \end{aligned} \quad (3.30)$$

Najpogosteje se uporablja Evklidska razdalja med vzorci, ki smo jo uporabljali v disertaciji in jo uporabljajo tudi druge metode: FCM algoritem in nevronska omrežja. Omenjeno razdaljo podaja enačba (3.19), zapišemo jo tudi v naslednji obliki:

$$d(\mathbf{x}_k, \mathbf{x}_l) = \|\mathbf{x}_k, \mathbf{x}_l\| = \sqrt{(\mathbf{x}_k - \mathbf{x}_l)^T \cdot (\mathbf{x}_k - \mathbf{x}_l)} \quad (3.31)$$

Na podlagi izračunanih razdalj med vzorci lahko določimo roje z grupiranjem vzorcev v odvisnosti od podobnosti. Zaradi ogromno različnih kombinacij združevanja vzorcev so se oblikovali postopki rojenja, ki temeljijo na zaporedju.

Prvi člen zaporedja, ki je najnižji v hierarhiji rojenja, predstavlja delitev vzorcev  $\mathbf{X}$  v  $C = N$  rojev. V vsakem roju je natanko en vzorec. Zadnji člen zaporedja predstavlja en sam roj in je najvišji v hierarhiji. V vsakem koraku rojenja združimo dva najbolj podobna roja v en sam večji roj. Oba se v naslednjih korakih rojenja ne razdružita več. Po združitvi dveh rojev  $C_k$  in  $C_l$  v nov roj  $C_m$  je treba izračunati razdalje novega roja do ostalih rojev  $C_n$ . Različni postopki rojenja se razlikujejo med seboj v načinu izračuna razdalj  $D(C_m, C_n)$  med roji, vendar vsi temeljijo na postopku, ki ima za predpogoj, da se izračuna razdalja med vzorci  $d(\mathbf{x}_k, \mathbf{x}_l)$  [36], [39]. Postopek prikazuje Sl. 3.16.



Sl. 3.16: Prikaz postopka hierarhičnega rojenja.

Običajne razdalje med roji so [36], [39]:

- minimalna razdalja,
- maksimalna razdalja,
- povprečna razdalja,
- Wardova razdalja.

Ker minimalna razdalja tvori veriženje ali krči prostor ter maksimalna metoda širi prostor [29], smo v disertaciji preizkusili dva algoritma rojenja vzorcev, ki imata za podlogo povprečno (3.32) in Wardovo (3.33) razdaljo med roji:

$$D(C_r, C_s) = \frac{\sum_{i=1}^{n_r} \sum_{k=1}^{n_s} d(\mathbf{x}_{r,i}, \mathbf{x}_{s,k})}{n_r \cdot n_s}, \quad (3.32)$$

$$D(C_r, C_s) = \frac{n_r \cdot n_s \cdot d^2(\bar{x}_r, \bar{x}_s)}{n_r + n_s}, \quad (3.33)$$

kjer je:

$n_r$  – število vzorcev v združenem roju  $C_r$ ,

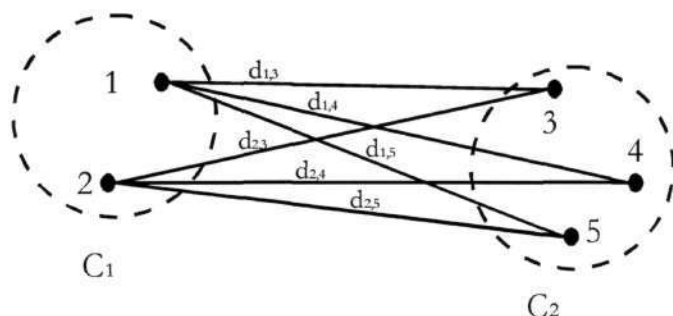
$n_s$  – število vzorcev v roju  $C_s$ ,

$d(x_{r,i}, x_{s,k})$  – razdalja med  $i$ -tim vzorcem roja  $C_r$  ter  $k$ -tim vzorcem roja  $C_s$ .

$d(\bar{x}_r, \bar{x}_s)$  – razdalja med središčema rojev  $r$  in  $s$ , ki jo zapišemo kot sledi:

$$\bar{x}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} x_i, \quad j = r, s. \quad (3.34)$$

Izračun povprečne razdalje med dvema rojema bomo ponazorili na preprostem primeru, sestavljenem iz dveh rojev. Prvi roj  $C_1$  sestavljata dva vzorca, medtem ko drugi roj  $C_2$  sestavljajo trije vzorci. Primer grafično prikazuje Sl. 3.17. Na sliki smo označili samo razdalje med vzorci sosednjih rojev, ki jih izračunamo z enačbo (3.29).



Sl. 3.17: Ponazoritev preprostega primera izračuna povprečne razdalje med dvema rojema, sestavljena iz dveh in treh vzorcev.

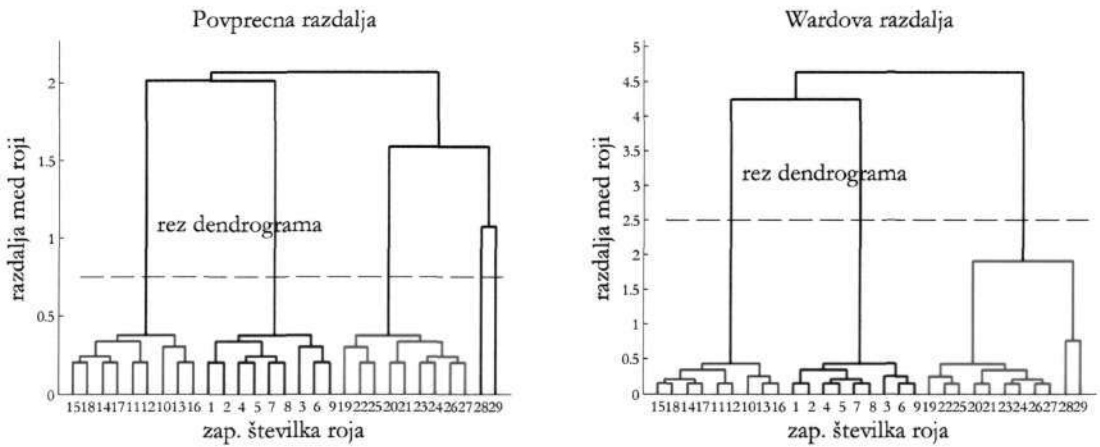
Povprečno razdaljo med rojema  $C_1$  in  $C_2$  na zgornji sliki izračunamo na naslednji način:

$$D(C_1, C_2) = \frac{d_{1,3} + d_{1,4} + d_{1,5} + d_{2,3} + d_{2,4} + d_{2,5}}{6} \quad (3.35)$$

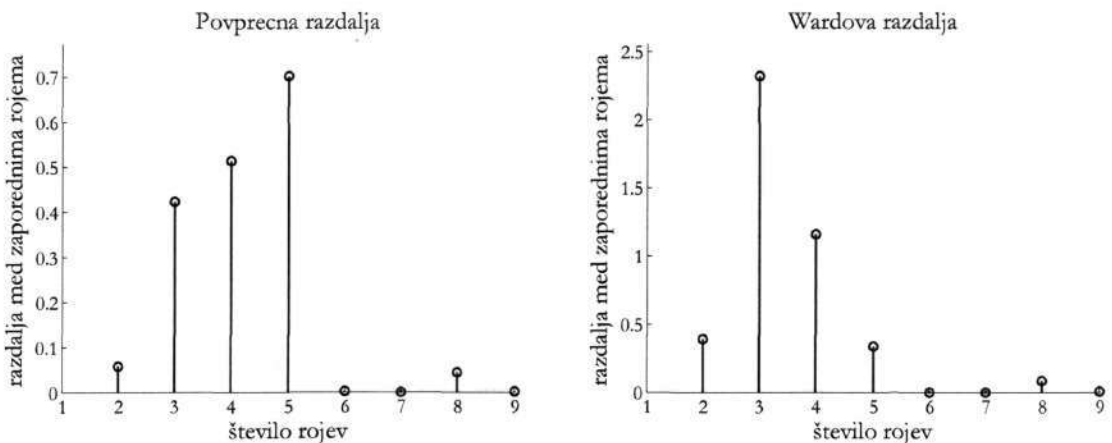
Rezultat rojenja lahko grafično ponazorimo z drevesom, ki ga imenujemo dendrogram. Sl. 3.18 ponazarja hierarhično rojenje vzorcev, predstavljenih na Sl. 3.15 za povprečno in Wardovo razdaljo med roji. Navpične črte rišemo sorazmerno razdalji rojev, ki se na tej stopnji združujejo v en roj. Krajše kot so navpične črte, bolj so vzorci v roju podobni; manjša je njihova medsebojna razdalja.

Enako, kot pri razvrščanju vzorcev z nehierarhičnimi metodami, npr. s FCM algoritmom, kot smo predstavili v razdelku 3.2.1, se tudi pri hierarhičnih metodah srečamo z vprašanjem optimalnega razreza dendrograma in določitve optimalnega števila rojev. Literatura [36] navaja, da rezanje navpičnih črt, kjer imamo največjo razdaljo med zaporednima rojema,

odkrije "naravne" roje v vzorcih. Tako določimo končno število rojev, v katere bomo podatke rojili. Določanje končnega števila rojev na podlagi razdalj med sosednjimi roji samo s pregledom dendrograma je pri velikem številu podatkov praktično nemogoče. Pomagamo si z grafičnim prikazom vrednosti razdalj med zaporednima rojema v odvisnosti od končnega števila rojev, ki ga prikazuje Sl. 3.19. Vidimo, da algoritem rojenja, ki uporablja povprečno razdaljo med roji, določa pet rojev kot optimalno število, medtem ko algoritem z Wardovo razdaljo predlaga tri optimalne roje. Razvrstitev vzorcev v optimalno število rojev, kakor jih predlaga posamezen algoritem, kažeta tudi dendrograma na Sl. 3.18. Na posameznem dendrogramu smo z vodoravno črtkano črto označili mesto, kjer smo dendrogram prerezali. Iz Sl. 3.18 se vidi, da uporaba povprečne razdalje med roji natančno pokaže tri jasno izražene roje in posebnega v vhodnih podatkih katerima dodeli lastna roja. Uporaba Wardove razdalje predlaga rojenje v tri roje in ne pokaže posebnega. Združi ju v njima najbližji roj. Ugotovimo, da je uporaba povprečne razdalje med rojema primerna za izločanje posebnega.



Sl. 3.18: Dendrogram rojenja vzorcev, Evklidska razdalje med vzorci ter povprečna in Wardova razdalja med roji.



Sl. 3.19: Razdalja med zaporednima rojema pri izračunu povprečne in Wardove razdalje med roji.

Postopek določanja končnega števila rojev z rezanjem dendrograma pri maksimalni razdalji med sosednjima rojema ne daje vedno najboljših rešitev, npr. postopek predlaga zelo majhno število končnih rojev, ki imajo velik raztros. Takrat poiščemo naslednjo največjo razdaljo med sosednjima rojema in analiziramo rezultate rojenja. Število rojev pri tej razdalji predstavlja naslednjo možno vrednost končnih rojev. To vidimo na Sl. 3.19, kjer se razdalja med sosednjima rojema pri izračunu povprečne razdalje med roji zopet poveča pri rojenju vzorcev v osem rojev.

Podrobnejša analiza in primerjava hierarhičnega rojenja z rezanjem največjih razdalj med sosednjima rojema dendrograma, glej Sl. 3.18, in FCM rojenja z upoštevanjem mer veljavnosti rojenja, glej Tab. 3.2, kaže naslednje. Povprečna razdalja med roji pri hierarhičnih metodah in kriterij  $f_2$  (3.25) pri FCM algoritmu natančno identificirata posebneža in jima določita lastna roja, kriterij  $f_1$  (3.24) pri FCM algoritmu posebneža združi v en skupen roj, medtem ko uporaba Wardove razdalje predlaga združitev posebnežev z najbližjim rojem. Ugotovitve veljajo za podatke na Sl. 3.15, ki so opisani samo z dvema podatkom v x-y ravnini in jih ne moremo posplošiti na poljubne podatke, da bi določili splošni kriterij, katera metoda in uporaba katere mere veljavnosti rojenja daje najboljše rezultate. Najprej moramo narediti analizo in na podlagi rezultatov določiti najprimernejšo število rojev in metodo rojenja. Pri tem moramo upoštevati cilj ali namen rojenja, npr. iskanje in izločanje posebnežev ali določanje splošnejših rojev, ki opisujejo širšo množico vzorcev.

### 3.2.3. Nevronska omrežja

Na nevronska omrežja (NO) lahko gledamo kot na množico paralelnih sistemov, sestavljenih iz velikega števila preprostih procesorjev z mnogo povezavami. Njihova glavna karakteristika je, da imajo sposobnost naučiti se kompleksnih nelinearnih odnosov med vhodi in izhodi ter se prilagoditi vhodnim podatkom [37].

Izmed različnih vrst NO-jev smo v nalogi uporabljali dve: samo-organizirajoča omrežja (ang. Self-Organising Map networks, SOM networks) in verjetnostna nevronska omrežja, VNO, (ang. Probability Neural Networks). Njihove osnovne značilnosti podajamo v naslednjih dveh razdelkih.

#### 3.2.3.1. Samo-organizirajoča omrežja

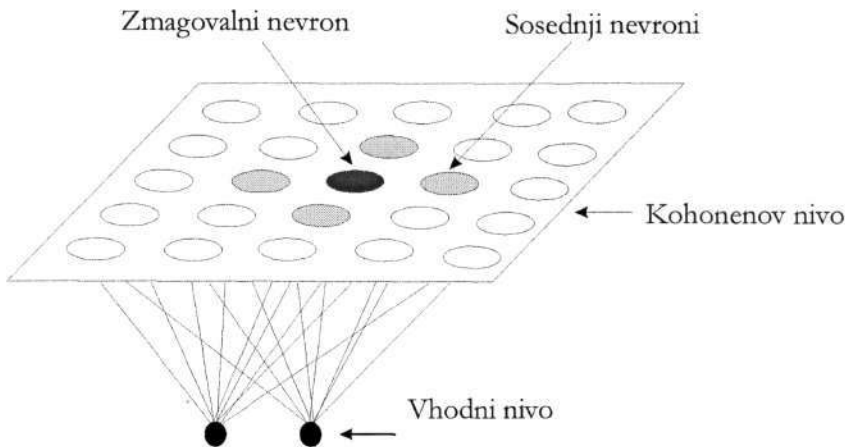
Samo-organizirajoča omrežja (SOM) so ena izmed oblik NO-jev, ki so namenjene predvsem za nenadzirano razvrščanje vzorcev. Sposobna so najti posamezne roje v podatkih in zgraditi urejen zemljevid, ki odraža podobnosti v množici vhodnih vzorcev. Osnovni gradniki SOM-a so nevroni, ki so navadno organizirani na način, da sestavljajo dvodimenzionalno mrežo. Za razliko od ostalih omrežij, SOM-i omogočajo medsebojne bočne povezave med nevroni. Ločimo več vrst oblik samo-organizirajočih mrež, ki se ločijo med seboj glede na lego nevronov in njihovimi medsebojnimi povezavami. Poznamo mreže s pravokotno, heksagonalno ali naključno razporejenimi nevroni. Njihova fizična lokacija na zemljevidu odraža relativne



podobnosti med vzorci. Na ta način ima načrtovalec možnost, da vizualno analizira ter razloži odnose med točkami ter s tem odnose med vzorci.

Načrtovanje in uporabo SOM omrežij lahko razdelimo na dva načina: prvi določa, da je število nevronov v omrežju enako številu rojev. Na primer, omrežje velikosti  $2 \times 2$  ima štiri nevrone, zato določa štiri roje. Pri drugem načinu izberemo število nevronov, ki je večje od števila rojev. Pri tem moramo narediti dodatno analizo rezultatov učenja SOM-a, da določimo ustrezno število rojev. Vizualno moramo pregledati zemljevid omrežja in na podlagi njegove medsebojne povezanosti določiti nevrone, ki so se med potekom učenja grupirali in tako določili podobne skupine.

SOM omrežja imajo dva nivoja, vhodnega ter Kohonenovega. Vhodni nivo je v celoti povezan z dvodimenzionalnim Kohonenovim nivojem. Shematsko strukturo SOM-a z zmagovalnim nevronom in njegovo okolico z radijem  $R = 1$  prikazuje Sl. 3.20. Med procesom učenja vhodni vzorci prehajajo v omrežje po procesnih elementih, nevronih v vhodnem nivoju. Množica  $N$  vzorcev  $\mathbf{X} = \{\mathbf{x}_k \mid k = 1, \dots, N\}$ , kjer ima posamezen vzorec  $\mathbf{x}_k = [x_{k,1}, x_{k,2}, \dots, x_{k,M}]$   $M$  parametrov, se primerja z nevroni v dvodimenzionalnem Kohonenovem nivoju. Vsakemu nevronu je pridružen utežni vektor  $\mathbf{w}_i = [w_{i,1}, w_{i,2}, \dots, w_{i,M}]$  prav tako reda  $M$ , kjer je  $w_{i,j}$  vrednost uteži nevrone  $i$  v povezavi z  $j$ -tim parametrom vhodnega vzorca.



Sl. 3.20: Struktura SOM-a z zmagovalnim nevronom in okolico z radijem  $R = 1$ .

Z učenjem nevroni posodablajo vrednost uteži glede na topološke odnose med vhodnimi vzorci. Neuron z minimalno razdaljo (3.36) do vhodnega vzorca, t.j. zmagovalni ali vzbujeni nevron  $i^*$ , prilagodi utež tako, da je bližje vhodnemu vzorcu.

$$d_{k,i^*} = \min_i \{ \|\mathbf{x}_k - \mathbf{w}_i\| \} \quad (3.36)$$

Poleg spremembe uteži zmagovalnega ali vzbujenega nevrone algoritem prilagodi vrednosti uteži tudi vsem nevronom v njegovi okolici  $O_{i^*}(d)$ . Neuron  $i$  se nahaja v okolici  $O_{i^*}(d)$  vzbujenega nevrone  $i^*$ , če leži znotraj radija  $R$ :

$$i \in O_{i^*}(d) \mid d_{i,i^*} \leq R, \quad (3.37)$$

kjer je  $d_{i,i^*}$  razdalja sosednjih nevronov do vzbujenega nevrona.

Programsko orodje [42], ki smo ga uporabljali v nalogi, omogoča določanje različnih razdalj med nevroni, izmed katerih so tudi Evklidska in Manhattan razdalja. Da bi čim boljše primerjali v disertaciji obravnavane metode rojenja vzorcem, smo uporabili Evklidsko razdaljo določanja razdalj med nevroni.

V vsaki iteraciji  $t$ , ko damo na vhod omrežja vzorec  $\mathbf{x}_k$ , se uteži  $\mathbf{w}$  vzbujenega nevrona  $i^*$  ter nevronov v njegovi okolici  $O_{i^*}(d)$  prilagodijo v skladu s Kohonenovim pravilom, katerega za  $i$ -ti nevron definira enačba:

$${}_i\mathbf{w}(t) = {}_i\mathbf{w}(t-1) + b(t) \cdot (\mathbf{x}_k - {}_i\mathbf{w}(t-1)). \quad (3.38)$$

Drugi člen enačbe (3.38) predstavlja spremembo vrednosti uteži  ${}_i dw(t)$ , ki jo določimo na naslednji način:

$${}_i dw(t) = \alpha(t) \cdot g(A,R) \cdot (\mathbf{x}_k - {}_i\mathbf{w}(t-1)), \quad (3.39)$$

kjer:

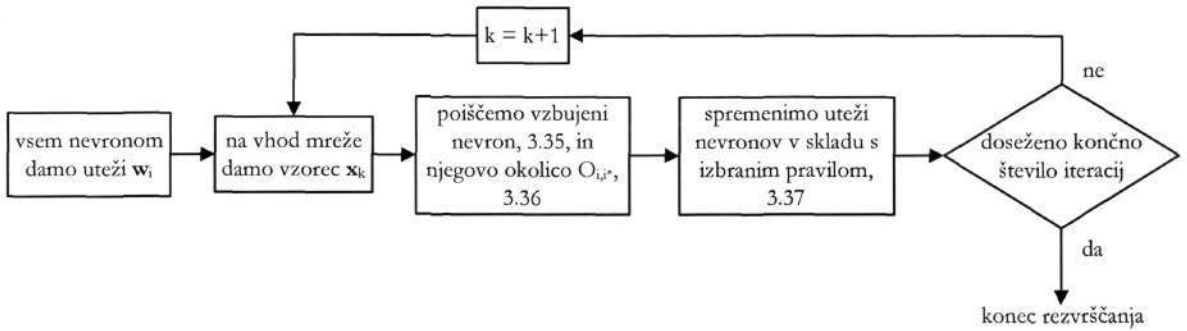
$\alpha(t)$  – predstavlja mero učenja, ki se s časom zmanjšuje in zavzema vrednosti med nič in ena ( $0 < \alpha(t) < 1$ ),

$g(A,R)$  – predstavlja aktivacijski faktor, ki zavisi od vzbujenosti nevrona  $A$  ter velikosti njegove okolice  $R$ . Okolica  $R$  ki se s časom zmanjšuje. Zadnji faktor predstavlja razdaljo med vhodnim vektorjem ter utežmi vzbujenega nevrona.

Odvisnost aktivacijskega faktorja, ki smo ga uporabljali v delu, od parametrov  $A$  in  $R$  podaja enačba (3.40) in zavzema vrednosti ena, če je nevron vzbujen ( $A = 1$ ),  $\frac{1}{2}$ , če se nevron nahaja znotraj okolice  $O_{i^*}(d)$  vzbujenega nevrona ter nič v vseh ostalih primerih [41].

$$g(A,R) = \begin{cases} 1, & \text{če je } A = 1, \\ \frac{1}{2}, & \text{če je } A = 1 \text{ in } d_{i,i^*} \leq R, \\ 0, & \text{za ostale primere.} \end{cases} \quad (3.40)$$

Na ta način se uteži nevronu hitro spreminjajo proti vrednostim vhodnega prostora, kjer se nahajajo vhodni vzorci. Podobni vzorci povzročijo, da se uteži sosednjim nevronom spreminjajo v isti smeri ter s tem povzročijo grupiranje nevronov in s tem njihovo topološko orientacijo v prostoru. Nevroni SOM-a se med procesom učenja orientirajo približno z enakomerno medsebojno razdaljo, če se vzorci pojavljajo z enakomerno porazdelitvijo v določenem delu vhodnega prostora. Podobno velja tudi, če se vhodni vzorci ponavljajo z ustrešno frekvenco. Tako se SOM med procesom učenja nauči kategorizirati topologijo kakor tudi porazdelitev vzorcev [41]. Diagram potek učenja SOM-a prikazuje naslednja slika.



Sl. 3.21: Diagram poteka učenja SOM-a.

Rezultat učenja je omrežje, ki je sposobno rojiti vhodne vzorce. Po končanem učenju je treba omrežje simulirati. Simulacije dajeje vektor, ki podaja indeks vzbujenega nevrona. Vsak nevron predstavlja eno izmed končnih skupin ali je del nje, če imamo večje število nevronov kot zelenih rojev ali skupin. Ko vemo, kateri nevroni sestavljajo posamezen roj, lahko določimo tudi to, kateri vzorci pripadajo izbranemu roju. SOM podaja jasno ali "trdo" pripadnost vzorca posameznemu roju.

Za doseganje optimalnih rezultatov razvrščanja vzorcev s SOM-om moramo uporabiti mere veljavnosti rojenja, npr. tiste, ki smo spoznali že v razdelku 3.2.1.1. Pri tem moramo biti nekoliko pozorni, saj ločimo dva postopka. Če SOM sestavlja več nevronov, kakor je zelenih končnih skupin, uporabimo mere veljavnosti rojenja za določanje, kateri nevroni sestavljajo posamezno skupino in kolikšno naj bi bilo optimalno število skupin. V drugem primeru določamo število skupin neposredno, zato moramo določiti samo optimalno število nevronov. Prvi postopek je zahtevnejši. Poleg tega bi te rezultate težje primerjali z rezultati ostalih dveh postopkov rojenja opisanih v razdelku 3.2.1 in 3.2.2. Brez pomena je primerjati rojenje istih vzorcev v različno število skupin.

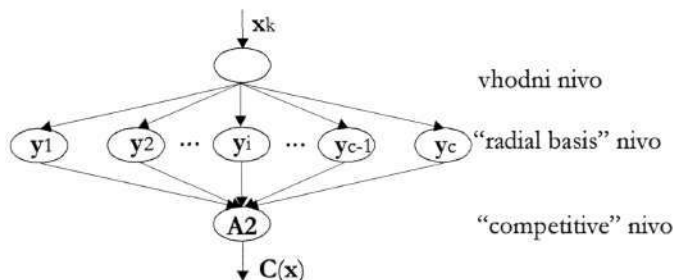
Do sedaj smo spoznali teoretične podlage in nekatere lastnosti metod rojenja vzorcev v roje ali skupine in mere veljavnosti rojenja, ki jih bomo v disertaciji uporabili. V nadaljevanju predstavljamo še eno izmed metod, ki sodi v široko področje razpoznavanja vzorcev. To so verjetnostna nevronska omrežja. Na tem mestu jih predstavljamo, ker predstavljajo način razpoznavanja vzorcev, temelječ na nevronskih omrežjih, vendar jih za rojenje OD-jev nismo uporabili. Njihova koristnost se je pokazala pri dodeljevanju NOD-ov posameznim UO-jem, kar bomo podrobneje prikazali v razdelku 3.3.

### 3.2.3.2. Verjetnostna Nevronska Omrežja

Verjetnostna nevronska omrežja [42], [43], VNO, so ena izmed oblik "vnaprejšnjih" omrežij (ang. feed forward networks). Značilno zanje je, da se odzivajo na vhodne vzorce s procesiranjem podatkov od enega nivoja k drugemu brez vzvratne zveze. Temeljijo na ocenjevanju gostote verjetnosti in se uporabljajo predvsem za razvrščanje vzorcev v vnaprej znane skupine (nadzirano razvrščanje). Njihovo shematsko strukturo prikazuje Sl. 3.22.

Sestavljene so iz več medsebojno povezanih nevronov v zaporednih nivojih; vhod, radial basis nivo – prvi nivo ter tekmovalni nivo (ang. competitive layer) – drugi nivo.

Vhod distribuirira vhodni vzorec  $\mathbf{x}_k = [x_{k,1}, x_{k,2}, \dots, x_{k,M}]$ , kjer je  $M$  število parametrov vzorca, "radial basis" nivoju, kjer število nevronov ustreza številu "učnih" vzorcev  $\mathbf{Y} = [y_1, y_2, \dots, y_C]$ . Vrednosti uteži  $w_{1,i}$  nevronov v prvem nivoju zavzemajo vrednosti "učnih" vzorcev  $y_i = [y_{i,1}, y_{i,2}, \dots, y_{i,M}]$ , kjer  $i$  predstavlja njihovo število,  $i = 1, \dots, C$ .



Sl. 3.22: Shematska struktura verjetnostnega nevronskega omrežja.

Ko damo na vhod VNO-ja nov vzorec  $\mathbf{x}_k$ , omrežje izračuna izhod iz "radial basis" nivoja, ki ga predstavlja vektor  $\mathbf{a1}_k$ , katerega  $i$ -ti element izračunamo z naslednjo enačbo:

$$a_{1,k,i} = \exp \left[ - \left( d_i(\mathbf{x}_k, \mathbf{w}_{1,i}) \cdot b_i \right)^2 \right], \quad (3.41)$$

kjer je:

$d_i(\mathbf{x}_k, \mathbf{w}_{1,i})$  – Evklidska (3.19) razdalja vhodnega vzorca  $\mathbf{x}_k$  do uteži  $\mathbf{w}_{1,i}$ ,

$b_i$  – sprožilni prag nevrona, ki definira njegovo občutljivost. Izračunamo ga kot sledi:

$$b_i = \text{sqrt} \left[ \frac{-\log(0.5)}{\mu} \right], \quad (3.42)$$

kjer je  $\mu$  občutljivostni parameter, ki ga definira uporabnik.

Gleda na enačbo (3.41), posamezni vhodni vzorec  $\mathbf{x}_k$ , ki je "bližje" določenemu učnemu vzorcu in ima manjšo razdaljo  $d_i(\mathbf{x}_k, \mathbf{w}_{1,i})$ , ga v vektorju  $\mathbf{a1}_k$  predstavlja vrednost bližje ena. Rezultat prvega nivoja, ko damo na vhod vse vhodne vzorce, ki jih želimo razvrstiti, je matrika  $\mathbf{A1}$  dimenzij  $C * N$ , kjer  $N$  predstavlja število vhodnih vzorcev.

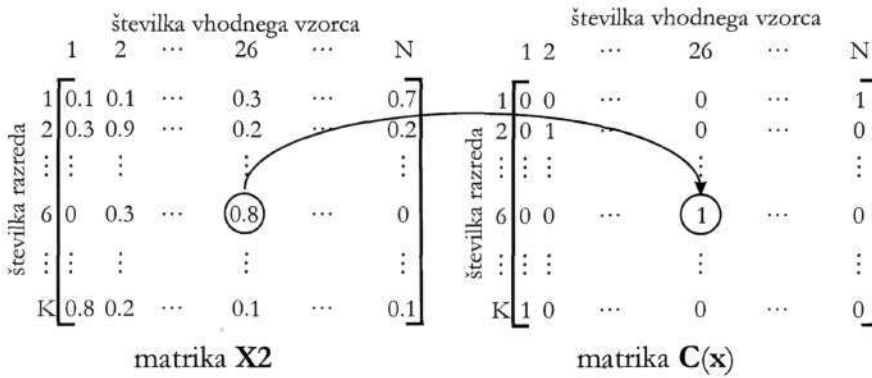
Število nevronov  $K$  v drugem nivoju določa število razredov "učnih" vzorcev in predstavlja kar njihovo dejansko število  $K = C$ . Tako vsak nevron predstavlja enega od razredov, v katerega želimo razvrstiti vhodne vzorce. Vrednosti uteži  $\mathbf{w}_2$  v drugem nivoju so postavljene v skladu z vrednostmi, ki jih podaja ciljna matrika  $\mathbf{T}$ , dimenzije  $K * K$ . Vsak vektor ciljne matrike zavzema vrednosti ena samo v vrsticah, povezanih z vhodnim razredom, druge imajo vrednosti nič. Na primer, utež tretjega nevrona, ki predstavlja tretji razred, zavzame vrednost

ena samo v tretji vrstici, drugje ima vrednost nič. Izhod drugega nivoja predstavlja matrika  $A_2$  dimenzije  $K * N$ , ki jo izračunamo kot sledi:

$$A_2 = T \cdot A_1 \quad (3.43)$$

Prenosna funkcija drugega nivoja se imenuje "compete". Njena naloga je, da pretvori matriko  $A_2$  tako, da določi vrednost ena v izhodni matriki  $C(x)$  samo na tistih mestih, kjer ima posamezen stolpec matrike  $A_2$  maksimalno vrednost. Postopek vidimo na sliki Sl. 3.23, ki shematično prikazuje delovanja prenosne funkcije "compete". Na primer, matrika  $X_2$  v 26. stolpcu zavzema maksimalno vrednost v vrstici šest. Prenosna funkcija "compete" dodeli vrednost ena istoležnemu elementu matrike  $C(x)$ . Pomen dodelitve je, da VNO klasificirajo 26. vzorec iz množice vhodnih vzorcev  $X = [x_1, x_2, \dots, x_N]$  v šesti razred. Na takšen način omrežje razvrsti vhodne vzorce v tiste razrede, ki imajo največjo verjetnost, da jim vzorec pripada.

Načrtovanje VNO-jev je relativno enostavno ter neodvisno od učenja, kakor pri ostalih tipih NO-jev. Njihove uteži predstavljajo prototipe vzorcev, na podlagi katerih razvrščamo ostale vzorce, zato ni treba prilagajati uteži in pragov. Uporabljamo jih lahko sproti in ko podamo učne vzorce, lahko omrežje začne razvrščati nove vzorce. Če imamo dovolj učnih vzorcev, VNO-ji konvergirajo k Bayes-ovi strategiji razvrščanja. Poleg tega imajo veliko sposobnost posploševanja.



Sl. 3.23: Shematski prikaz delovanja prenosne funkcije "compete". Transformacija matrike  $X_2$  v matriko  $C(x)$ .

Največjo pomanjkljivost VNO-jev je v tem, da zahtevajo en nevron za vsak učni vzorec in da so počasnejše kot ostali tipi NO-jev, saj zahtevajo veliko izračunov. Treba je izračunati razdaljo vsakega vhodnega vzorca  $x_k$  do vseh "učnih" vzorcev, kar z večanjem števila vhodnih ali "učnih" vzorcev zahteva obsežnejše izračune. Kljub temu VNO-ji predstavljajo uporabno orodje za nadzirajoče razvrščanje, ko ne potrebujemo zelo hitrih računskih operacij ali imamo majhno število vhodnih oziroma "učnih" vzorcev. Takšna je tudi uporaba VNO-jev v pričujočem delu.

### 3.3. Dodelitev NOD-ov posameznim UO-jem

#### 3.3.1. Značilnosti odjemalčeve dejavnost

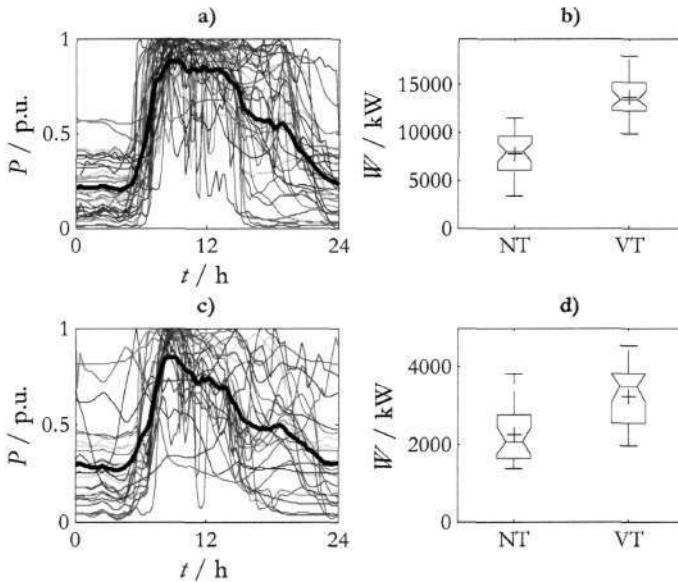
V prejšnjih dveh razdelkih 3.1 in 3.2 smo podrobneje predstavili potek oblikovanja podobnih skupin ali rojev z uporabo valčne analize in metod nenadziranega rojenja. Zadnji korak predstavlja določitev NOD-e in njihova dodelitev posameznim UO-jem. V pregledu literature s področja oblikovanja NOD-ov smo videli, da so avtorji, ki so uporabljali metode rojenja, določali samo tipične skupine odjemalcev v odvisnosti od oblike njihovih OD-jev. Ob tem ne predlagajo posebnega postopka, ki bi bil enostaven in učinkovit. Dodatno merjenje OD-jev od tri do šest mesecev [24], [25], po našem mnenju, ne predstavlja časovno učinkovite in cenovno ugodne metode. Menili smo, da področje dodeljevanja in s tem oblikovanja NOD-jev še ni bilo zadovoljivo rešeno. Zato smo si zadali nalogo, da oblikujemo primeren postopek, ki bi odpravil to pomanjkljivost. Pri tem smo za podlago vzeli podatke o posameznem UO-ju, ki so razpoložljivi in hitro dostopni. Oprli smo se na tiste podatke, ki bi doprinesli k oblikovanju splošnega pravila dodeljevanja OD-jev. Analizirali smo naslednje podatke, ki jih distribucijska podjetja imajo o UO-jih in so bili z vidika dodeljevanja OD-jev za nas zanimivi:

- registrirana dejavnost,
- porabljena mesečna energija v visoki tarifi  $W_{VT}$  ter nizki tarifi  $W_{NT}$ ,
- mesečna konična moč  $P_k$ .

Zadnja dva podatka sta po našem mnenju manj uporabna. Konična moč  $P_k$  je samo ena vrednost znotraj celega meseca in ne pove nič o obliki odjemalčevega OD-ja. V nekaterih primerih je celo zavajajoča, glej Sl. 3.10. Porabljeni mesečni energiji v visoki  $W_{VT}$  in nizki  $W_{NT}$  tarifi nosita več informacij o odjemalčevem OD-ju kakor njegova konična moč, vendar imajo odjemalci s podobnim razmerjem med obema energijama zelo različne OD-je. S povprečenjem moči odjemalca znotraj definiranega intervala, kar merjenje energije v nekoliko drugačni interpretaciji tudi predstavlja, izgubimo informacijo o celotni dinamiki porabe in s tem o obliki OD-ja. Primer razvrščanja odjemalcev na podlagi porabljenih mesečnih energij prikazuje Sl. 3.24. Na sliki vidimo normirane OD-je in njihovo povprečno vrednost dveh skupin odjemalcev, ki imajo podobno mesečno porabo električne energije, grafa a) in c) ter obseg, mediano in povprečno vrednost (označili smo jo s +) porabljene električne energije v obeh tarifah, grafa b) in d). Iz slike vidimo, da imajo odjemalci s podobno, relativno homogeno porabo energije zelo različne OD-je. Razlog je v načinu ali dinamiki odjema in posledično obliki OD-ja. Odjemalci, ki obratujejo dlje časa pri nižji moči, porabijo enako količino energije kakor odjemalci, ki obratujejo krajši čas z višjo močjo. Primerjava grafov a) in c) ter b) in d) kaže, da imajo skupine odjemalcev z zelo različno mesečno porabo energije podobne povprečne OD-je. Druga skupina odjemalcev na Sl. 3.24 porabi približno tretjino energije odjemalcev prve skupine, vendar sta njuna povprečna diagrama zelo podobna.



Pomanjkljivosti kažeta, da dodeljevanje in oblikovanje NOD-ov na podlagi primerjanja mesečne porabe energije ne daje zadovoljivih rezultatov. Ker nas zanima oblika OD-jev, ki jih odjemalci imajo, menimo, da takšen postopek dodeljevanja ni primeren.



Sl. 3.24: Primerjav oblik OD-jev pri razvrščanju odjemalcev na podlagi porabljene mesečne energije v nizki tarifi NT in visoki tarifi VT.

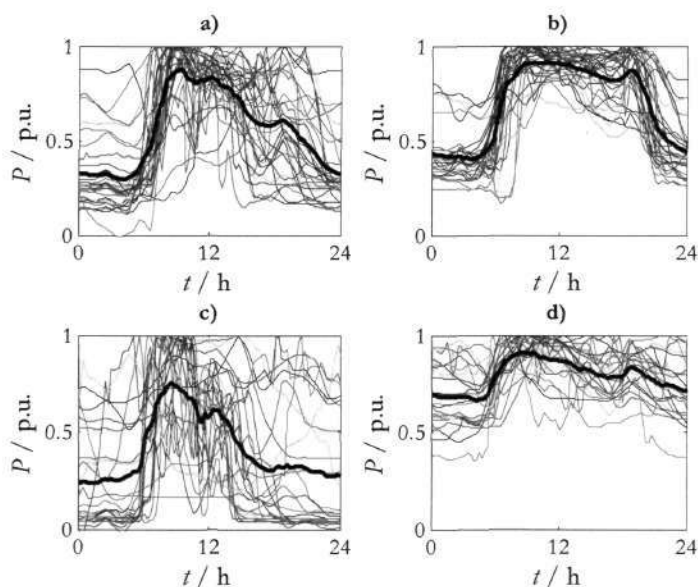
Izkušnje strokovnjakov na podlagi dolgoletne prakse na področju obratovanja distribucijskega omrežja in dobave električne energije odjemalcem kažejo, da so odjemalčeve obratovalne ure primernejše za njihovo razvrščanje v različne razrede kakor samo porabljena električna energija. Npr. odjemalec, ki ima približno 3500 obratovalnih ur letno, ima dvoizmenski obratovalni cikel. Razlog za takšne izkušnje je relativno preprost. Obratovalne ure vsebujejo dva podatka o odjemalčevi porabi, tj. konično moč  $P_k$  in mesečno energijo  $W_{MES}$  in podajajo, kolikšno število ur v mesecu bi odjemalec obratoval, če bi konstantno obratoval s konično močjo. Definiramo jih kot sledi:

$$H_M = \frac{W_{MES}}{P_k} \quad (3.44)$$

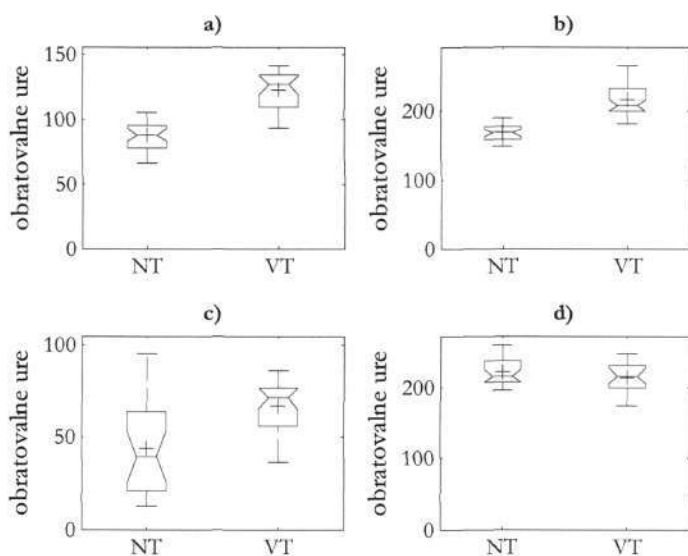
in jih nadalje razdelimo na mesečne obratovalne ure visoke  $H_{MVT}$  in nizke  $H_{MNT}$  tarife.

Rezultate razvrščanja odjemalcev na podlagi obratovalnih ur visoke in nizke tarife prikazujeta Sl. 3.25 in Sl. 3.26. Iz slik se vidi, da razvrščanje obratovalnih ur določa skupine UO-jev s homogenejšimi OD-ji, grafa b) in d) na Sl. 3.25, od razvrščanja mesečnih energij, vendar imamo še vedno skupine z raznolikimi OD-ji, grafa c) in a). Trivialno rešitev problema predstavlja izločitev omenjenega grafa iz rezultatov določanja in dodeljevanja OD-jev, npr. graf c) predstavlja zelo nehomogeno skupino odjemalcev, katere povprečni OD izkazuje veliko standardni odklon. Z izločitvijo izgubimo podatek, kakšen OD imajo odjemalci z obratovalnimi urami, ki jih opisuje izločena skupina, npr. ne vemo oblike OD-ja odjemalcev,

ki imajo v povprečju 45 obratovalnih ur v nizki tarifi in 67 obratovalnih ur v visoki tarifi mesečno, graf c) Sl. 3.26. Ker pri tem omejimo splošnost metode, nekaterim skupinam odjemalcev ne moremo dodeliti NOD-ov, takšna rešitev ni primerna. Menimo, da uporaba mesečnih obratovalnih ur ne zadostuje zahtevi po preprosti in univerzalni metodi.



Sl. 3.25: Normirani OD-ji; razvrščanje odjemalcev na podlagi mesečnih obratovalnih ur.



Sl. 3.26: Obseg, mediana in povprečna vrednost (označena s +) mesečnih obratovalnih ur v visoki VT in nizki NT tarifi; razvrščanje odjemalcev na podlagi mesečnih obratovalnih ur.

Kratka analiza uporabe porabljene energije in obratovalnih ur kaže omejeno možnost njune uporabe pri oblikovanju splošne in učinkovite metode dodeljevanja NOD-ov. Z njima ne moremo oblikovati tipičnih skupin odjemalcev, ki imajo različne povprečne OD-je z majhno standardnim odklonom, z njimi ne moremo tudi zajeti celotne populacije odjemalcev. Kot

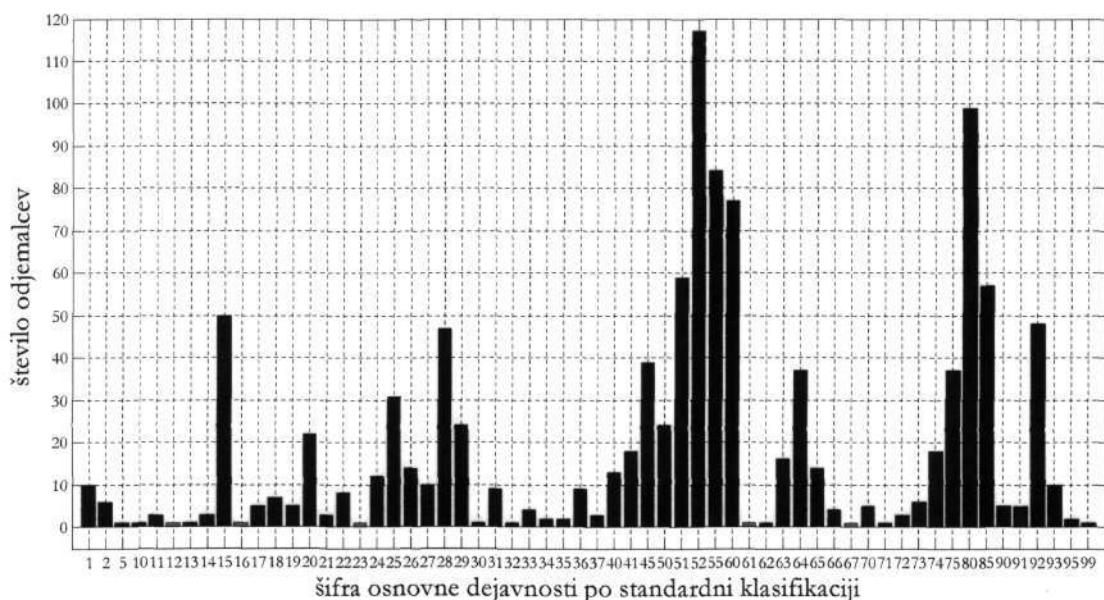
primerna rešitev se je pokazala uporaba tipa odjemalčeve dejavnosti, katere postopek prikazuje naslednji razdelek.

### 3.3.2. Uporaba tipa odjemalčeve dejavnosti za dodeljevanje NOD-ov

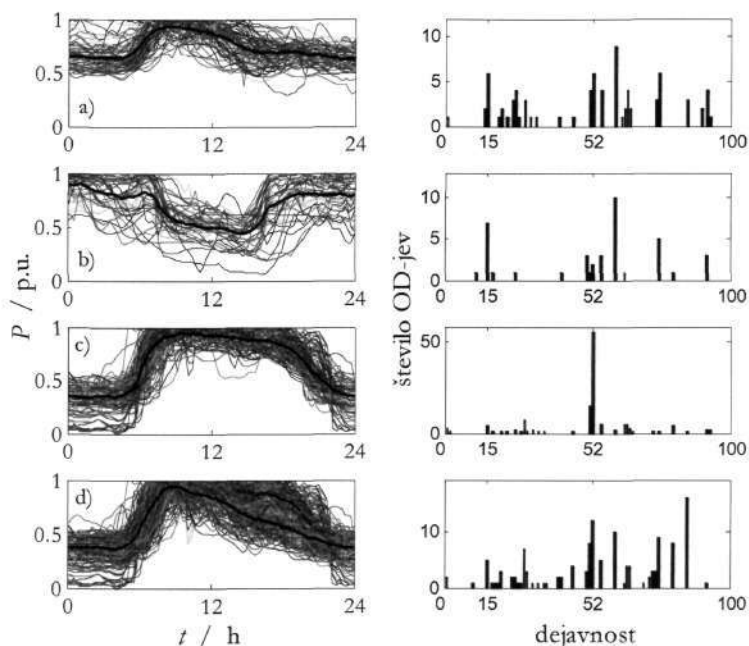
Kot smo že omenili, predstavlja odjemalčeva registrirana dejavnost primeren parameter, na podlagi katerega smo določili, kateri NOD pripada posameznemu ali skupin UO-jev. Dejavnosti odjemalcev določa Standardni klasifikat dejavnosti (SKD), ki se sklada z evropsko klasifikacijo ter zavzema 60 glavnih dejavnosti označenih od 1 do 99, vendar ne zavzema vseh vrednosti. Dejavnost posameznega UO-ja smo izbrali kot edini preostali razpoložljiv in enostavno pridobljiv podatek, ki ga imajo distribucijska podjetja. Preostala dva podatka in njuna kombinacija so se izkazali za manj primerni. Njegova uporaba daje jasen kriterij, kateri NOD pripisati posameznemu UO-ju.

V ta namen smo najprej analizirali, katerim dejavnostim pripadajo UO-ji, od katerih smo pridobili meritve OD-jev. Analizo zastopanosti števila UO-jev v glavni dejavnosti podaja Sl. 3.27. Iz slike se vidi, da so najbolj zastopane dejavnosti 15, 50, 51, 52 in 80. Pomen posamezne šifre in dejavnosti prikazuje Tab. 1 v prilogi B. Relativno mnogo je dejavnosti, ki so zelo slabo zastopane z enim ali dvema UO-jema. Nekatere dejavnosti celo niso zastopane in so označene z rdečo barvo. Te dejavnosti so 12, 16, 23, 61 in 67.

Prvi korak pri uporabi registrirane dejavnosti je namenjen iskanju korelacije med obliko povprečnega OD-ja posamezne skupine ter registrirane dejavnosti odjemalcev, ki to skupino sestavljajo. Pri tem naj še enkrat poudarimo, da smo UO-je razvrščali, rojili, glede na podobnost diagramov. Za rojenje smo uporabili FCM algoritem ter določili "trdo" pripadnost OD-ja v posamezni skupini. Nekatere rezultate iskanja korelacije med skupinami in tipom registrirane dejavnosti v grafični obliki prikazuje Sl. 3.28.



Sl. 3.27: Število odjemalce glede na njihovo dejavnost.



Sl. 3.28: Korelacija med obliko OD-ja in dejavnostjo UO-jev.

V levi strani zgornje slike vidimo posamezne skupine, z oznakami z zaporedno črko od a do d, skupaj s OD-ji, ki jih sestavljajo ter povprečnim OD-jem, ki je označen z debelejšo črto. Na desni strani slika prikazuje število OD-jev, ki sestavljajo posamezno skupino v odvisnosti od dejavnosti UO-jev. Na abscisi so šifre glavnih SKD-jev, na ordinati število OD-jev, ki sestavljajo povprečen diagram za posamezen glavni SKD. Iz slike vidimo korelacijo med obliko povprečnega OD-ja ter registrirano dejavnostjo odjemalcev. Ugotovimo, da sta npr. dejavnosti 15 in 52 zastopani v vseh štirih skupina. Zaradi pojavljanja iste dejavnosti v različnih podobnih skupinah nastopijo težave pri enoumnem dodeljevanju NOD-ov posameznim UO-jem. Podobne ugotovitve bi lahko podali tudi za nekatere druge dejavnosti. Predstavljena analiza kaže na dejstvo, da je korelacija med registrirano dejavnostjo posameznega UO-ja ter obliko njegovega OD-ja šibka. Do podobne ugotovitve so prišli tudi avtorji v referenci [24], [25].

Kljub nizki stopnji korelacije vseeno predlagamo uporabo registrirane dejavnosti za dodelitev NOD-a posameznemu odjemalcu, saj je ta podatek trenutno edina informacija o aktivnosti, ki jo imamo za UO-je in edina pot do enoumnega dodeljevanja NOD-a posameznemu UO-ju. Pokazali smo že, da sta preostala podatka praktično neuporabna. Dodeljevanje NOD-ov odjemalcem predstavlja zadnji korak v celotnem postopku oblikovanja NOD-ov. Nalogo smo rešili v dveh korakih. Najprej smo s postopkom opisanem v razdelkih 3.1 in 3.2 merjene podatke razvrstili v skupine s podobnimi OD-ji. Povprečni diagrami posamezne skupine predstavljajo "učno" množico vzorcev pri razvrščanju s VNO-jev iz razdelku 3.2.3.2, medtem ko število skupin določa število nevronov v drugem nivoju. Vzporedno s tem smo določili povprečne diagrame posameznih glavnih dejavnosti. Pri tem smo uporabili meritve OD-jev

odjemalcev, ki smo jih "obdelali" v skladu s postopkom, opisanim v razdelku 3.1. Povprečni diagrami glavnih dejavnosti predstavljajo vhodne vzorce v VNO.

Drugi korak predstavlja razvrstitev povprečnih diagramov glavnih dejavnosti v skupine, ki so rezultat rojenja podobnih diagramov. Tiste povprečne OD-je skupin, v katere razvrstimo povprečne diagrame glavnih dejavnosti, imenujemo **Nadomestne Obremenitvene Diagrame** (NOD-e).

### 3.4. Ugotovitve

V poglavju smo predstavili novo metodo določanja NOD-ov, ki uporablja algoritme za razpoznavanje vzorcev. Bistvena njena prednost je, da omogoča enostaven in pregleden postopek dodeljevanja NOD-ov odjemalcem samo na podlagi njihove dejavnosti.

Izdelava NOD-ov temelji na predpripravi podatkov, kjer smo z valčki družine symlet merjene OD-je najprej gladili in jih nato glede na njihovo konično moč normirali. Poglavje predstavlja tri algoritme za razvrščanje vzorcev, ki smo jih testirali na simulacijskem primeru treh strnjenih skupin vzorcev in dveh posebnježev. Pri razvrščanju s FCM algoritmom smo na testnih podatkih testirali dve meri veljavnosti rojenja, ki dajeta enake rezultate. Število skupin, v katere se vzorci razvrščajo pri hierarhičnih metodah, smo določili na podlagi razdalje med zaporednima skupinama. Pri testiranju se je algoritem, ki izračunava povprečno razdaljo med skupinami pokazal, kot primernejši v primerjavi z algoritmom, ki izračunava Wardovo razdaljo, saj jasno pokaže na posebneža. Algoritem z Wardovo razdaljo predlaga združitev posebnježev z najbližjo skupino.

Uporaba različnih algoritmov izračuna razdalje med skupinami je odvisna od namena razvrščanja. Če iščemo posebneže v množici podatkov, je primerna povprečna razdalja, medtem ko daje izračun Wardove razdalja boljše rezultate, če želimo določiti splošne skupine, ki se morajo medsebojno jasno razlikujejo. Zato menimo, da je pri izdelavi NOD-ov izračun Wardove razdalje primernejši od izračuna povprečne razdalje, kar bomo pokazali v razdelku 5.1.

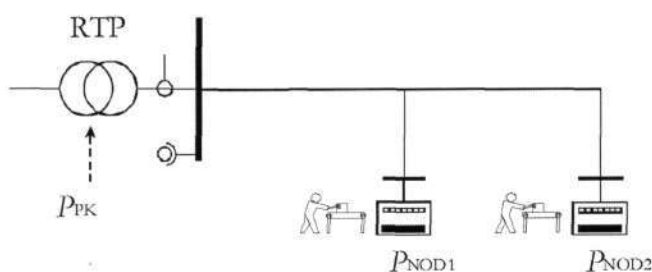
Analiza dodeljevanja NOD-ov posameznim UO-jem na podlagi informacij, ki jih imajo distribucijska podjetja o odjemalcih, je pokazala, da je registrirana dejavnost zadosten podatek za dodeljevanje NOD-ov UO-jem. Porabljena mesečna energija v visoki in nizki tarifi, konična moč ter mesečne obratovalne ure so se pokazali kot neprimerni podatki pri oblikovanju NOD-ov.

## 4. Nova metoda ugotavljanja porabe električne energije

### 4.1. Izpeljava metode in kriteriji, ki jih mora upoštevati

Predlagana metoda temelji na znani preostali krivulji  $P_{PK}$ , ki jo določa (2.3) in NOD-ih, ki smo jih dodelili odjemalcem v skladu s postopkom iz razdelka 3.3. Pri izpeljavi matematičnega modela bomo obravnavali samo en dan, kjer bomo predpostavili, da poznamo dnevno porabo električne energije posamezne skupine odjemalcev.

Ker je uporaba podatkov iz realnega distribucijskega omrežja zaradi vpliva gospodinjskih odjemalcev in preostalega odjema, za katere še nimamo NOD-ov, neprimerna, bomo izpeljavo metode pokazali na preprostem in poenostavljenem primeru distribucijskega omrežja. Omrežje sestavlja RTP, kjer imamo sprotne meritve ter dve skupini odjemalcev brez sprotnih meritev, ki se na distribucijsko omrežje priključujejo na različnih nivojih. Omrežje prikazuje Sl. 4.1. Preostalo krivuljo v RTP-ju smo simulirali z vsoto dejanskih meritev porabe upravičenih odjemalcev, ki smo jih pridobili od slovenskih distribucijskih podjetij. Poudariti moramo, da meritev, na podlagi katerih smo primer simulirali, v realnosti ne poznamo.



Sl. 4.1: Primer preprostega distribucijskega omrežja.

Kljub izpeljavi metode na preprostem primeru in upoštevanju nekaterih poenostavitev splošnost njene uporabe ni omejena, saj jo relativno enostavno razširimo na poljubno število skupin odjemalcev. Kriteriji in omejitve, katerim mora predlagan model optimizacije zadostovati, so splošni in veljajo tudi, če obdobje opazovanja razširimo na en teden ali mesec.

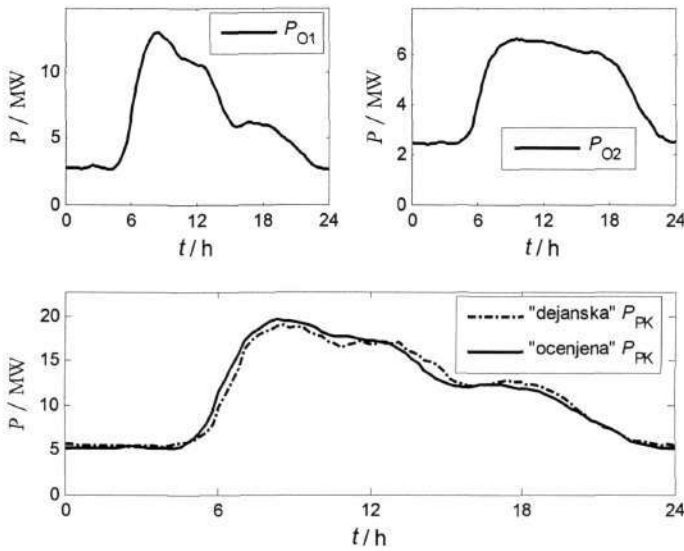
Predlagana metoda, ki smo jo poimenovali optimalno razdeljevanje preostalega odjema z uporabo NOD-ov, sestavljajo trije večji koraki:

- ocena preostale krivulje in OD-jev skupin odjemalcev,
- prilagoditev ocenjenih OD-jev skupin odjemalcev z uporabo primerne optimizacijskega postopka, da odpravimo odstopanja med izračunano in ocenjeno preostalo krivuljo,



- določitev OD-jev posameznih odjemalcev ali dobaviteljev.

V prvem koraku ocenimo OD-je skupin odjemalcev  $P_{O_i}$  in preostalo krivuljo  $P_{OPK}$  z uporabo sintetičnega postopka, ki smo ga podrobneje predstavili v razdelku 2.3.2. Istočasno bi v realnem distribucijskem sistemu določili tudi dejansko preostalo krivuljo. Z besedo "dejansko" označujem preostalo krivuljo, ki jo izračunamo z uporabo enačbe (2.3) in predstavlja splošno dinamiko spreminjanja porabe vseh odjemalcev brez sprotnih meritev. Ker nimamo realnega sistema, smo dejansko preostalo krivuljo simulirali z meritvami OD-jev analiziranih odjemalcev. Rezultate prvega koraka za primer preprostega omrežja iz Sl. 4.1 grafično prikazuje Sl. 4.2. V analizo smo vključili odjemalce, katerim smo dodelili prvi in peti NOD iz Sl. 5.18.



Sl. 4.2: Primer ocenjene in dejanske preostale krivulje  $P_{PK}$ .

Iz zgornje slike se vidijo odstopanja ocenjene preostale krivulje od dejanske, ki so posledica NOD-ov in ki predstavljajo boljši ali slabši približek dejanskim OD-jem.

Prvi korak predlaganega postopka, katerega rezultate za dve skupini odjemalcev prikazuje Sl. 4.2, v splošnem matematično modeliramo ali opišemo z naslednjo enačbo:

$$\mathbf{z} = \mathbf{h}(\mathbf{X}) + \mathbf{e}, \quad (4.1)$$

kjer je:

- $\mathbf{z}$  –  $m$  dimenzionalni vektor, ki predstavlja dejansko preostalo krivuljo  $P_{PK}$ ,
- $\mathbf{X}$  –  $m \times n$  dimenzionalno matriko rekonstruiranih OD-jev skupin odjemalcev  $P_{NOD_i}$ ,
- $\mathbf{h}(\cdot)$  – ustrezna linearna funkcija; v našem primeru je to vsota,
- $\mathbf{e}$  –  $m$  dimenzionalni vektor odstopanj med dejansko in ocenjeno preostalo krivuljo,
- $n$  – število skupin odjemalcev brez sprotnih meritev, ki določajo preostalo krivuljo.

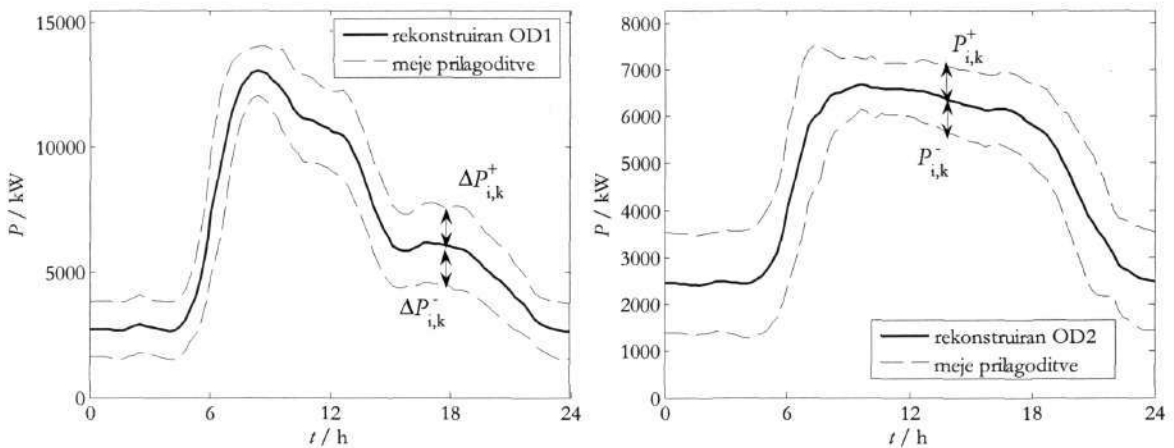
Funkcija  $h(\mathbf{X})$  predstavlja splošen matematični zapis ocene preostale krivulje z uporabo sintetičnega postopka, ki smo jo definirali že z (2.4) in (2.5) v razdelku 2.3.2. S (4.1) prikazan način ali oblika zapisa prvega koraka predlaganega postopka predstavlja znano obliko enačbe, ki se uporablja na področju ocenjevanja stanja v EES-u. Takšen zapis problema jasno nakazuje postopke in možne rešitve drugega koraka predlaganega postopka optimalne razdelitve preostalega odjema. Cilj tega koraka je z uporabo primerne matematičnega modela in optimizacijskaga algoritma prilagoditi rekonstruirane OD-je tako, da minimiziramo ali odpravimo odstopanja  $e$  iz (4.1), kar zapišemo z naslednjo enačbo:

$$e = z - h(\mathbf{X}) = \min \quad (4.2)$$

Pri tem težimo, da so popravki OD-jev čim manjši, saj v nasprotnem primeru izgubimo informacijo o lastnostih posameznih skupin odjemalcev določenih z NOD-i. Rezultat optimizacije so nekoliko prilagojeni OD-ji posamezne skupine odjemalcev, ki jih bomo za potrebe disertacije imenovali "dejanski" OD-ji.

Na področju ocenjevanja stanja so raziskovalci predlagali različne optimizacijske metode, namenjene minimizaciji odstopanja  $e$ , med katerimi so uporaba metode uteženih najmanjših absolutnih vrednosti (ang. Weighted Least Absolute Value) [44], [45] in linearnega ali mehkega linearnega programiranja [46], [47]. Linearna narava funkcije  $h(\cdot)$  v (4.1) kaže, da je za naš primer najprimernejša uporaba linearnega [48], [49] programiranja. Dodatni razlogi za uporabo optimizacijskih metod linearnega programiranja so še enostavnost zapisa kriterijske funkcije in omejitev ter hitra konvergenca h globalnemu minimumu ali maksimumu.

Predlagano metodo optimalnega razdeljevanja preostale krivulje smo izpeljali iz (4.1) z uporabo nekaterih predlogov v [47]. Glavno idejo oblikovanja optimizacijskega modela prikazuje Sl. 4.3.



Sl. 4.3: Princip prilagoditve rekonstruiranih NOD-ov.

V matematični model uvedemo dodatni spremenljivki  $P_{i,k}^+$  in  $P_{i,k}^-$ , ki predstavljata pozitiven in negativen popravek rekonstruiranega OD-ja skupine odjemalcev  $k$  v časovnem intervalu  $i$ .

Tako dobimo  $2 \cdot m \cdot n$ , kjer je  $n$  število skupin odjemalcev in  $m$  število 15-minutnih intervalov, dodatnih spremenljivk, ki se morajo nahajati znotraj v naprej določene meje prilagoditev, kot kaže Sl. 4.3. Na podlagi uvedbe dodatnih spremenljivk oblikujemo matematični model, za katerega velja:

- popravljena obremenitev ne sme prekoračiti postavljene meje negotovosti rekonstruiranih OD-jev skupin odjemalcev;
- porabljen energija odjemalcev se s prilagajanjem OD-jev ne sme spremeniti; ploščina pod dejanskim OD-jem je enaka ploščini pod rekonstruiranim OD-jem;
- pozitiven in negativen popravek sta pozitivnega predznaka,  $P_{i,k}^+ \geq 0$  in  $P_{i,k}^- \geq 0$ ;
- izračunana in ocenjena preostala krivulja sta po optimizaciji enaki;  $\mathbf{z} = \mathbf{h}(\mathbf{X}), \Rightarrow \mathbf{e} = 0$ ;
- rekonstruirane OD-je želimo čim manj spremeniti.

Zapisane lastnosti predstavljajo osnovne pogoje, ki jim mora predlagana metoda zadostiti. Na podlagi prvih štirih pogojev oblikujemo omejitve matematičnega modela, medtem ko na podlagi zadnjega pogoja določimo kriterijsko funkcijo.

Zadnji korak predlaganega postopka predstavlja določitev OD-jev posameznih odjemalcev ali njihovih dobaviteljem. Odjemalcev OD določimo z razdelitvijo dejanskega OD-ja skupine odjemalcev, kateri pripada v odvisnosti od njegovega deleža energije, medtem ko dobaviteljev OD določimo s seštevkom OD-jev odjemalcev, katerim dobavlja energijo.

Do sedaj smo zelo na kratko opredelili posamezna korake predlagane metode določanja porabe odjemalcev brez sprotnih meritev. Definirali smo pogoje, ki jih moramo vključiti v matematični model, na podlagi katerega bomo izvedli optimizacijo. Zaradi linearnosti funkcije  $h(\cdot)$  v (4.1) predlagamo uporabo linearnega programiranja. V nadaljevanju sledi podrobnejši opis drugega koraka, saj sta prvi in tretji korak preprosta in ne potrebujeta dodatnih pojasnil.

## 4.2. Optimizacija rekonstruiranih OD-jev

### 4.2.1. Splošno o optimizaciji in linearnem programiranju

Splošno obliko optimizacijskega problema, kjer iščemo minimalno vrednost kriterijske funkcije, zapišemo [48]:

$$\text{najdi } \mathbf{x} = [x_1, \dots, x_n]^T, \text{ ki minimizira } f(\mathbf{x}), \quad (4.3)$$

kjer upoštevamo enačbe in neenačbe omejitvev:

$$h_j(\mathbf{x}) = 0, \quad j = 1, \dots, r, \quad (4.4)$$

$$g_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m, \quad (4.5)$$

kjer je:

$\mathbf{x}$  –  $n$  dimenzionalni vektor optimizacijskih ali odločitvenih spremenljivk,

$f(\mathbf{x})$  – kriterijska ali ciljna funkcija,

$g(\mathbf{x})$  – sistem neenačb omejitev,

$h(\mathbf{x})$  – sistem enačb omejitev,

$m, r$  – število neenačb in enačb omejitev.

Vektor odločitvenih spremenljivk  $\mathbf{x}$  je dopustna rešitev, če zadosti predpisanim omejitvam  $g(\mathbf{x})$  in  $h(\mathbf{x})$ . Če pri tem kriterijska funkcija doseže svoj minimum, pravimo, da je  $\mathbf{x}$  optimalna rešitev. Optimizacijski problem, zapisan z enačbami (4.3) do (4.5), imenujemo omejen optimizacijski problem. Če omejitev nimamo,  $m = 0, r = 0$ , imamo optimizacijski problem brez omejitev. Glede na vrsto ali lastnosti funkcij  $f(\cdot), g(\cdot)$  in  $h(\cdot)$  in njihovih koeficientov ločimo različne optimizacijske probleme, za reševanje katerih obstaja cela vrsta optimizacijskih metod, kakor so npr. linearno, nelinearno, kvadratično, geometrično ali stohastično programiranje. Izbiro optimizacijske metode pogojuje vrsta optimizacijskega problema, njegova velikost in želena natančnost rešitve.

V disertaciji smo zaradi linearnosti kriterijske funkcije, enačb in neenačb omejitev uporabili linearno programiranje. Splošni problem linearnega programiranja izrazimo v matematični obliki takole [49]: določiti je treba vrednosti optimizacijskih spremenljivk  $\mathbf{x}$ , ki zadoščajo pogojem nenegativnosti  $\mathbf{x} \geq 0$  in linearnim neenačbam ali enačbam omejitev tako, da ima kriterijska funkcija  $f(\mathbf{x})$  ekstrem, to je minimum ali maksimum. V matrični obliki ga za namen iskanja minimuma kriterijske funkcije zapišemo kot:

$$\text{kriterijska funkcija:} \quad \min_{\mathbf{x}} f(\mathbf{x}) = \min_{\mathbf{x}} \mathbf{c}^T \mathbf{x}, \quad (4.6)$$

$$\begin{aligned} \text{omejitve:} \quad & \mathbf{A}_{\text{eq}} \mathbf{x} = \mathbf{b}_{\text{eq}}, \\ & \mathbf{Ax} \leq \mathbf{b}, \\ & \mathbf{x} \geq 0. \end{aligned} \quad (4.7)$$

Za reševanje optimizacijskega problema je ugodno, da vsako neenačbo iz (4.7) prevedemo v enačbo, kar rešimo z vpeljavo dopolnilnih spremenljivk. Tega koraka uporabnikom optimizacijskih metod ni treba posebej izvajati, ker so sodobna računalniška jedra, namenjena linearnemu programiranju, sposobna sama pretvoriti neenačbe v enačbe. Uporabnik mora zapisati ustrezno kriterijsko funkcijo, enačbe in neenačbe ter izbrati primeren optimizacijski algoritem ali metodo. V disertaciji smo uporabili računalniško jedro iz v programskem paketu Matlab [50], ki omogoča uporabo dveh metode iskanja optimalne rešitve, tj. metode simpleksov in metode notranje točke. Prva metoda išče rešitve optimizacijskega problema s premikanjem po robu dopustnega območja, poliedra, definirane z omejitvami iz enega oglišča v drugo. Pri tem v vsakem koraku išče minimalno vrednost kriterijske funkcije  $f(\mathbf{x})$ . Metoda notranje točke začne iskati rešitev znotraj prostora dopustnih rešitev in se postopoma premika navzven proti optimalni rešitvi. Algoritem je primeren za iskanje optimalne rešitve, če imamo veliko število optimizacijskih spremenljivk in omejitev in je veliko hitrejši.

### 4.2.2. Določitev meje popraviljanja rekonstruiranih OD-jev

Rekonstruirane OD-je  $P_{OK,i}$  posameznih skupin odjemalcev smo določili z "vrivanjem" energije pod NOD-e, ki predstavljajo srednjo vrednost ali matematično upanje vseh normiranih OD-jev, razvrščenih v isto skupino. Zaradi tega imajo NOD-i določeno naključnost ali raztros, ki ga določa standardni odklon. Ta podatek smo uporabili kot podlago za določanje meje, znotraj katere smo še dovoljevali prilagajanje rekonstruiranih OD-jev posameznih skupin odjemalcev. Pri tem smo uporabili Gaussov zakon širjenja pogreškov, ki ga za splošen primer zapišemo:

$$\sigma_f^2 = \sum_{j=1}^k a_j^2 \cdot \sigma_j^2, \quad (4.8)$$

kjer je  $\sigma_f^2$  disperzija odvisne spremenljivke  $z$ , ki je funkcija neodvisnih spremenljivk  $x_j$ ;  $z$  disperzijo  $\sigma_j^2$ ,  $j = 1, \dots, k$ . Koeficienti  $a_j$  predstavljajo parcialne odvode spremenljivke  $z$  po  $x_j$ .

V disertaciji smo za neodvisne spremenljivke določili NOD-e, kjer poznamo standardni odklon  $\sigma_{NOD}$ . Odvisne spremenljivke predstavljajo rekonstruirani OD-ji. Njihov standardni odklon  $\sigma_{OD}$  smo za  $i$ -ti interval  $k$ -te skupine odjemalcev določili z:

$$\sigma_{i,k} = \frac{W_{ODk}}{W_{NODk}} \cdot \sigma_{NODi,k} = c_k(W) \cdot \sigma_{NODi,k}, \quad (4.9)$$

kjer je:

$W_{ODk}$  – energija skupine odjemalcev, ki jim je dodeljen  $k$ -ti NOD,

$W_{NODk}$  – površina pod  $k$ -tim NOD-om, ki predstavlja njegovo energijo.

V disertaciji smo predpostavili, da so normirani OD-ji normalno porazdeljeni v posamezni skupini. Za določanje zgornje in spodnje meje prilagajanja rekonstruiranih OD-jev smo uporabili verjetnost zaupanja, da se dejanski OD skupine odjemalcev nahaja znotraj intervala  $P_{OK,i} \pm l \cdot \sigma_{k,i}$ . Vrednosti parametra  $l$  za nekatere intervale zaupanja podaja Tab. 4.1.

Tab. 4.1: Vrednost parametra  $l$  za nekatere intervale zaupanja.

Interval zaupanja	50 %	68,3 %	86,6 %	95,4 %	99,7 %
1	0,675	1	1,5	2	3

### 4.3. Zapis matematičnega modela linearnega programiranja

Do sedaj smo si ogledali splošen zapis matematičnega modela linearnega programiranja in postopek določanja zgornje in spodnje meje negotovosti rekonstruiranih OD-jev. V nadaljevanju sledi zapis glavnih pogojev iz začetnega dela razdelka 4, ki jim mora predlagana metoda določanja OD-jev zadostiti, v matematični obliki.

Na podlagi zahteve, da rekonstruirane OD-je skupin odjemalcev čim manj spreminjamo, ker želimo ohraniti informacije o splošnih značilnostih porabe posameznih skupin odjemalcev, ki jim pripišemo NOD, zapišemo kriterijsko funkcijo, kjer iščemo njeno minimalno vrednost. Poiskati moramo minimalno vrednost vsote vseh pozitivnih  $P^+$  in negativnih  $P^-$  popravkov rekonstruiranih OD-jev, kar zapišemo kot:

$$\min_{P^+, P^-} f(P_{i,k}^+, P_{i,k}^-) = \sum_{i=1}^m \sum_{k=1}^p (P_{i,k}^+ + P_{i,k}^-), \quad (4.10)$$

kjer je  $m$  število 15 minutnih intervalov,  $p \leq n$  število analiziranih skupin odjemalcev in  $n$  število NOD-ov.

Omejitve (4.7) iz splošnega zapisa linearnega programiranja določimo na podlagi preostalih glavnih pogojev, ki jim mora predlagana metoda zadostiti, kjer zahtevamo:

Porabljen energija odjemalcev se s prilagajanjem OD-jev ne sme spreminjati. Energija odjemalcev, ki jim je dodeljen isti NOD  $W_{\text{NOD}k}$ , mora biti po optimizaciji enaka energiji odjemalcev pred optimizacijo:

$$W_{\text{NOD}k \text{ po optimizaciji}} = W_{\text{NOD}k \text{ pred optimizacijo}}, \quad \forall k. \quad (4.11)$$

Iz (4.11) izpeljemo prvi pogoj za  $P^+$  in  $P^-$ , ki pravi, da mora biti vsota njune razlike po vseh 15 minutnih intervalih  $m$  enaka nič za vse skupine odjemalcev  $k$ , kar zapišemo z:

$$\sum_{i=1}^m (P_{i,k}^+ - P_{i,k}^-) = 0, \quad \forall k. \quad (4.12)$$

Po optimizaciji morata v vsakem časovnem intervalu  $i$  biti dejanska  $P_{\text{DPK}i}$  in ocenjena  $P_{\text{OPK}}$  preostala krivulja enaki,  $\mathbf{z} = h(\mathbf{X}) \Rightarrow \mathbf{e} = 0$ . Na podlagi tega pogoja zapišemo naslednjo enačbo za pozitivne in negativne popravke moči:

$$\sum_{k=1}^p (P_{i,k}^+ - P_{i,k}^-) = P_{\text{DPK}i} - \sum_{k=1}^p P_{\text{O}i,k}, \quad \forall i, \quad (4.13)$$

kjer  $\sum_{k=1}^p P_{\text{O}i,k}$  predstavlja ocenjeno preostalo krivuljo  $P_{\text{OPK}}$  za  $i$ -ti časovni interval.

Vrednosti pozitivnih in negativnih popravkov moči morajo biti pozitivne in ne smejo prekoračiti postavljenih meja negotovosti rekonstruiranih OD-jev, kar zapišemo:

$$P_{i,k}^+, P_{i,k}^- \geq 0, \quad \forall i, k, \quad (4.14)$$

$$\begin{aligned} P_{i,k}^+ &< P_{\text{O}i,k}^{\max} - P_{\text{O}i,k}, \\ P_{i,k}^- &< P_{\text{O}i,k} - P_{\text{O}i,k}^{\min}, \end{aligned} \quad \forall i, k, \quad (4.15)$$



kjer sta  $P_{O_i,k}^{\max}$  in  $P_{O_i,k}^{\min}$  zgornja in spodnja meja negotovosti rekonstruiranih OD-jev iz razdelka 4.2.2. Iz (4.15) določimo zgornje meje za popravke moči:

$$P_{i,k}^+, P_{i,k}^- < 1 \cdot \sigma_{i,k} \quad \forall i,k. \quad (4.16)$$

Vrednosti faktorja  $l$  za nekatere intervale zaupanja podaja Tab. 4.1, medtem ko standardno deviacijo izračunamo po (4.9). Ker želimo rekonstruirane OD-je čim manj spremeniti, smo v disertaciji izbrali minimalno vrednost parametra  $l = 1$ . Pogoj za njegovo izbiro je bil konvergenca optimizacije.

Z enačbo (4.10), ki predstavlja kriterijsko funkcijo, in enačbami (4.12), (4.13), (4.14) in (4.16), ki predstavljajo omejitve, smo glavne pogoje, ki jim mora predlagana metoda zadostiti, zapisali v matematični obliki. Naslednji korak predstavlja njihov zapis v matrični obliki, primerni za uporabo v ustreznem računalniškem jedru namenjena linearnemu programiranju. Ker je matrični zapis z malo spretnosti relativno enostaven in hkrati odvisen od uporabljenega računalniškega jedra, ga v disertaciji ne bomo posebej obravnavali.

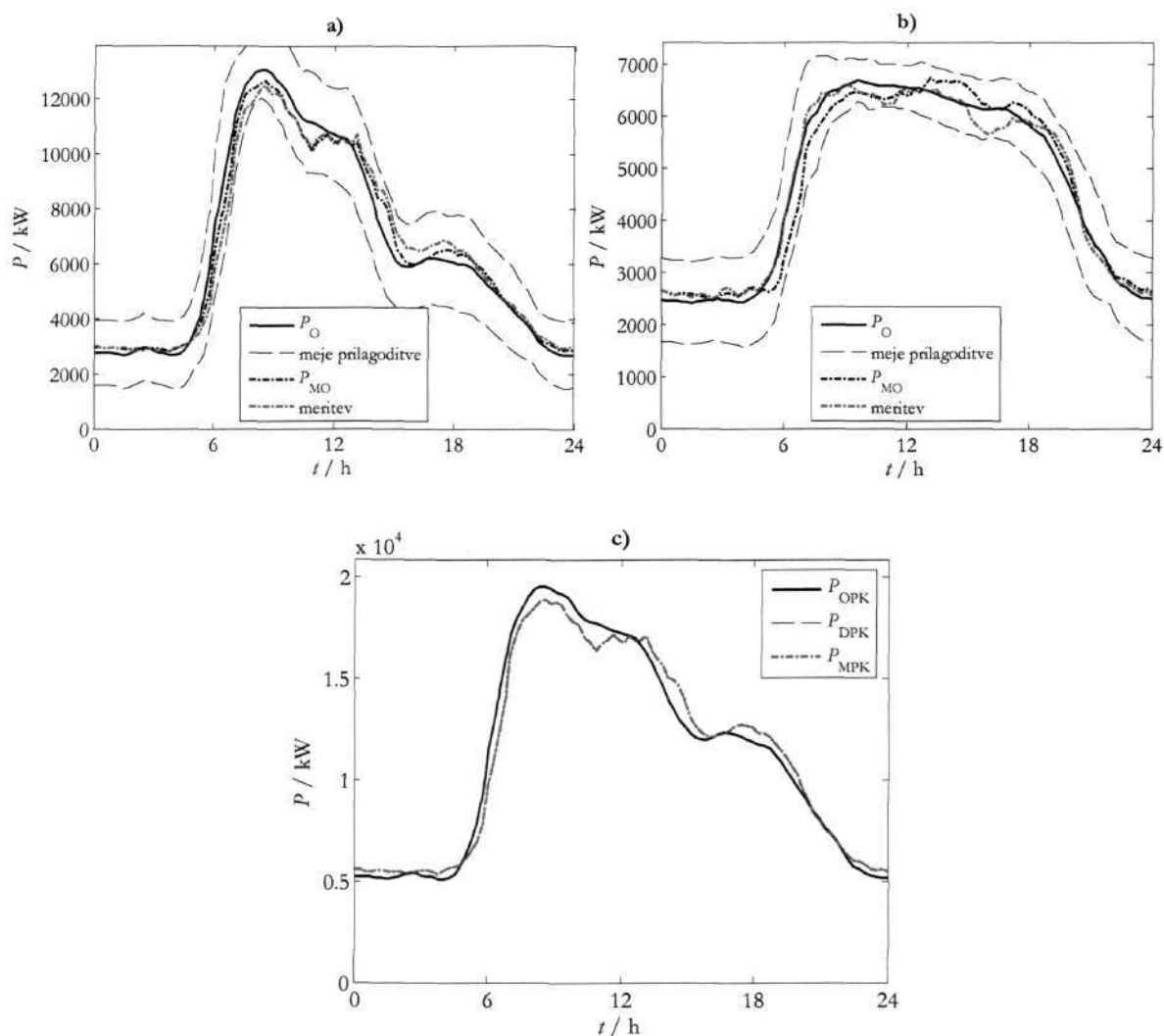
Rezultat optimizacije so optimalni pozitivni  $P_{i,k}^+$  in negativni  $P_{i,k}^-$  popravki rekonstruiranih OD-jev. V vsakem časovnem intervalu  $i$  za skupino odjemalcev  $k$  ima eden izmed njiju vrednost nič. Modificirane ali optimirane OD-je  $P_{MO}$  skupin odjemalcev izračunamo s prištevanjem in odštevanjem optimalnih popravkov moči od rekonstruiranih OD-jev  $P_O$ , kar zapišemo:

$$P_{MO_i,k} = P_{O_i,k} + P_{i,k}^+ - P_{O_i,k}^-, \quad \forall i,k. \quad (4.17)$$

Te spremenjene OD-je skupin odjemalcev proglašimo za "dejanske" OD-je odjemalcev brez sprotnih meritev. Na njihovi podlagi bi DO določal OD-je posameznim odjemalcem ali dobaviteljem, ki bi nadalje služili npr. za ugotavljanje odstopanj napovedane porabe električne energije od "dejanske".

Rezultate optimizacije OD-jev dveh skupin odjemalcev s Sl. 4.2 prikazuje Sl. 4.4, kjer grafa a) in b) prikazujeta OD-je skupin odjemalcev, medtem ko graf c) prikazuje preostalo krivuljo v RTP-ju. Iz grafov a) in b) se vidi, da optimirana OD-ja posamezne skupine odjemalcev  $P_{MO}$  sledita rekonstruiranima OD-jema  $P_O$  in se nahajata znotraj postavljenih meja prilagoditve. Za primerjavo rezultatov optimizacije smo na grafih a) in b) podali tudi vrednosti meritev OD-jev. Primerjava dodatnih analiz kaže, da smo z optimizacijo vsoto absolutnih odstopanj optimiranega OD-ja od merjenega OD-ja na grafu a) zmanjšali za približno 58 %, medtem ko so se odstopanja grafa b) povečala za 38 %. Z optimizacijo smo zmanjšali skupna odstopanja vseh optimiranih OD-jev od meritev. Ta odstopanja lahko z nekaj spretnosti še dodatno zmanjšamo, če za pozitivne in negativne popravke obremenitev upeljemo uteži v kriterijsko funkcijo. Uteži morajo biti za drugo skupino odjemalcev z najmanj odstopanji večje od uteži prve skupine odjemalcev. Tako se odstopanja te skupine ne spremenijo, saj so njeni popravki obremenitev enaki nič, tj. optimiran OD je enak rekonstruiranemu OD-ju. Ker se OD druge skupine odjemalcev ne spremeni, se mora dodatno spremeniti OD prve skupine odjemalcev,

da zadostimo pogoju optimizacije, določenem z (4.13). Tako njegovo obliko "približamo" meritvam in dodatno zmanjšamo odstopanja.



Sl. 4.4: Optimirani OD-ji.

Iz grafa c) na zgornji sliki se vidi, da sta dejanska  $P_{DPK}$  in modificirana ali optimirana  $P_{MPK}$  preostala krivulja po optimizaciji enaki, kar zadostuje pogoju (4.13). Energiji preostale krivulje pred in po optimizaciji se razlikujeta za  $2,96 \cdot 10^{-11}$  kWh, kar je zanemarljiva vrednost in lahko rečemo, da optimizacija zadostuje tudi pogoju (4.11).

Z vpeljavo optimizacije smo odpravili glavno pomanjkljivost sintetičnega postopka, tj. odpravili smo neravnotežje med porabo celotnega distribucijskega omrežja, določeno s sintetičnim in analitičnim postopkom, glej razdelek 2.4. Pri tem smo upoštevali, da imajo odjemalci različne oblike OD-jev, ki jih določajo njim dodeljeni NOD-i in da z analitičnim postopkom natančneje določimo preostalo krivuljo porabe v celotnem distribucijskem omrežju.

Zapišemo lahko, da smo s predlagano metodo analitično določeno preostalo krivuljo optimalno razdelili na posamezne skupine odjemalcev ob upoštevanju njihovih NOD-ov.

Poleg tega optimizacija zmanjša vsoto absolutnih odstopanj "dejanskih" OD-jev skupin odjemalcev od njihovih meritev, kar predstavlja drugo pozitivno lastnost vpeljave optimizacije. Ta ugotovitev je splošna in velja za poljubno število skupin odjemalcev, ne glede na to, da smo do nje prišli na podlagi analize samo dveh skupin. Predlagana metoda predstavlja optimalno kombinacijo obeh znanih metod določanja porabe odjemalcev brez sprotnih meritev in uspešno odpravlja njune temeljne pomanjkljivosti.

Ob tem moramo poudariti, da podobne analize, kakršno smo naredili na podlagi OD-jev, ki jih prikazuje Sl. 4.4, v realnem distribucijskem sistemu težko naredimo. Razlog je v neobstoju meritev odjemalčevih OD-jev. Naj še enkrat opomnimo, da je predlagana metoda določanja porabe namenjena odjemalcem brez sprotnih meritev.

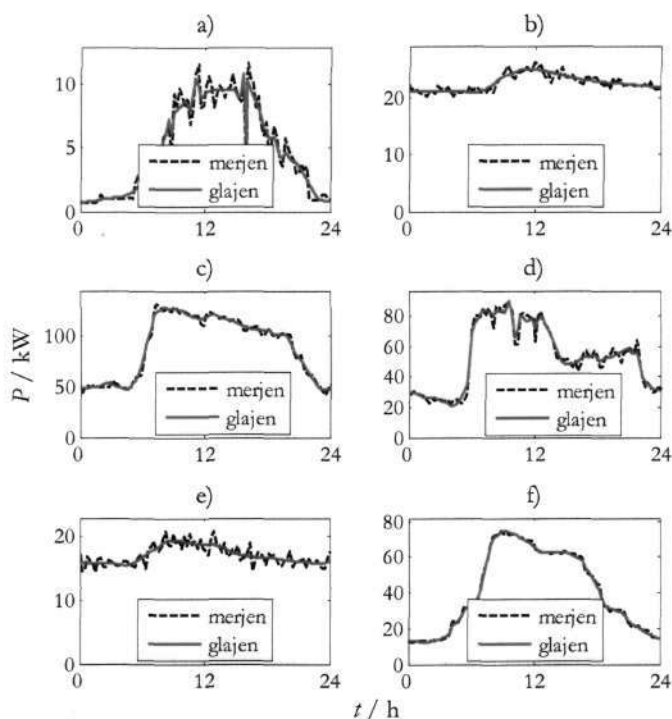
## 5. Rezultati

V predhodnih dveh razdelkih smo predstavili teoretične podlage predlaganih metod določanja porabe električne energije odjemalcev brez sprotnih meritev. Ob tem smo podali nekatere rezultate, ki so bili nujni za razumevanje obravnavanega področja. Ugotovitve na podlagi teh rezultatov so nas usmerjale h končnim rešitvam zastavljenega problema, katerih rezultate podajamo v nadaljevanju.

### 5.1. Rezultati oblikovanja NOD-ov

Novo metodo oblikovanja in dodeljevanja NOD-ov posameznim UO-jem iz poglavja 3 smo testirali na 1095 v mesecu decembru 2003 izmerjenih OD-jih odjemalcev. Meritve OD-jev smo pridobili od slovenskih distribucijskih podjetij. Iz meritev smo najprej določili povprečne OD-je za posamezni tip dneva v tednu. Rezultati v nadaljevanju, temeljijo na povprečnih OD-jih srednjih delavnikov. Sredo smo izbrali zato, ker naj bi po izkušnjah strokovnjakov iz distribucijskih podjetij predstavljala najbolj tipični delovni dan v tednu.

Pred rojenjem v podobne skupine smo OD-je gladili in normirali. Rezultate glajenja in izločanja šuma za šest OD-jev označenih s črkami od a) do f) prikazuje Sl. 5.1.



Sl. 5.1: Rezultati glajenja OD-jev.

Iz slike se vidi, da v razdelku 3.1.1 predstavljen algoritem zelo dobro gladi OD-je, če vsebujejo relativno dosti šuma, grafa a) in e). Če "merjen OD" ne vsebuje dosti šuma, je glajen OD skoraj popolnoma enak merjenemu OD-ju, grafa c) in f). Glajeni OD-ji dobro "sledijo" merjenim; nimamo zamikov med obema diagramoma pri jutranjem relativno strmim porastu in večernem upadu porabe, grafa c) in d). Stopnja glajenja je lokalno spremenljiva. Deli OD-ja, ki imajo več šuma, so bolj zglajeni, medtem ko sta merjen in glajen OD v ostalih časovnih intervalih praktično enaka. Takšen primer prikazuje graf d) na zgornji sliki. Vidimo, da je predlagan postopek glajenja dobro zgladil OD med 14 in 22 uro, medtem ko je glajenje v jutranjih urah praktično minimalno.

Naslednji korak pri določanju NOD-ov predstavlja normiranje OD-jev. Ker je ta korak enostavno izvedljiv, vse vrednosti OD-ja delimo z maksimalno vrednostjo (3.11), mu na tem mestu ne bomo posvečali dodatne pozornosti in se bomo v nadaljevanju posvetili rezultatom določanja števila skupin, v katere bomo OD-je rojili.

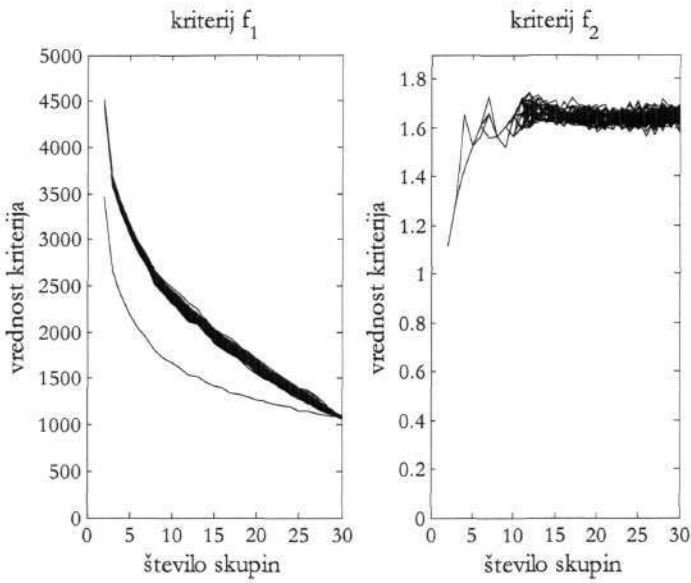
### 5.1.1. Določanje števila skupin OD-jev

V razdelku 3.2.2 smo zapisali, da je pri uporabi različnih postopkov in algoritmov za razpoznavanje vzorcev pomembno vedeti, kaj je cilj iskanja optimalne rešitve. Samo na podlagi definirane cilja lahko določimo najprimernejši postopek razpoznavanja vzorcev. V tem okviru smo si zadali cilj najti ustrezno število skupin OD-jev, ki se med seboj razlikujejo in predstavljajo najširšo množico različnih oblik OD-jev. Čeprav naš namen ni iskanje posebnih vzorcev v množici OD-jev moramo uporabiti tisti algoritem rojenja, ki je posebnije sposoben razpoznati, vendar jih mora razvrstiti samo v nekaj skupin, ki jih v nadaljevanju zanemarimo.

Določanje števila skupin smo najprej iskali s FCM algoritmom in z uporabo mer veljavnosti rojenja. Ker je FCM algoritem predstavnik nehierarhične metode, ki zahtevajo določanje želenega števila rojev pred začetkom rojenja, smo ga morali podati vnaprej. Pri splošni predstavitvi metod razpoznavanja vzorcev, razdelek 3.2, smo z roji označevali vzorce, razvrščene v posamezno skupino. Pri tem roji predstavljajo OD-je, ki sestavljajo posamezno skupino. V nadaljevanju bomo namesto rojev uporabljali izraz skupine.

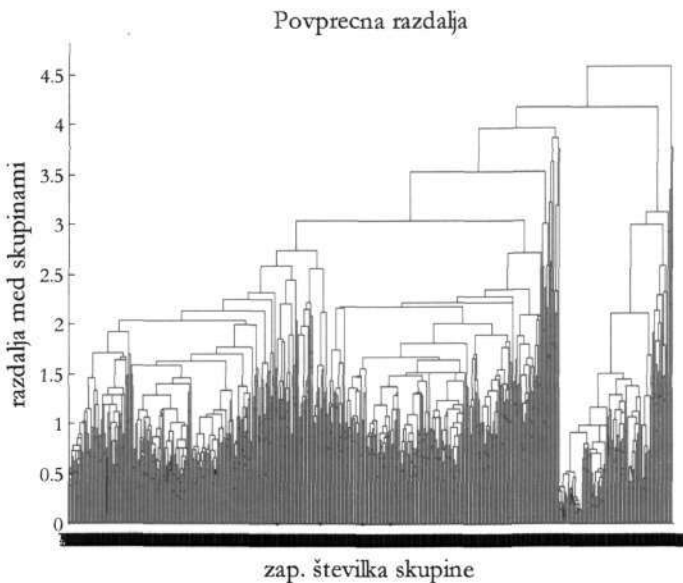
Zaradi vnaprej neznanega števila skupin smo njihovo število spreminjali od vrednosti dva do 30 pri različnih izbirah začetne pripadnostne matrike  $U$ . Število izbir je bilo 100. Za vsako izbiro pripadnostne matrike in vsako posamezno število skupin smo izračunali vrednosti kriterijskih funkcij  $f_1$  (3.24) in  $f_2$  (3.25). Njuna poteka prikazuje Sl. 5.2. Iz slike vidimo, da je potek vrednosti kriterijskih funkcij  $f_1$  in  $f_2$  v odvisnosti od števila skupin, v katere rojimo OD-je, za večino primerov različnih inicializacij pripadnostne matrike podoben. Relativno velika odstopanja imamo samo v enem od primerov pri uporabi kriterija  $f_1$ . Kriterij  $f_2$  se je za razliko od razvrščanja vzorcev, predstavljenih v razdelku 3.2.1.2, kjer daje dobre rezultate, pri razvrščanju OD-jev, ki jih opisuje 96 parametrov ali vrednosti, izkazal kot neprimeren. Literatura [36] navaja, da je rojenje, pri katerem je vrednost kriterij  $f_2$  večje od 0,5, neustrezno. Iz Sl. 5.2 vidimo, da je vrednost  $f_2$  vedno večja od ena neodvisno od izbire pripadnostne

matrike  $U$  in števila skupin. Na podlagi tega sklepamo, da omenjenega kriterija pri rojenju OD-jev ne moremo uporabiti.



Sl. 5.2: Potek kriterijev  $f_1$  in  $f_2$  v odvisnosti od števila skupin.

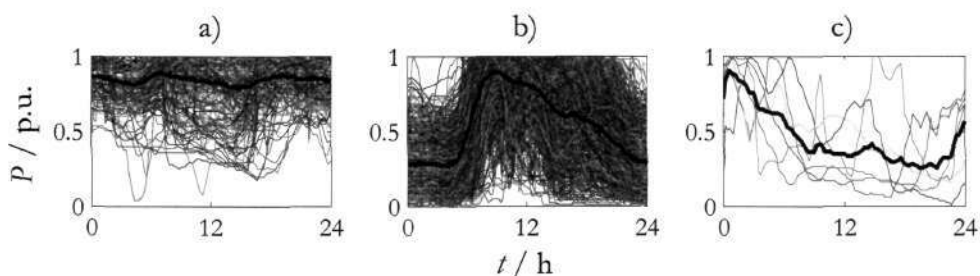
Iz zgornje slike vidimo, da koleno kriterija  $f_1$  ni tako izrazito, kakor smo ga videli na Sl. 3.13, zato je določitev števila skupin, pri katerem nastopa, subjektivna. Na razpolago imamo več možnosti izbire števila skupin, ki se po našem mnenju gibljejo med devet in petnajst. Manjše kot je število skupin, večji je raztros OD-jev v skupini. Vidimo, da je določanje kolena kriterija  $f_1$  mehko, zato smo najprimernejše število skupin preverili s hierarhičnimi metodami. Glajene in normirane OD-je smo razvrstili v skladu s postopkom iz razdelka 3.2.2. Najprej smo uporabili algoritem, ki temelji na izračunu povprečne razdalje med roji, skupinami. Rezultate v obliki drevesa združevanja ali dendrograma prikazuje Sl. 5.3.



Sl. 5.3: Dendrogram rojenja OD-jev. Povprečna razdalja med skupinami.

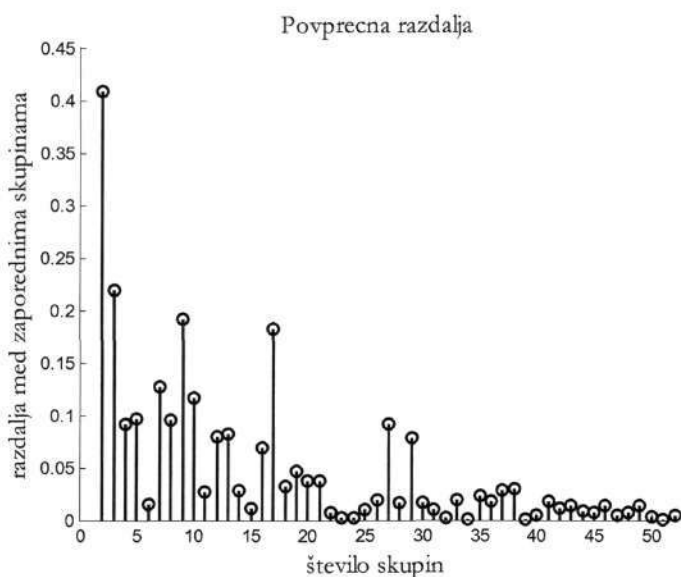


Iz Sl. 5.3 ne moremo jasno določiti primerne števila skupin, vendar se vidi, da algoritem, ki uporablja povprečno razdaljo, nove OD-je raje priključi že obstoječi skupini, kakor da oblikuje novo skupino, ki bi se razlikovala od preostalih skupin. Jasno se vidijo predvsem tri skupine, ki se združujejo na najvišjem nivoju. Razdelitev OD-jev v te tri skupine skupaj s povprečnim OD-jem posamezne skupine kaže Sl. 5.4. Večina, 890, OD-jev je šla v skupino b), ki ima posledično velik raztros, medtem ko skupino c), ki predstavlja posebnost, sestavlja samo šest OD-jev.



Sl. 5.4: Rezultati razvrstitve OD-jev v tri skupine; povprečna razdalja med skupinami.

Preprosta analiza kaže, da rojenje OD-jev v samo tri skupine ne daje dobrih rezultatov. OD-je moramo razvrstiti v več skupin, katerih število smo določali na podlagi velikosti razdalje med sosednjima skupinama, ki jo prikazuje Sl. 5.5. Na sliki smo prikazali vrednosti razdalj samo za prvih 52 skupin.

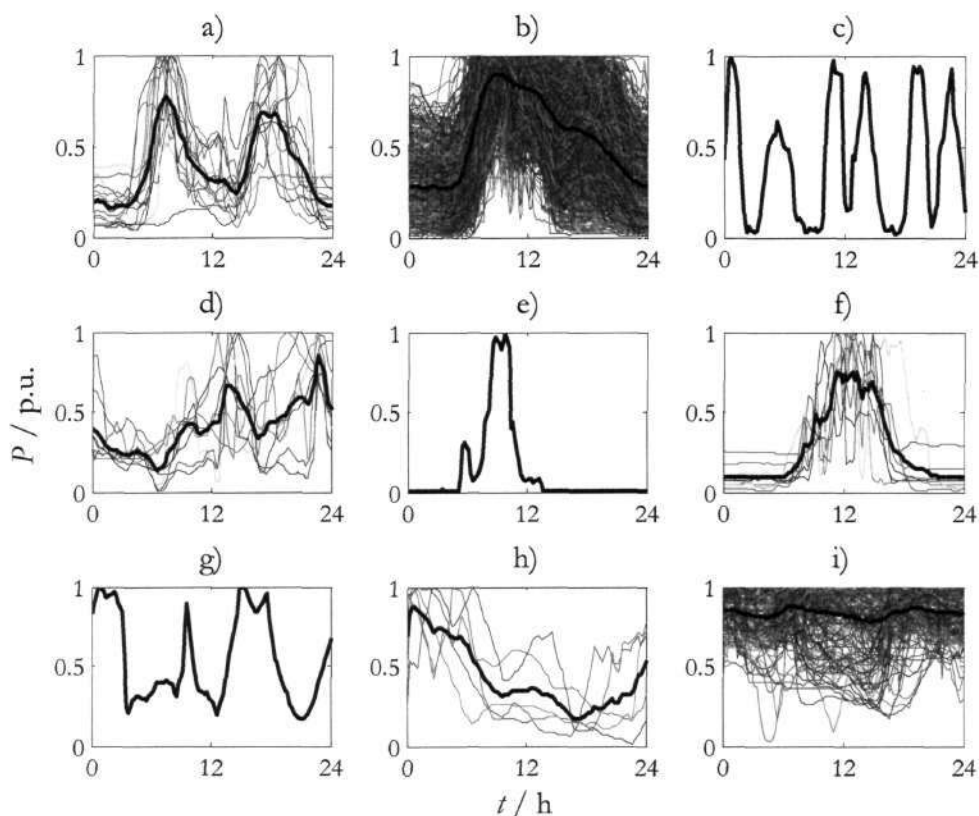


Sl. 5.5: Razdalje med zaporednima skupinama v odvisnosti od končnega števila skupin.

Povprečna razdalja med skupinami.

Na Sl. 5.5 se razdalje med sosednjima skupinama z večanjem števila končnih skupin v splošnem zmanjšujejo. Kljub temu se jasno vidi število skupin, tj. 3, 9, 17, 27 in 29, pri katerih so razdalje med zaporednima skupinama večje in predstavljajo naslednja potencialno primerna končna števila skupin, v katere naj bi rojili OD-je. Njihovo končno število določimo na

podlagi natančnejšega pregleda OD-jev, ki posamezno skupino sestavljajo. Zaradi velikega raztrosa pri razvrščanju OD-jev samo v tri skupine smo v naslednjem koraku analizirali razvrstitev OD-jev v devet skupin. Rezultate kaže Sl. 5.6. Vidi se, da imamo še vedno dve močno zastopani skupini b) in i) s 858 in 199 OD-ji, ki imata še vedno velik raztros ter sedem skupin s skupno 38 OD-ji. Tri skupine sestavlja samo en OD.

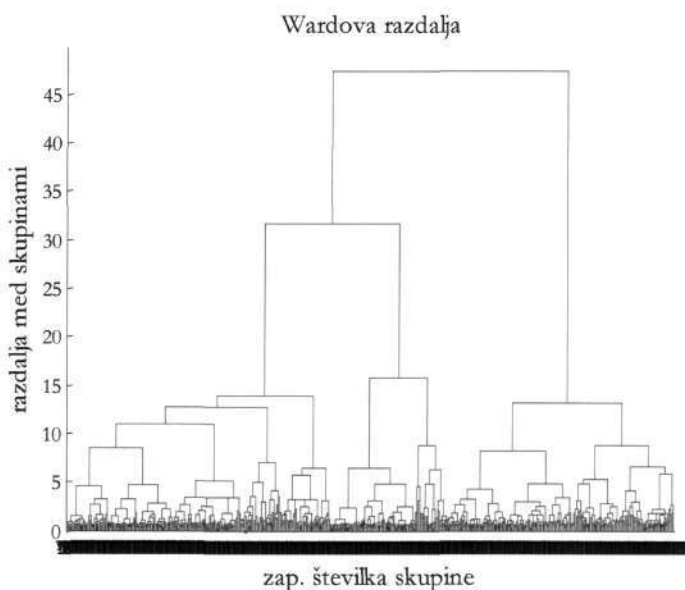


Sl. 5.6: Rezultati razvrstitve OD-jev v devet skupin; povprečna razdalja med skupinami.

Analiza rezultatov razvrstitve kljub trikratnemu povečanju števila skupin kaže, da nismo zmanjšali raztrosa najbolj zastopanih skupin. Še več, skupini a) in c) na Sl. 5.4 sta ostali nespremenjeni, prikazujeta ju skupini h) in i) na Sl. 5.6. S povečanjem skupin na devet smo iz skupine b) na Sl. 5.4 izločili šest skupin, ki predstavljajo posebneže. Naslednji korak je razvrstitve OD-jev v večje število skupin. Rezultate razvrstitve OD-jev v 27 skupin prikazujeta Sl. 1 in Sl. 2, ki smo jo podali v prilogi. Iz slik se vidi, da večina skupin predstavlja posebneže, ki smo jih s povečanjem števila skupin izločili iz skupine b) in i) na Sl. 5.6.

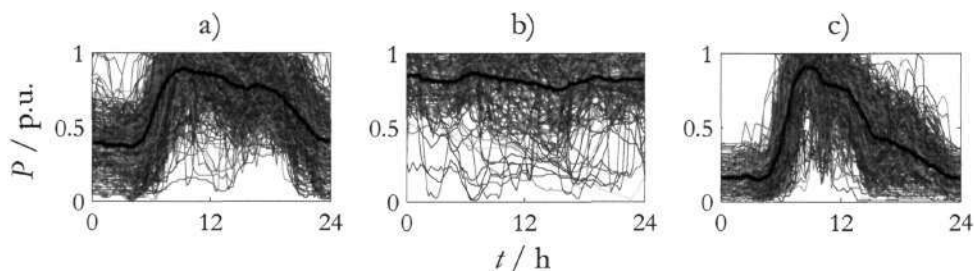
Na podlagi predstavljenih rezultatov ugotavljamo, da rojenje OD-jev z izračunom povprečne razdalje med skupinami ne oblikuje splošnih skupin, temveč išče posebneže, ki jim dodeli samostojne skupine. Te skupine na višjem nivoju združevanja pridruži že obstoječim skupinam in ne oblikuje posebne skupine, ki bi predstavljala samo posebneže. To je v nasprotju s ciljem določanja končnega števila skupin, ki smo ga opredelili v začetnem odstavku tega razdelka, kjer smo zapisali, da ni naš namen iskanje posebnežev v množici OD-jev, temveč oblikovanje splošnih skupin. Da dobimo ustrezno število skupin z relativno nizkim

raztrosom, moramo OD-je razvrstiti v veliko skupin. Menimo, da algoritem, ki izračunava povprečno razdaljo med skupinami, ni primeren za rojenje OD-jev in določanje ustreznega števila končnih skupin. Zato smo uporabili algoritem z Wardovo razdaljo. Rezultate v obliki drevesa združevanja ali dendrograma prikazuje Sl. 5.7.



Sl. 5.7: Dendrogram rojenja OD-jev. Wardova razdalja med skupinami.

Iz dendrograma na Sl. 5.7 se vidi, da algoritem razvrsti OD-je v tri ločene skupine, ki se nadalje delijo v manjše podskupine. Rezultate razdelitve v tri skupine prikazuje Sl. 5.8.

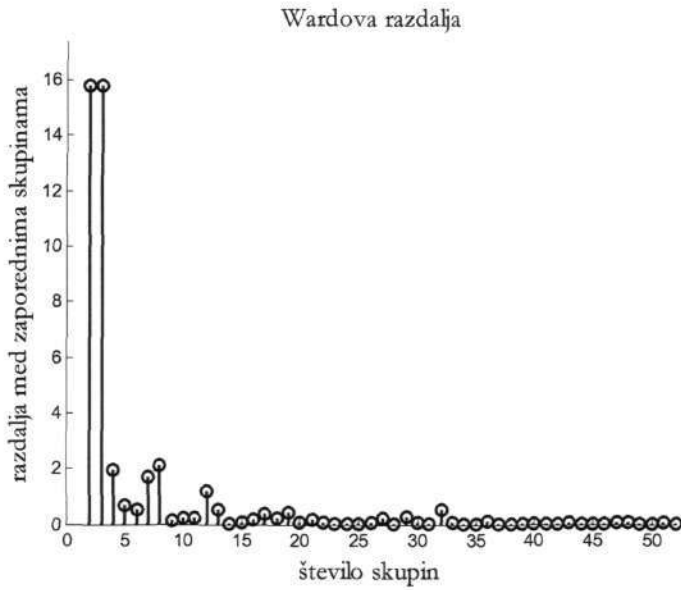


Sl. 5.8: Rezultati razvrstitve OD-jev v tri skupine; Wardova razdalja med skupinami.

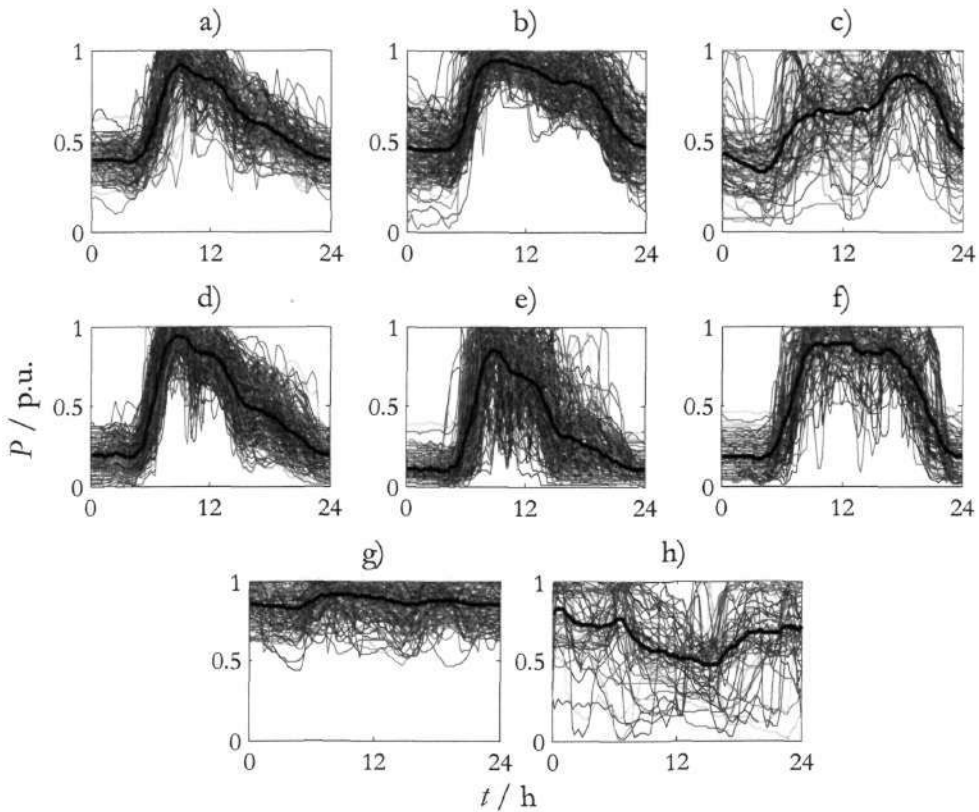
Iz zgornje slike se vidi, da se skupine med seboj jasno razlikujejo. Posamezno skupino sestavlja 475, 211 in 409 OD-jev, kar ima za posledico njihov velik raztros. Da zmanjšamo raztros skupin, OD-je razvrstimo v več skupin. Naslednje število potencialno zanimivih skupin določimo na podlagi razdalje med zaporednima skupinami, ki jih prikazuje Sl. 5.9. Na sliki smo podobno kot v predhodnem primeru prikazali vrednosti razdalj samo za prvih 52 skupin.

Iz Sl. 5.9 se vidi, da sta razvrstitvi v dve ali tri skupine praktično enakovredni, saj je razdalja med zaporednima skupinama praktično enaka, vendar bi z združitvijo skupine a) in b) na Sl. 5.8 močno povečali raztros združene skupine. Za nadaljnjo analizo so zanimive še razvrstitve

OD-jev v 8, 12 ali 32 skupin. Razvrstitve v 32 skupin ne bomo posebej analizirali saj je veliko število skupin v nasprotju s zahtevami iz razdelka 2.5.2, ki jih morajo NOD-i izpolnjevati. Rezultate razdelitve OD-jev v osem skupin prikazuje Sl. 5.10.



Sl. 5.9: Razdalje med zaporednima skupinama v odvisnosti od končnega števila skupin. Wardova razdalja med skupinami.



Sl. 5.10: Rezultati razvrstitve OD-jev v osem skupin; Wardova razdalja med skupinami.

Iz Sl. 5.10 se vidi, da razvrstitev OD-jev v osem skupin daje jasne skupine, ki se razlikujejo med seboj, vendar imamo še vedno skupine, ki imajo dokaj velik raztros. Takšne skupine so predvsem c), e) in h). Zato menimo, da je primerneje OD-je razvrstiti v več skupin. Poleg tega se iz slike vidi, da razvrstitev z izračunom Wardove razdalje oblikuje skupine posebnježev, (glej skupino h), ter ne razvrsti vsakega posebnježa v samostojno skupino, kakor smo to videli pri izračunu povprečne razdalje.

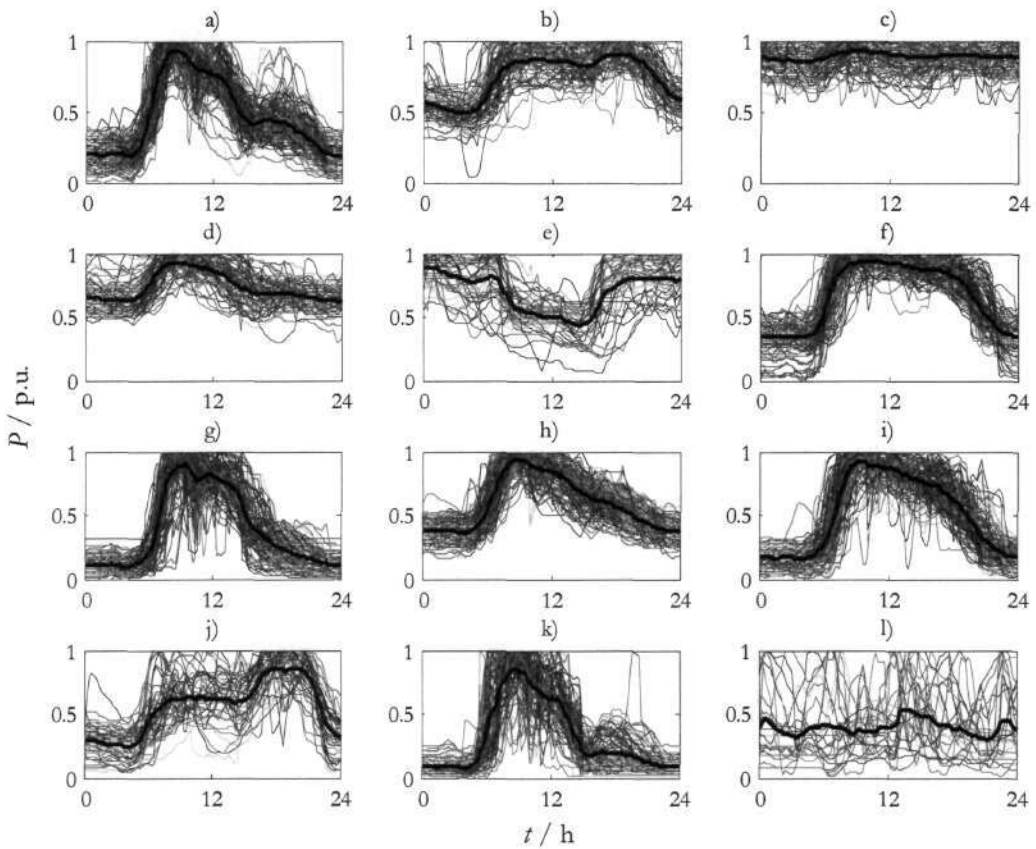
Na podlagi zgornjih ugotovitev menimo, da je 12 najprimernejše število skupin, v katere bomo OD-je razvrstili s hierarhičnimi metodami. Število sovпада z okvirnim številom skupin, ki smo jih določili s FCM algoritmom in kriterijem  $f_1$  na Sl. 5.2. Zato menimo, da je 12 končno in optimalno število skupin, v katere bomo OD-je razvrščali. Rezultate razvrstitev z uporabo vseh treh predlaganih metod podajamo v nadaljevanju.

### 5.1.2. Rezultati razvrščanja OD-jev s FCM algoritmom

V predhodnem razdelku smo z uporabo mer veljavnosti rojenja in razdalj med sosednjima skupinama določili 12 za optimalno število skupin. Poleg določitve optimalnega števila skupin moramo pri razvrščanju s FCM algoritmom določiti tudi optimalno izbiro pripadnostne matrike  $U$  ter potenčnega faktor mehčanja  $m$ .

V literaturi, ki obravnava področje razvrščanja s FCM algoritmom, avtorji ne navajajo postopkov, kako določiti ali izbrati ustrezno vrednost potenčnega faktorju mehčanja  $m$ . Navajajo samo vrednosti, ki so jih uporabljali. V delu smo uporabili  $m = 1,2$  na podlagi analize, v kateri smo spreminjali njegovo vrednost. Rezultat povečevanja parametra  $m$  je zmanjševanje števila različnih oblik povprečnih OD-jev. Pri vrednosti  $m = 2$  smo dobili že dve skupini odjemalcev, ki imata različni obliki povprečnih OD-jev, vendar se ponovita: prva skupina se ponovi dvakrat, medtem ko se druga skupina ponovi trikrat. Cilj pri razvrščanju je v čim večjem medsebojnem razlikovanju skupin, zato smo za  $m$  izbrali podano vrednost.

Za optimalno smo določiti tisto izbiro, ki je po končni razvrstitvi in upoštevanju "trde" pripadnosti OD-jev v posameznih skupinah imela minimalno vrednost kriterija  $f_1$ . Rezultate razvrščanja prikazuje Sl. 5.11, kjer vidimo 12 skupin, ki smo jih prikazali na posamezni manjši sličici z zaporedno črko od a) do l). Ordinatne osi predstavljajo čas v urah, medtem ko abscisne osi predstavljajo moč  $P$  v p.u. Za vsako izmed skupin smo izračunali povprečni diagram (PD) vseh OD-jev v skupin ter ga predstavili z debelejšo črno črto. Iz slike se vidi, da so nekatere skupine zelo jasne ter dobro zastopane kot npr. skupina a), c), f), ali h), medtem ko so med njimi tudi skupine, ki so slabše zastopane z OD-ji. Takšne skupine so e), j) in l). Podrobnejše rezultate zastopanosti OD-jev v posamezni skupini podaja Tab. 5.1. Iz nje se vidi, da imamo tri skupine, ki imajo štiri ali manj odstotni delež vseh razvrščenih OD-jev. Med njimi predstavlja skupina l) samo posebnježe, ki imajo zelo različne oblike OD-jev. Preostali dve skupini e) in j), ki sta prav tako manj zastopani z OD-ji in imata v povprečju večjo standardno deviacijo, sestavljajo OD-ji dokaj podobnih, vendar ne pogosto zastopanih oblik.



Sl. 5.11: Rezultati razvrstitve s FCM algoritmom.

Tab. 5.1: Število in delež OD-jev v posamezni skupini.

skupina	število OD-jev	delež / %	skupina	število OD-jev	delež / %
a)	150	14	g)	101	9
b)	75	7	h)	132	12
c)	125	11	i)	113	10
d)	79	7	j)	44	4
e)	40	4	k)	84	8
f)	129	12	l)	23	2

Za razločevanje skupin in njihovih povprečnih OD-jev med seboj so pomembni naslednji podatki:

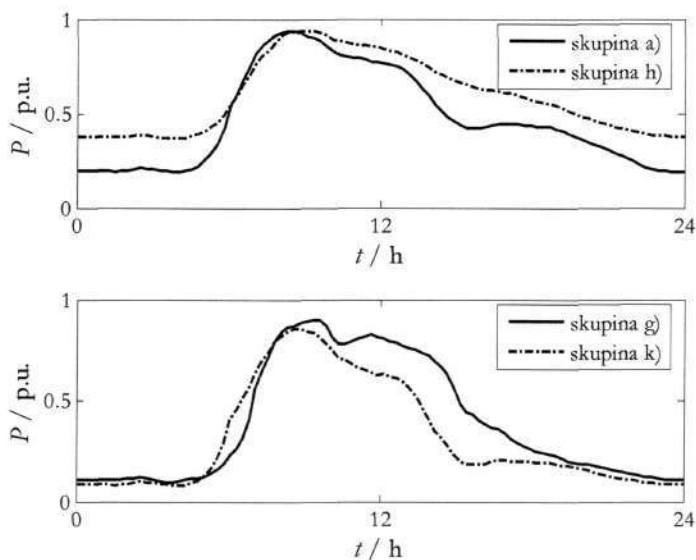
- širina koničnega dela,
- strmina dviga in padca obremenitve,
- višina pasovnega dela,
- razgibanost vršnega dela.

Vpliv teh podatkov nazorno prikazuje Sl. 5.12, kjer vidimo primerjavo povprečnih diagramov skupin a) in h) ter g) in k). Prva dva diagrama prikazujeta vpliv višine pasovnega dela na OD. Vidimo, da odjemalci začnejo in končajo obratovati ob istem času, vendar imajo prvi nižjo



nočno ali pasovno porabo. Odjemalci skupine a) imajo hitrejši začetni padec porabe, ki se v nadaljevanju za določen čas nekoliko dvigne in nato zopet pade, medtem ko poraba odjemalcev skupine h) pada enakomerno.

Vpliv širine koničnega dela in razgibanosti vršnega dela povprečnega OD-ja kažeta grafa g) in k). Vidimo, da ima skupina g) širši in bolj razgiban konični del, saj se poraba po začetnem padcu ponovno dvigne, medtem ko se poraba pri skupini k), potem ko doseže maksimalno vrednost, dokaj enakomerno znižuje do 15. ure.

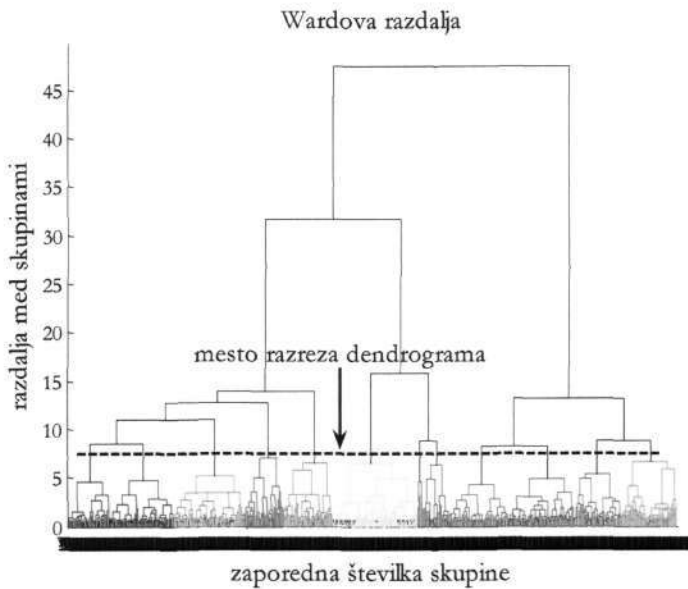


Sl. 5.12: Primerjava med povprečnimi diagrami skupin a) in h) ter g) in k).

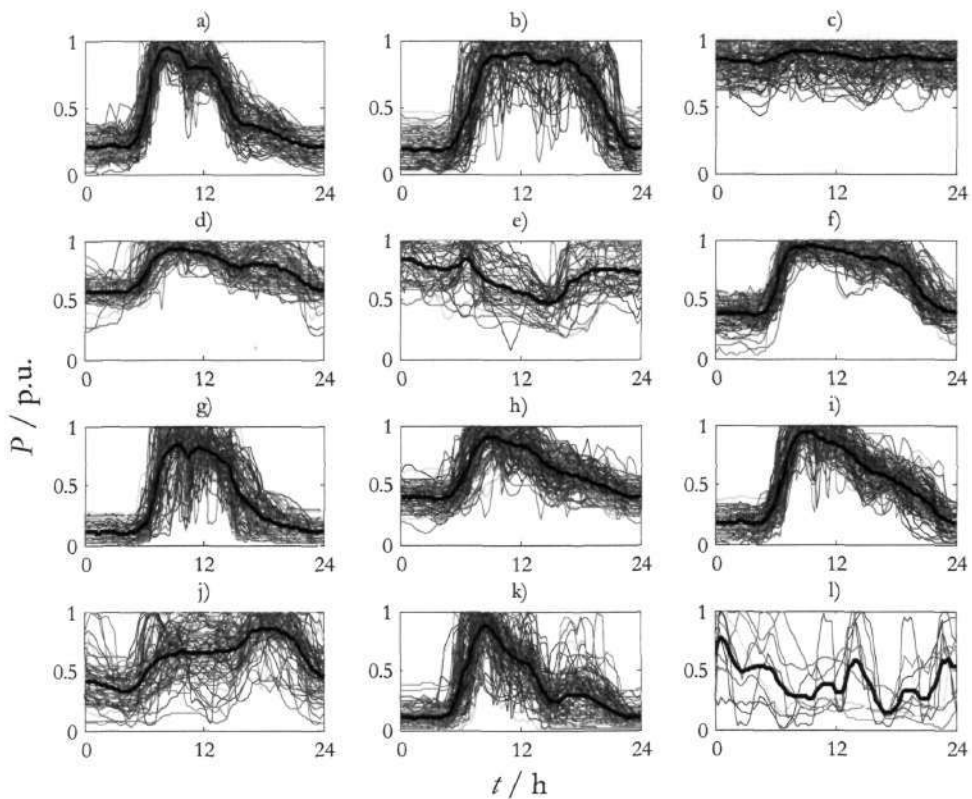
### 5.1.3. Rezultati razvrščanja s hierarhično metodo

Druga metoda razvrščanja, ki smo jo preizkusili, je hierarhična metoda z izračunom Wardove razdalje med skupinami. Rezultate razvrščanja ali končne skupine OD-jev določimo na podlagi razreza dendrograma pri znani razdalji med skupinami. Za lažjo predstavitev razreza in določitve ustrezne razdalje si dendrogram iz Sl. 5.7 oglejmo še enkrat. Prikazuje ga Sl. 5.13. Če dendrogram razrežemo pri razdalji 8 med skupinami, kakor smo to označili na zgornji sliki, ta razpade na 12 samostojnih skupin, ki jih prikazuje Sl. 5.14. Iz dendrograma se razločno vidi katere skupina OD-jev nastanejo iz treh skupin OD-jev na Sl. 5.8, ko povečamo končno število skupin. Skupino a) na Sl. 5.8, ki predstavlja odjemalce s trapeznim OD-jem, predstavljajo skupine b), d), f), h) in j) na Sl. 5.14. Skupino b) s konstantnim ali pasovnim odjemom predstavljajo skupine c), l) in e) na Sl. 5.14, medtem ko skupino c) na Sl. 5.8 predstavljajo skupine a), i), g) in k) na Sl. 5.14. Rezultate smo označili na podoben način, kakor pri razvrščanju s FCM algoritmom. Iz Sl. 5.14 se vidi, da so nekatere skupine jasne ter dobro zastopane kot npr. skupina a), c), f), ali h), medtem ko so med njimi tudi skupine, ki so slabše zastopane z OD-ji. Takšne skupine so e) in l). Podrobnejše rezultate zastopanosti OD-jev v skupini podaja Tab. 5.1. Iz nje se vidi, da imamo za razliko od rezultatov FCM algoritma samo

dve skupini, ki imata štiri ali manj odstotni delež vseh OD-jev. Skupina l) ima zelo malo OD-jev, samo 9.



Sl. 5.13: Dendrogram rojenja OD-jev; Wardova razdalja med skupinami; rezec dendrograma na 12 skupin.



Sl. 5.14: Rezultati razvrstitve s hierarhično metodo; Wardova razdalja med skupinami.

Tab. 5.2: Število in delež OD-jev v posamezni skupini.

skupina	število OD-jev	delež / %	skupina	število OD-jev	delež / %
a)	106	10	g)	82	7
b)	88	8	h)	122	11
c)	155	14	i)	121	11
d)	76	7	j)	73	7
e)	47	4	k)	100	9
f)	116	11	l)	9	1

Pri razvrščanju s hierarhično metodo smo želeli preveriti optimalno število skupin, v katere smo OD-je razvrstili. Uporabili smo razdaljo med sosednjima skupinama, na podlagi katere smo določili 12 za končno število skupin, vendar bi lahko izbrali tudi katero izmed mer veljavnosti rojenja iz razdelka 3.2.1.1. Kriterij  $f_2$  se je izkazal za neprimerne pri razvrščanju OD-jev, zato smo uporabili kriterij  $f_1$ . Rezultate izračuna kriterija  $f_1$  bomo podrobneje predstavili in komentirali v razdelku 5.1.5, kjer podajamo primerjavo metod razvrščanja.

#### 5.1.4. Rezultati razvrščanja s samoorganizirajočim omrežjem

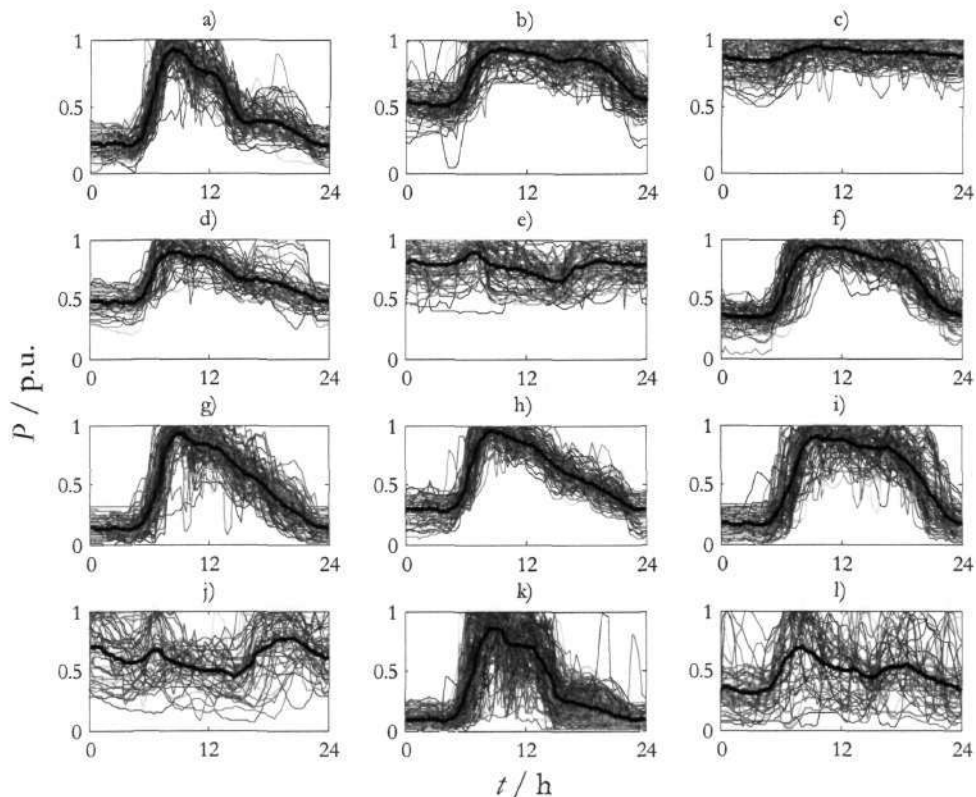
Za razvrščanje s SOM-om smo izbrali omrežje s številom nevronov, ki ustreza zelenemu številu rojev ali skupin. Pri tem predstavlja vsak nevron eno skupino. Zaradi primerjave rezultatov razvrščanja s preostalima dvema metodama smo izbrali 12 nevronov, ki so bili razporejeni v dvodimenzionalno mrežo. V postopku iskanja optimalne strukture nevronskega omrežja smo spreminjali njihovo dimenzijo med  $3 \times 4$  in  $4 \times 3$ , strukturo med pravokotno in heksagonalno ter vrednost radija  $R$  in števila iteracij učenja. Optimalno omrežje smo določili z uporabo kriterija  $f_1$ . Razvrstitev OD-jev in s tem nevronske omrežje, ki je imelo minimalno vrednost kriterija  $f_1$ , smo izbrali za najboljše.

Na podlagi tega kriterija smo za optimalno določili pravokotno strukturo omrežja dimenzije  $3 \times 4$  z radijem  $R = 1,2$  ter število iteracij učenja 2000. Rezultate razvrščanja s SOM-om prikazuje Sl. 5.15. Vidimo, da so nekatere skupine jasne ter dobro zastopane kot npr. skupina a), c), f), ali h), medtem ko so med njimi tudi skupine, ki so slabše zastopane z OD-ji. Takšne skupine so j) in l). Poleg tega imamo tri skupine, ki imajo večji raztros, skupine j), k) in l). Podrobnejše rezultate zastopanosti OD-jev v skupini podaja Tab. 5.3.

Tab. 5.3: Število in delež OD-jev v posamezni skupini.

skupina	število OD-jev	delež / %	skupina	število OD-jev	delež / %
a)	89	8	g)	93	8
b)	107	10	h)	96	9
c)	121	11	i)	95	9
d)	58	5	j)	45	4
e)	63	6	k)	157	14
f)	121	11	l)	50	5

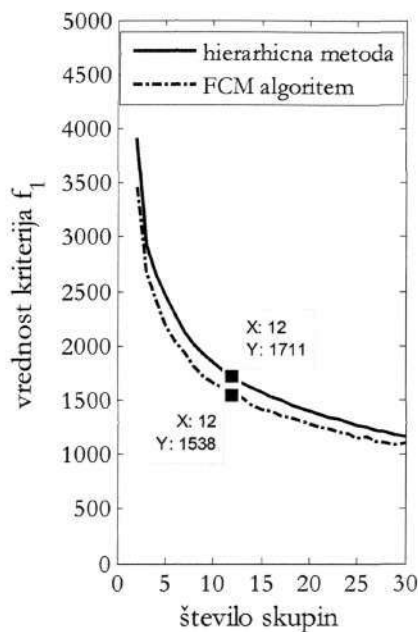
Iz zgornje tabele se vidi, da samo eno skupino sestavlja štiri ali manj odstotni delež vseh OD-jev. Skupina l) je prav tako slabše zastopana, medtem ko ima skupina k) največ OD-jev izmed vseh skupin neodvisno od uporabljene metode razvrščanja.



Sl. 5.15: Rezultati razvrstitve s SOM-om.

### 5.1.5. Primerjava metod razvrščanja

Do sedaj smo rezultate razvrščanja z uporabo treh različnih metod predstavili posamezno. Da določimo najprimernejšo metodo razvrščanja, bomo rezultate medsebojno primerjali. Metode smo primerjali na podlagi vrednosti kriterija  $f_1$  (3.24), ki predstavlja vsoto kvadratnih odstopanj OD-jev razvrščenih v skupino od povprečnega OD-ja skupine. Boljša je tista metoda razvrščanja, ki ima manjšo vrednost kriterija. FCM algoritem in hierarhično razvrščanje z izračunom Wardove razdalje med skupinam smo primerjali pri različnem številu skupin. Rezultate primerjave, ko spreminjamo število skupin od dva do 30 kaže Sl. 5.16. Vidi se, da je razvrščanje s FCM algoritmom boljše od razvrščanja s hierarhično metodo neodvisno od končnega števila skupin. Primerjava poteka kriterija  $f_1$  pri razvrščanju z obema metodama kaže, da oba poteka izkazujeta koleno pri 12 skupinah. Kar dodatno potrjuje pravilnost izbire optimalnega števila skupin. Še več, ugotavljamo, da je kriterij  $f_1$  splošen in neodvisen od izbrane metode razvrščanja ali vzorcev, ki jih razvrščamo. Tega ne moremo trditi za kriterij  $f_2$ , ki daje boljše rezultate kot kriterij  $f_1$  pri razvrščanju vzorcev, opisanih z dvema vrednostma, medtem ko ni primeren za razvrščanje OD-jev, opisanih s 96 vrednostmi.



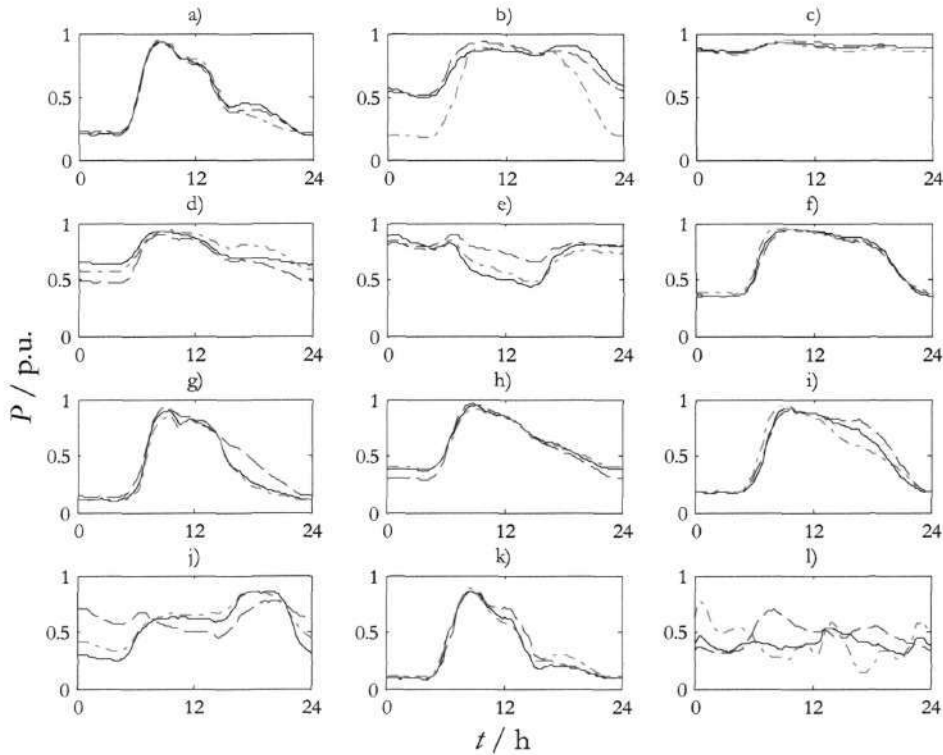
Sl. 5.16: Primerjava vrednosti kriterija  $f_1$  za hierarhično metodo in FCM algoritem.

Pri razvrščanju OD-jev s SOM-om nismo računali odvisnosti kriterija  $f_1$  od končnega števila skupin, v katere bi razvrščali OD-je. Analizirali smo samo razvrstitev v 12 skupin, kjer je kriterij  $f_1$  znašal 1704. Pri ostalih dveh metodah je  $f_1$  pri razvrstitvi v 12 skupin znašal 1538 za FCM algoritem in 1711 za hierarhično metodo, kar smo tudi označili na Sl. 5.16. Na podlagi primerjave vrednosti  $f_1$  ugotavljamo, da je FCM algoritem najprimernejši za razvrščanje OD-jev. Poleg tega razvrščanje s FCM algoritmom omogoča določitev ustreznega praga pripadnosti OD-jev  $u_{ij}$ , na podlagi katerega lahko izločimo vse OD-je, ki imajo maksimalno pripadnost v končni skupini manjšo od praga. Tako bi dodatno izločili posebneže, vendar tega v delu nismo uporabili.

Kljub ugotovitvi na podlagi kriterija  $f_1$ , da je FCM algoritem najprimernejša metoda razvrščanja, podajamo tudi grafično primerjavo povprečnih OD-jev skupin. Prikazuje jo Sl. 5.17, kjer smo s črno polno črto označili rezultate FCM algoritma, z modro črtkano črto rezultate hierarhične metode in rdečo črto rezultate SOM-a. Iz slike se vidi, da skupine, dobljene z vsemi tremi metodami, ločijo naslednje lastnosti:

- Skupine imajo praktično enake povprečne OD-je, glej skupini c) in f);
- Skupine imajo na posameznem delu enake OD-je, medtem ko imajo v popoldanskem in večernem delu manjša odstopanja, glej skupine a), i) in k). Takšna skupina je tudi h), ki ima manjša odstopanja v jutranjem in večernem delu;
- Skupine imajo primerljive povprečne OD-je, glej skupine d), e) in l);
- Dve metodi imata primerljive ali praktično enake OD-je, medtem ko tretja metoda ne oblikuje primerljivega OD-ja ali ima večja odstopanja, glej skupine b), g) in j). V prvem primeru SOM-i ne oblikujejo primerljive skupine, saj je OD primerljiv z OD-ji skupine i) in ne skupine b). Pri skupinah g) in j) vidimo, da FCM algoritem in

hierarhična metoda oblikujejo enake ali primerljive OD-je, medtem ko imata OD-ja, dobljena s SOM-i, večja ali manjša odstopanja.



Sl. 5.17: Primerjava rezultatov treh metod razvrščanja; FCM algoritem - črna črta, hierarhična metoda - rdeča črta, SOM - modra črta.

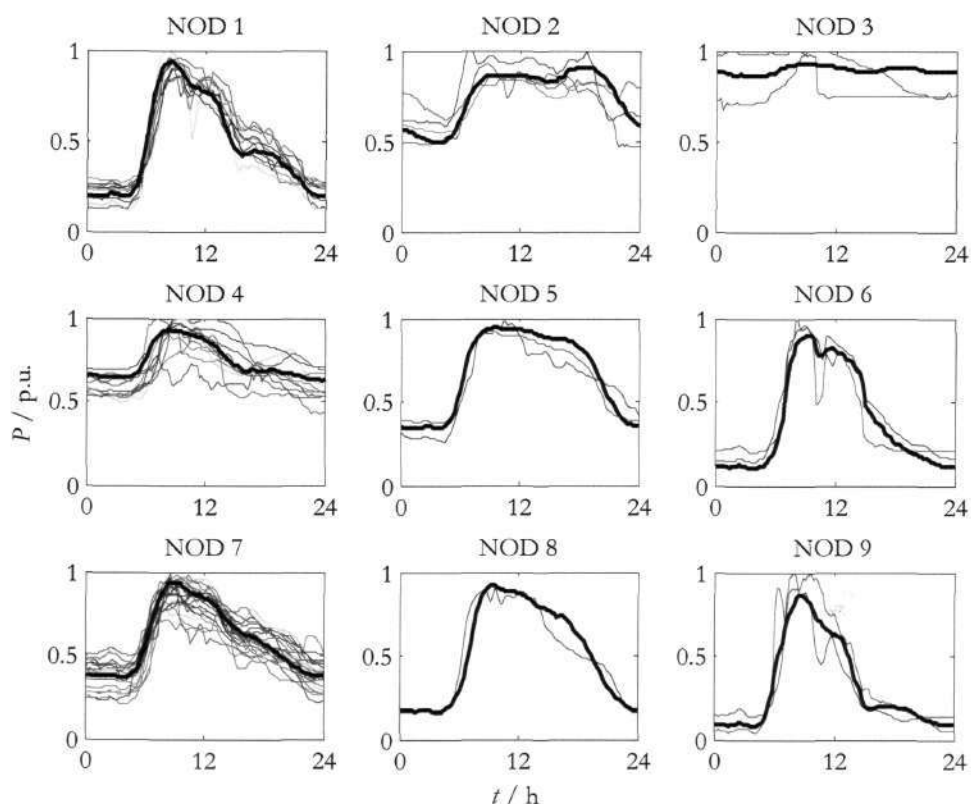
### 5.1.6. Rezultati določitve in dodelitev NOD-ov posameznim UO-jem

V tem razdelku predstavljamo rezultate zadnjega koraka oblikovanja NOD-ov. Postopek smo teoretično predstavili v razdelku 3.3. Najprej smo izračunali povprečne OD-je glavnih dejavnosti ter jih nato razvrstili z uporabo verjetnostnih nevronske omrežji v skupine, ki smo jih določili s FCM algoritmom. Rezultat razvrščanja je devet NOD-ov, ki jih prikazuje Sl. 5.18. Na sliki imajo NOD-i zaporedne številke od ena do devet. Prikazani so s poudarjeno črto. Ostali diagrami predstavljajo povprečne OD-je posameznih registriranih dejavnosti. Iz slike vidimo, da nekateri NOD-i zastopajo več dejavnosti, medtem ko NOD 8 zastopa samo dejavnost 22. Tab. 5.4 podaja, katere dejavnosti zastopa posamezen NOD.

Primerjava NOD-ov, ki jih prikazuje spodnja slika, ter povprečnih OD-jev posameznih podobnih skupin, ki jih prikazuje Sl. 5.11 kaže, da treh podobnih skupin nismo predlagali za NOD-e. Te skupine sestavljajo OD-ji relativno netipičnih oblik in imajo v povprečju večji raztros. To še posebej velja za skupino l) na Sl. 5.11. Vse tri skupine zastopa najmanj OD-jev, skupaj samo 107, kar predstavlja približno 10 % vseh analiziranih OD-jev. Na drugi strani NOD-i zastopajo 90 % vseh analiziranih OD-jev, kar je eden izmed ciljev njihovega oblikovanja. Že v razdelku 2.5.2 smo zapisali, da morajo OD-ji predstavljati čim širšo



populacijo odjemalcev, zato menimo, da so predlagani NOD-i dovolj splošni in primerni za nadaljnjo uporabo.



Sl. 5.18: Končni NOD-i.

Tab. 5.4: Glavne dejavnosti v posameznem NOD-u,

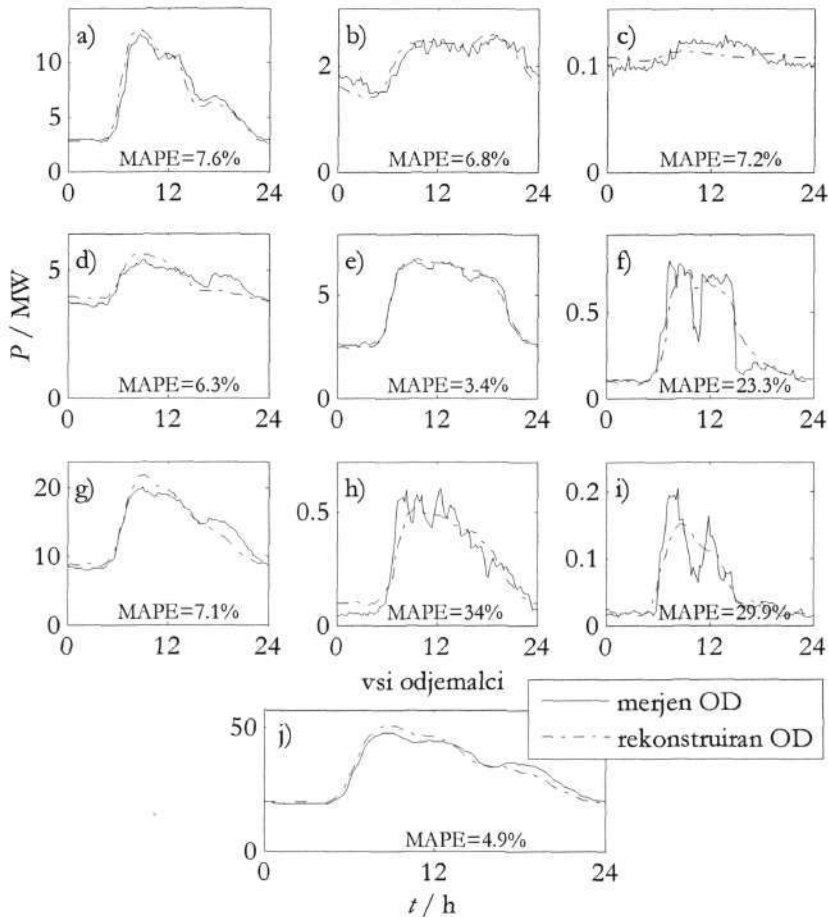
NOD	Glavne dejavnosti po SKD 2003
1	2 19 20 26 29 34 45 80 85 91
2	5 21 41 92
3	13 99
4	11 30 32 40 60 62 64 70 72 90
5	35 52
6	36 37
7	1 10 14 15 17 18 24 25 27 28 31 33 50 51 55 63 65 66 73 74 75 93
8	22
9	71 95

## 5.2. Primerjava meritev in rekonstruiranih OD-jev

Celoten postopek določanja NOD-ov smo hoteli tudi preveriti na dejansko izmerjenih OD-jih odjemalcev. Poudariti moramo, da primerjav, katerih rezultate bomo prikazali v tem razdelku, ne moremo izvajati v realnem distribucijskem sistemu. Razlog je preprost. Nimamo meritev OD-jev odjemalcev, saj so NOD-i namenjeni odjemalcem brez sprotnih meritev. Kljub temu

smo NOD-e testirali in jih primerjali z dejanskimi meritvami, ki smo jih pridobili od slovenskih distribucijskih podjetij. Analizirali smo delavnike 1095 odjemalcev v mesecu novembru 2003.

Glede na predlagan postopek dodeljevanja NOD-ov smo oblikovali devet skupin odjemalcev, katerim smo dodelili NOD-e. Iz meritev smo izračunali OD-je posameznih skupin odjemalcev in jih primerjali z njihovimi rekonstruiranimi OD-ji. Predpostavili smo, da poznamo dnevno energijo odjemalcev. Rezultate primerjave za en delavnik za posamezno skupino odjemalcev in vse odjemalce skupaj kaže Sl. 5.19.



Sl. 5.19: Primerjav meritev in rekonstruiranih OD-jev.

Iz Sl. 5.19, na kateri smo podali tudi povprečna procentualna odstopanja (ang. Mean Average Percentage Error, MAPE) rekonstruiranega OD-ja od merjenega, se vidi, da en rekonstruirani OD, glej graf e), v povprečju zelo malo odstopa od meritev. Pet OD-jev odstopa v povprečju med 6 % in 8 %, medtem ko imajo trije OD-ji, glej grafe f), h) in i), v povprečju dokaj velika odstopanja. Razlog za tako velika odstopanja je v majhnem številu analiziranih odjemalcev, ki sestavljajo te tri skupine. Ne glede na večja odstopanja teh skupin vidimo, da oblika rekonstruiranih sledi obliki merjenih OD-jev. Če bi imeli več odjemalcev, bi zaradi vpliva faktorja istočasnosti bila odstopanja manjša. To se vidi iz primerjave merjenega in rekonstruiranega OD-ja vseh odjemalcev. Primer nazorno kaže, da je uporaba NOD-ov

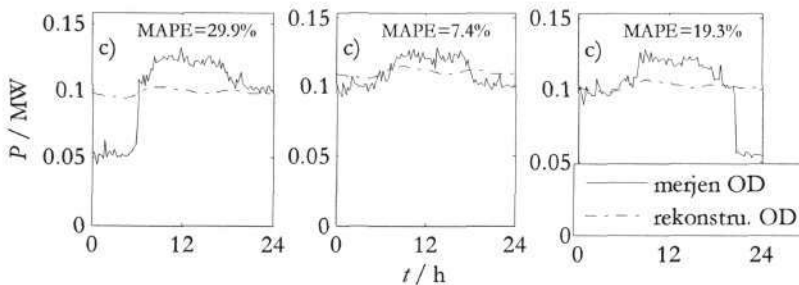
smiselna samo pri hkratni obravnavi večjega števila odjemalcev in ne posameznega odjemalca. Iluzorno je pričakovati, da bi bili sposobni izdelati takšne NOD-e, s katerimi bi bili sposobni relativno natančno opisati porabo posameznega odjemalca.

V sklopu preverjanja NOD-ov smo analizirali vse delavnike v mesecu novembru 2003. Pri tem smo opazovali MAPE med merjenim in rekonstruiranim OD-jem vseh skupin. Rezultate za minimalno, maksimalno in povprečno vrednost odstopanj 20 delavnikov prikazuje Tab. 5.5.

Tab. 5.5: Število in delež OD-jev v posamezni skupini.

skupina	MAPE		
	min.	max.	povprečje
a)	6,5	13,1	8,46
b)	5,6	8,6	7,18
c)	2,4	29,9	12,95
d)	5,2	7,8	6,6
e)	2,8	4,3	3,37
f)	13	41,6	22,77
g)	5,2	11,8	7,66
h)	15,8	47,2	28,34
i)	25,2	43,7	33,47
j)	3,6	8,5	5,41

Iz zgornje tabele se vidi, da je v povprečju večina odstopanj posameznih skupin primerljiva z odstopanji prikazanimi na Sl. 5.19. Skupine a), b), d) in e) imajo majhen razpon med minimalnimi in maksimalnimi odstopanji, kar kaže, da so OD-ji teh skupin odjemalcev za delavnik med seboj dokaj podobni. Skupine f), h) in i) imajo večji razpon, zato so njihovi OD-ji od delavnika do delavnika zelo spremenljivi, vendar je to pričakovano zaradi majhnega števila odjemalcev, vključenih v analizo. Za nadaljnjo analizo je zanimiva predvsem skupina c), katere razpon med minimalnimi in maksimalnimi odstopanji je prav tako dokaj velik. Podrobnejša analiza kaže, da imamo večja odstopanja samo v ponedeljek in petek. Razloge za porast odstopanj v teh dveh dnevih kaže Sl. 5.20, kjer smo prikazali rekonstruirane in merjene OD-je treh delavnikov; ponedeljek, sreda in petek. Iz slike se vidi, da odjemalci del ponedeljka in del petka obratujejo s polovično močjo, medtem ko smo mi predvidevali, da obratujejo vse delavnike v pasu.



Sl. 5.20: Primerjava rekonstrukcije treh delavnikov skupine c).

Prikazan primer je zelo specifičen in bi ga ob večjem številu analiziranih odjemalcev težje opazili, saj bi se zaradi vpliva faktorja istočasnosti odstopanja zmanjšala. Še enkrat moramo poudariti, da prikazanih primerjav v realnem distribucijskem sistemu zaradi neobstoja meritev ni mogoče izvesti. Kljub temu je primer poučen, ker kaže na previdnost pri uporabi NOD-ov, ter da so v nekaterih primerih trenutna odstopanja dejanskih OD dokaj velika.

Primerjava OD-jev je mogoča smo na nivoju celotnega distribucijskega podjetja, kjer obstajajo meritve v posameznih RTP-jih in je možen izračun odstopanj med merjenim in rekonstruiranim OD-jem.

### 5.3. Rezultati optimizacije obremenitvenih diagramov

Do sedaj smo predstavili rezultate oblikovanja NOD-ov in primerjavo rekonstruiranih OD-jev posameznih skupin odjemalcev z dejanskimi meritvami. Čeprav smo rezultate optimizacije OD-jev dveh skupin odjemalcev in nekatere ugotovitve podali že ob predstavitvi in izpeljavi matematičnega modela optimizacije v razdelku 4.3, v nadaljevanju prikazujemo rezultate optimizacije vseh devetih skupin odjemalcev. V analizo smo vključili en delavnik 1095 odjemalcev v mesecu novembru 2003. Predpostavili smo, da se vsi odjemalci napajajo iz enega distribucijskega omrežja, kje poznamo dejansko preostalo krivuljo  $P_{DPK}$ .

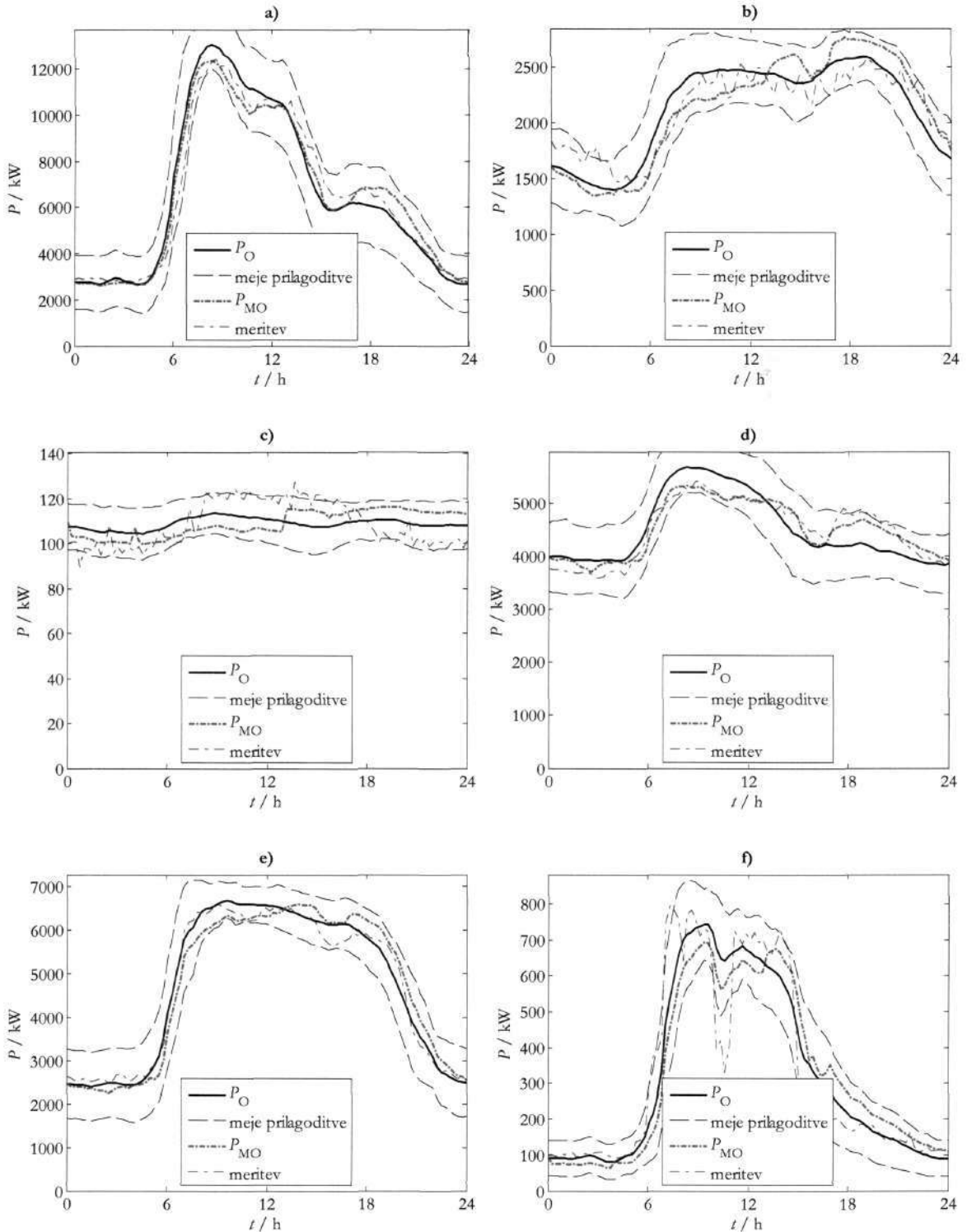
Glede na postopek dodeljevanja NOD-ov iz razdelka 3.3 smo oblikovali devet skupin odjemalcev. Z uporabo sintetičnega postopka smo določili rekonstruirane OD-je  $P_O$  posameznih skupin odjemalcev in izračunali ocenjeno  $P_{OPK}$  preostalo krivuljo. V skladu z modelom iz razdelka 4.3 smo zapisali ustrezne matrične enačbe za omejitve in kriterijsko funkcijo ter jih uporabili kot vhodne parametre v Matlab-ovi funkciji *linprog* [50], ki je namenjena reševanju problemov linearnega programiranja. Za določanje optimalne rešitve optimizacijskega problema smo uporabili metodo notranje točke.

Rezultate optimizacije OD-jev prikazujeta Sl. 5.21 in Sl. 5.22, kjer grafe posameznih skupin odjemalcev označujejo črke od a) do i), medtem ko črka j) označuje graf preostale krivulje celotnega distribucijskega omrežja. Grafi OD-jev a) do i) poleg rekonstruiranih  $P_O$ , označeni so s polno črno črto, in optimiranih  $P_{MO}$ , rdeča črtkana črta, prikazujejo tudi meje, črna črtkana črta, znotraj katerih smo dopuščali modifikacijo rekonstruiranih OD-jev in izmerjene OD-je odjemalcev, črtkana modra črta. Slednji služijo samo kot referenca za vrednotenje končnih rezultatov optimizacije ter za določitev "dejanske" preostale krivulje. Ocenjeno preostalo krivuljo na grafu j) označuje črna črta, medtem ko dejansko  $P_{DPK}$  in optimirano  $P_{MPK}$  preostalo krivuljo označujeta rdeča in črna črtkasta črta.

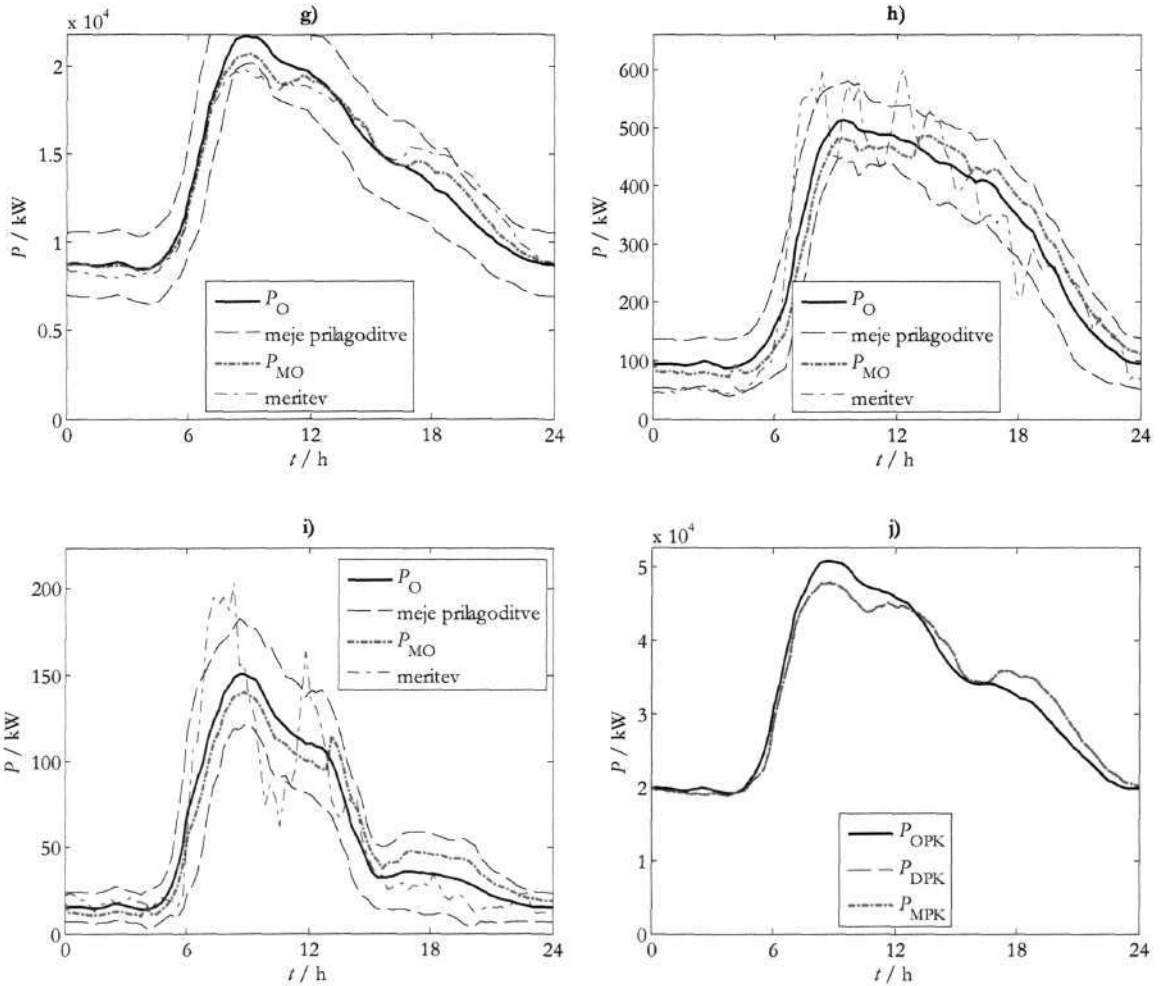
Na podlagi rezultatov optimizacije razdeljevanja preostale krivulje z uporabo NOD-ov, ki jih prikazujeta spodnji sliki, ugotavljamo naslednje:

- (i) optimirani OD-ji se nahajajo znotraj postavljenih meja popravkov in se v splošnem bolj prilagajajo merjenim OD-jem. To sklepamo na podlagi vsote absolutnih odstopanj optimiranih OD-jev od merjenih. Z optimizacijo smo vsa absolutna odstopanja

zmanjšali za 22,3 %, pri čemer smo procentualno največ odstopanj 56 % in 52 % zmanjšali skupinama odjemalcev, katerih OD-ja prikazujeta grafa d) in g) na spodnji sliki, kar se vidi iz njihovih potekov optimiranega in merjenega OD-ja. Kljub zmanjšanju vsote odstopanj smo z optimizacijo nekaterim skupinam odstopanja povečali. Skupino z izrazitim povečanjem odstopanj prikazuje graf e). Ta skupina odjemalcev ima zelo natančen NOD in s tem tudi rekonstruiran OD, kar smo videli v razdelku 5.2.



Sl. 5.21: Rezultati optimizacije rekonstruiranih OD-jev; grafi skupin odjemalcev a) do f).



Sl. 5.22: Rezultati optimizacije rekonstruiranih OD-jev; grafi skupin odjemalcev g) do i) in preostala krivulja; graf j).

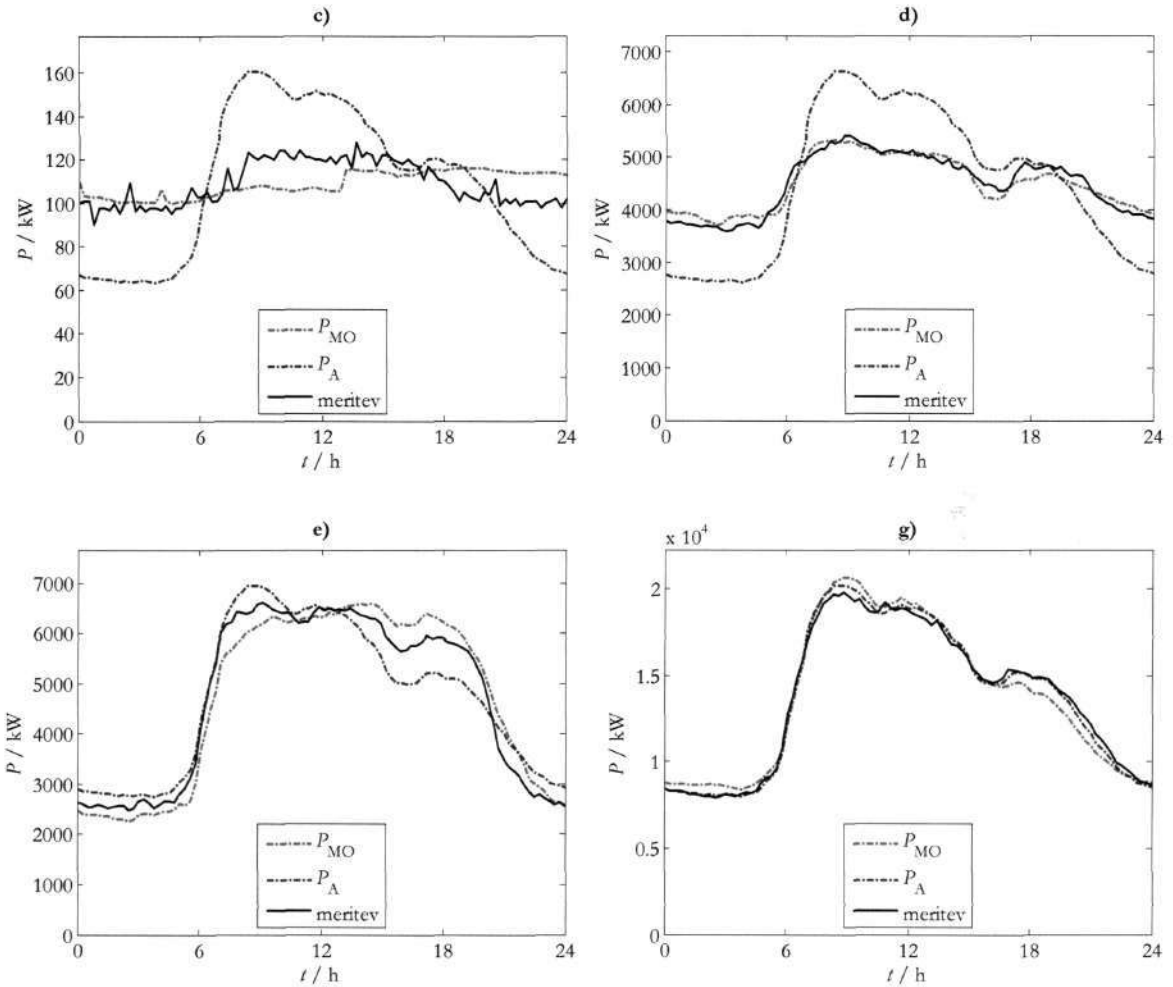
Z optimizacijo smo natančnost NOD-a zmanjšali. Težavo odpravimo z vpeljavo uteži v kriterijsko funkcijo za to skupino odjemalcev, kakor smo pokazali v razdelku 4.3. Tako ohranimo obliko rekonstruiranega OD-ja, tj. vsi pozitivni in negativni popravki moči te skupine odjemalcev so enaki nič. Posledično zmanjšamo vsoto odstopanj optimiranih OD-jev vseh skupin za dodatnih 8 % v primerjavi z optimizacijo brez uteži v kriterijski funkciji. Zaključujemo, da optimizacija z utežmi v kriterijski funkciji pri skupini odjemalcev z najnatančnejšim NOD-om najbolj zmanjša skupna odstopanja optimiranih OD-jev od merjenih;

- (ii) izbira standardnega odklona NOD-ov za določanje zgornje in spodnje meje optimizacije rekonstruiranih OD-jev, problematiko prikazuje razdelek 4.2.2, se je pokazala kot primerna rešitev, saj se nahaja večina merjenih OD-jev znotraj podanih meja;
- (iii) z optimizacijo smo odpravili odstopanja med ocenjeno preostalo krivuljo, določeno s sintetičnim postopkom iz razdelka 2.3.2 in dejansko preostalo krivuljo, določeno z analitičnim postopkom iz razdelka 2.3.3. Rezultate prikazuje graf j) na zgornji sliki.



Tako smo odpravili glavno pomanjkljivost sintetičnega postopka, ki jo podaja razdelek 2.4.

Primerjavo OD-jev nekaterih skupin odjemalcev iz Sl. 5.22, ki jih določa analitični postopek, modra črta, novo metodo iz razdelka 4, rdeča črta, in meritvami, črna črta, prikazuje Sl. 5.23.



Sl. 5.23: Primerjava OD-jev, določenih z analitičnim postopkom, predlagano metodo in meritvami za skupine odjemalcev c), d), e) in g) s Sl. 5.22.

Primerjava kaže, da uporaba nove metodologije določanja OD-jev iz poglavja 4 daje boljše rezultate, kakor analitični postopek za skupine c), d) in e), medtem ko to ne velja za skupino g), kjer daje nekoliko boljše rezultate analitični postopek. Za boljše smatramo tiste rezultate, ki imajo manjša vsota absolutnih odstopanj izračunanih OD-jev od meritev. Če se vprašamo, zakaj daje analitični postopek boljše rezultate kakor predlagana metoda, ugotovimo, da zaradi velikega deleža energije, 42 %, ki jo ima skupina g) v primerjavi z energijo vseh skupin. Tako skupina odjemalcev g) v največji meri določa obliko preostale krivulje, kar daje manjša odstopanja njenega analitično določenega OD-je od meritev. Pri vseh ostalih skupinah, tudi tistih OD-jev, ki jih nismo prikazali na zgornji sliki, daje predlagana metoda boljše rezultate.

Namen diskusije v tem razdelku je bilo pokazati, da predlagana metoda iz poglavja 4 daje boljše rezultate od obstoječih metod, tj. analitičnega in sintetičnega postopka, ker:

- predvideva različne oblike OD-jev posameznih skupin odjemalcev, ki jih določajo NOD-i,
- odpravlja odstopanja sintetično ocenjene od dejanske preostale krivulje, določene z enačbo (2.3) iz razdelka 2.3.1.

Poteke meritev OD-jev, ki jih prikazujejo Sl. 5.21 do Sl. 5.23 smo podali samo kot referenco za vrednotenje končnih rezultatov nove metode določanja porabe. Zavedati se moramo, da je namen disertacije določanje OD-jev odjemalcev brez sprotnih meritev, zato primerjava rezultatov nove metode z meritvami ni mogoča. Primerjamo lahko samo rezultate analitičnega in sintetičnega postopka ter nove metode določanje OD-jev. Privzeti moramo, da so rezultati nove metode končni.

## 5.4. Uporaba izsledkov disertacije za nadaljnje delo

Ključne ugotovitve disertacije smo podali skupaj z rezultati posameznega sklopa disertacije v predhodnih razdelkih pričujočega poglavja. V nadaljevanju podajamo nekatere smernice za nadaljnje delo, ki smo jih pridobili na podlagi ugotovitev disertacije. Rezultate iz predhodnih razdelkov utemeljujemo z izračuni in primerjavami na dejanskih meritvah obremenitev odjemalcev, ki smo jih pridobili od slovenskih distribucijskih podjetij. Izvajanje postopkov določanja NOD-ov iz poglavja 3 in optimiranja rekonstruiranih OD-jev iz poglavja 4 smo osredotočili na en tipičen dan, tj. delavnik v decembru in novembru. Pri tem smo analizirali OD-je upravičenih odjemalcev, ki imajo priključno moč med 41 kW in 300 kW. V slovenskem EES-u imamo mesečni obračunski interval plačevanja električne energije in odstopanj, zato moramo predlagane metode razširiti, da bodo omogočale njihov obračun, kakor ga udeleženci na trgu dejansko izvajajo v praksi.

Najprej moramo določiti NOD-e za preostale tipične dneve, kot so sobota, nedelja, praznični ter pred- in po-praznični dnevi. Pri prazničnih in ob-prazničnih dneh moramo biti pazljivi, saj je večina praznikov vezanih na datum, kar pomeni, da je praznik lahko na katerikoli dan v tednu. Za izdelavo NOD-ov bi lahko uporabili isti postopek, kakor za delavnike iz poglavja 3, vendar bi bila takšna rešitev nepraktična. Potrebovali bi dodatne podatke o odjemalcih, kakor so npr. obratovalni čas med vikendi ali prazniki, kar bi povzročalo dodatne težave pri dodeljevanju NOD-ov odjemalcem. Zgodi se, da del odjemalcev, ki imajo isti tip dejavnosti med vikendom obratujejo na enak način, kakor med delavniki, medtem ko ostali odjemalci med vikendom ne obratujejo, npr. trgovine. Težavo rešimo z opazovanjem porabe vseh odjemalcev, ki sestavljajo posamezno skupino ali NOD tudi za preostale tipične dneve in določimo njihove povprečne vrednosti, ki predstavljajo NOD-e za te tipične dni. Boljša rešitev je, če porabo merimo neprestano in določimo mesečne, sezonske in letne NOD-e.

V disertaciji smo NOD-e določili s povprečenjem vseh normiranih OD-jev, ki so se razvrstili v isto skupino. Iz metodološkega vidika je takšna odločitev upravičena, ker nas zanima samo

oblika NOD-ov in ne dejanska porabljena energija odjemalcev, ki jo merimo s števci. Tako smo OD-jem vseh odjemalcev pripisali enako "težo" v končni obliki NOD-a, kar po naše mnenju vnaša določeno napako pri uporabi NOD-ov in določanju rekonstruiranih OD-jev skupin odjemalcev. Jasno je, da imajo odjemalci z večjo močjo večji vpliv na končno obliko OD-jev, kar bi lahko upoštevali pri določanju oblike NOD-a.

Optimizacija rekonstruiranih OD-jev ne predstavlja večje težave. Povečati moramo število spremenljivk, ki predstavljajo pozitivne in negativne popravke moči, da opišemo porabo celotnega obračunskega obdobja, zapisati ustrezne enačbe, ki predstavljajo kriterijsko funkcijo ter omejitve in izvesti optimizacijo ter izračunati OD-je posameznih skupin odjemalcev. Na podlagi znanih OD-jev skupin odjemalcev določimo OD-je posameznih odjemalcev ali dobaviteljev glede na delež energije, ki jo porabijo. Ko poznamo dobaviteljeve OD-je, izvedemo obračun odstopanj dejanske porabe od napovedane v skladu s sprejetim pravilnikom [12].

V razdelku 2.3.1, kjer smo obravnavali preostalo krivuljo, smo celotni OD distribucijskega omrežja razdelili na tri obremenitvene krivulje ali diagrame: izgube, merjeni in nemerjeni odjemalci. Dodatna delitev nemerjenih odjemalcev je še na upravičene, v disertaciji smo se posvečali predvsem njim, ter neupravičene ali tarifne odjemalce. Do sedaj je dobava in obračun električne energije ter odstopanj za ti dve skupini potekal ločeno. S popolnim odprtjem trga z električno energijo, ko bodo tudi sedanji neupravičeni odjemalci prosto nastopali na trgu, bo obračun enoten. Zato moramo za to skupino odjemalcev tudi izdelati NOD-e. Poudariti moramo, da bi lahko NOD-e za te odjemalce uporabljali že sedaj, četudi trenutno obstaja različen obračun za tarifne odjemalce. Z NOD-i za te odjemalce bi si pomagali pri razdelitvi obremenitvene krivulje nemerjenih odjemalcev na del, ki pripada UO-jem in del, ki ga določa tarifni odjem.

Da bi v polni meri izkoristili prednosti nove metode določanja dinamike porabe odjemalcev brez sprotne meritve iz poglavja 4, moramo poznati NOD-e tudi tarifnih odjemalcev. Zato bomo v nadaljevanju, na podlagi izkušenj, ki smo jih pridobili z izdelavo NOD-ov UO-jev ter strokovne literature s tega področja, skušali podati osnovne smernice za izdelavo NOD-ov za tarifne odjemalce. Izdelava NOD-ov za to skupino odjemalcev se nekoliko razlikuje od izdelave NOD-ov za UO-je zaradi specifičnosti dinamike odjema. Zaradi manjše priključne moči in porabe ima ta skupina odjemalcev zelo razgibane OD-je. Njihovo splošno obliko zelo težko določimo na podlagi ene meritve, zato bi metode razvrščanja, ki smo jih v disertaciji uporabili za izdelavo NOD-ov UO-jev, težje neposredno uporabili za to skupino odjemalcev. Nasprotno, uporaba valčne multiresolucijske analize, namenjene glajenju OD-jev, bi bila dobrodošla. Na podlagi tega predlagamo sledeče smernice:

- (i) razdelitev odjemalcev v vnaprejšnje skupine. Pri tem bi si pomagali z obstoječimi tarifnimi razredi;

- (ii) pridobitev in analiza podatkov o letni porabi energiji vseh odjemalcev, razdeljenih na posamezne skupine. Takšna analiza bi bila smiselna tudi pri izdelavi NOD-ov za UO-je;
- (iii) določitev števila odjemalcev, katerim bi merili dinamiko odjema. Pri tem bi uporabili metode plastnega vzorčenja iz [15]-[18], kjer bi na podlagi standardnega odklona porabe električne energije določili število meritev dinamike odjema za vsako skupino odjemalcev posebej;
- (iv) merjenje porabe in izdelava NOD-ov ter ocenitev njihove natančnosti glede na porabljeno mesečno ali letno energijo ter dinamiko odjema.

NOD-e, izdelane na podlagi zgornjih točk, bi vključili v optimizacijski model iz razdelka 4.3 in določili dejanske OD-je vseh nemerjenih odjemalcev in ne samo UO-jev, kakor smo pokazali v disertaciji. Pri tem bi morali optimizacijski model nekoliko prilagoditi in razširiti, da bi upošteval pavšalni obračun električne energije gospodinjstev. V tem okviru bi lahko uporabili mehko linearno programiranje [51], [52].

## 6. Sklep

### 6.1. Rezultati dela

Odprtje trga z električno energijo je prineslo veliko novih vprašanj ter priložnosti udeležencem, ki nastopajo na njem. Uporaba obremenitvenih diagramov ni novo področje, saj so jih vrsto let uporabljali že v preteklosti. Nova je ideja uporabe odremenitvenih diagramov kot pripomočkov pri spodbujanju trgovanja z električno energijo, kar nas je vodilo k oblikovanju nove metode določanja dinamike porabe odjemalcev brez sprotnih meritev.

V disertaciji smo zasnovali in predstavili novo metodo določanja obremenitvenih diagramov odjemalcev brez sprotnih meritev, ki uporablja prednosti več obstoječih metod. Metoda upošteva analitično določeno preostalo krivuljo, ki predstavlja splošno in dokaj natančno obliko odjema vseh nemerjenih odjemalcev določenega distribucijskega omrežja, ter nadomestnih obremenitvenih diagramov (NOD-ov), ki podajajo dinamiko odjema posameznih skupin odjemalcev. Ker temelji na optimizaciji z uporabo linearnega programiranja, ki najde globalni optimum, je metoda hitra in enostavna. Predlagana metoda odpravlja slabost analitičnega postopka, ki predvideva enako obliko obremenitvenih diagramov vseh odjemalcev, kar se največkrat ne sklada z realnostjo, saj imajo odjemalci diagrame različnih oblik. Poleg tega metoda odpravlja morebitna dodatna tveganja distribucijskega operaterja za neravnotežja med porabo določeno s sintetičnim in analitičnim postopkom. V disertaciji smo prednosti nove metode utemeljevali na dejanskih meritvah OD-jev. Pri tem smo analizirali samo en delavnik, kjer smo predpostavili, da poznamo dnevno porabo električne energije odjemalcev. Kljub predstavitvi rezultatov samo na podlagi analize enega dne lahko predlagano metodo z dodajanjem spremenljivk relativno enostavno razširimo na daljše časovno obdobje, npr. en mesec, kar ustreza tudi sedanjemu obračunskemu obdobju.

Nova metoda določanja porabe odjemalcev brez sprotnih meritev temelji na NOD-ih, ki so izdelani na podlagi meritev diagramov posameznih odjemalcev. Obstoječe metode njihovega določanja temeljijo na dveh načinih. Prvi predvideva uporabo večletnega merjenja porabe odjemalcev, razdeljenih v znane skupine, medtem ko drugi način predlaga uporabo metod razpoznavanja vzorcev, vendar ne podaja postopka, kako NOD-e dodeliti odjemalcem. Disertacija predlaga kompromis med obema načinoma določanja NOD-ov, ki temelji na metodah razpoznavanja vzorcev, in jih dodeljuje odjemalcem na podlagi njihove registrirane dejavnosti. V tem okviru smo analizirali OD-je upravičenih odjemalcev in predlagali postopke, ki so se izkazali za učinkovite, kar smo pokazali s primerjavo rekonstruiranih diagramov posameznih skupin odjemalcev in njihovih izmerjenih vrednosti. Pred razvrščanjem smo merjene diagrame gladili z valčno multiresolucijsko analizo in normirali.

Na podlagi izsledkov, ki smo jih pridobili pri oblikovanju optimizacijskega modela določanja dinamike porabe odjemalcev brez sprotnih meritev pri izdelavi NOD-ov za upravičene odjemalce, in na podlagi strokovne literature s tega področja smo predlagali nadaljnje naloge, ki so potrebne za učinkovito uporabo zasnovane metode v praksi. Najpomembnejša in hkrati najzahtevnejša med njimi je izdelava NOD-ov za tarifne odjemalce.

## 6.2. Izvirni prispevki disertacije

Najpomembnejše izvirne prispevke doktorske disertacije pri raziskavi metod za ugotavljanje porabe električne energije odjemalcev brez sprotnih meritev lahko strnemo v naslednjih točkah:

- Določitev trenutno najustreznejše metode za razvrščanja obremenitvenih diagramov upravičenih odjemalcev;
- Preprost postopek dodeljevanja nadomestnih obremenitvenih diagramov upravičenim odjemalcem brez sprotnih meritev;
- Nova metoda za razdelitev obremenitve iz preostale krivulje obremenitev z uporabo nadomestnih obremenitvenih diagramov in postavitvev kriterijev, ki jih razdelitev mora upoštevati;
- Optimalna razdelitev preostale krivulje obremenitve z upoštevanjem standardnega odklona.



## 7. Literatura

- [1] EU, European Union Council Directive 96/92/EC, 1996 [http://europa.eu.int/eur-lex/en/lif/reg/en\\_register\\_1230.html](http://europa.eu.int/eur-lex/en/lif/reg/en_register_1230.html).
- [2] Energetski zakon (EZ), Klasifikacijska številka: 320-01/98-0007/0001, Objava v Ur. listu 79/99, 30/9/99 (stran 12378).
- [3] H.L. Willis, T.D. Vismor, R.W. Powell, "Some Aspects of Sampling Load Curve Data on Distribution Systems" IEEE Transactions on Power Systems, Vol. PAS-104, No. 11, November 1985.
- [4] T. G. Werner, "Load Profiling in Germany", TELMARK, 1st Discussion Forum, 2-4 September 2002, London, UK, razpoložljivo na: [http://www.telmark.org/2002Sep/2-2\\_Werner.pdf](http://www.telmark.org/2002Sep/2-2_Werner.pdf).
- [5] "Pravila za delovanje trga z električno energijo", Borzen, organizator trga z električno energijo, Ljubljana 2003.
- [6] D. Gerbec, S. Gašperič, Obremenitveni diagrami nizkonapetostnih odjemalcev električne energije. Peta konferenca slovenskih elektroenergetikov, Bled, 28. - 30.maja 2001. Zbornik referatov. Ljubljana: Društvo SLOKO CIGRE, 2001, del 2, str. 85-91.
- [7] "Metering, Load Profiles and Settlement in Deregulated Markets", System Tariff Issues Working Group, EURELECTRIC, ref: 2000-220-0004, March 2000, <http://public.eurelectric.org/Content/Default.asp>.
- [8] D. Gerbec, S. Gašperič, I. Šmon, F. Gubina, "Allocation of the Load Profiles to Consumers Using Probabilistic Neural Networks", članek sprejet v objavo v IEEE Transactions on Power Systems.
- [9] D. Gerbec, S. Gašperič, I. Šmon, F. Gubina, "Determining the load profiles of consumers based on fuzzy logic and probability neural networks", IEE proceedings Generation transmission and distribution, vol. 151, no. 3, pp. 395-400, 2004.
- [10] A. Capasso, W. Grattier, R. Lamedica, A. Prudenzi, "A Bottom-Up Approach to Residential Load Modeling", IEEE Transactions on Power Systems, vol. 9, No. 2, pp. 957-964, May 1994.
- [11] A. Seppälä, "Load research and load estimation in electricity distribution", doktorska disertacija, Technical research centre of Finland, VTT Publications 289, 1996.
- [12] "Pravilnik o načinu obračunavanja odstopanj oddaje ali odjema električne energije od vozniških redov", Uradni list Republike Slovenije, št. 83/27.9.2002/st. 9171-9177.
- [13] "Uredba o splošnih pogojih za dobavo in odjem električne energije", Uradni list 117/2

- 
- [14] C. F. Walker, J. L. Pokoski, "Residential Load Shape Modelling Based on Customer Behaviour", IEEE Transactions on Power Apparatus and Systems, vol. 104, no. 7, pp. 1703-1711, July 1985.
- [15] C. S. Chen, J. C. Hwang, Y. M. Tzeng, C. W. Huang, M. Y. Cho, Determination of Customer Load Characteristics by Load Survey System at Taipower", IEEE Transactions on Power Delivery, vol. 11, No. 3, p.p. 1430-1436, July 1996.
- [16] C. S. Chen, J. C. Hwang, and C. W. Huang, "Application of Load Survey Systems to Proper Tariff Design", IEEE Transactions on Power Systems, vol. 12, No. 4, pp. 1746-1751, Nov. 1997.
- [17] C. S. Chen, M.S. Kang, J. C. Hwang, and C. W. Huang, "Synthesis of power system load profiles by class load study", Electrical Power and Energy Systems, vol. 22, p.p. 325-330, 2000.
- [18] C. S. Chen, M.S. Kang, J. C. Hwang, and C. W. Huang, "Implementation of the Load Survey System in Taipower", IEEE Transmission and Distribution Conference, IEEE. vol. 1, pp.300-304, Piscataway, NJ, USA, 1999.
- [19] J. A. Jardini, C. M. V. Tahan, M. R. Gouvea, S. U. Ahn, F.M. Figueiredo, "Daily Load Profiles for Residential, Commercial and Industrial Low Voltage Consumers", IEEE Transactions on Power Delivery, vol. 15, No. 1, pp. 375-380, January 2000.
- [20] J. A. Jardini, H.P. Schmidt, C.M.V. Tahan, C.C.B. de Oliveira, S.U. Ahn, "Distribution Transformer Loss of Life Evaluation: A Novel Approach Based on Daily Load Profiles", IEEE Transactions on Power Delivery, Vol. 15, No. 1, pp.361-366, January 2000.
- [21] E. Bompard, E. Carpaneto, G. Chicco, R. Napoli, F. Piglionne, P. Postolache, M. Scutariu, "Stratified Sampling of the Electricity Customers for Setting up a Load Profile Surey", Proceedings of RIMAPS Conference, Madeira, Portugal, September 2000.
- [22] S. V. Allera, A. G. Horsburgh, "Load Profiling for Energy Trading and Settlements in the UK Electricity Markets", Proceedings DistribuTECH, 1998, <http://www.electricity.org.uk/services/lrg/dadsm2a2.html>.
- [23] M. Scutariu, C. Toader, P. Postolache, "Investigations on Daily Load Curve Characterisations", Proceedings of RIMAPS Conference, Porto, Portugal, September 2001.
- [24] G. Chicco, R. Napoli, P. Postolache, M. Scutariu, C. Toader, "Electric Energy Customer Characterisation for Developing Dedicated Market Strategies", in Proceedings of IEEE Porto Power Tech Conference, 2001.
- [25] G. Chicco, R. Napoli, P. Postolache, M. Scutariu, C. Toader, "Customer Characterization Options for Improving the Tariff Offer", IEEE Transactions on Power Systems, vol. 18, no. 1, pp. 381-387, February 2003.
-

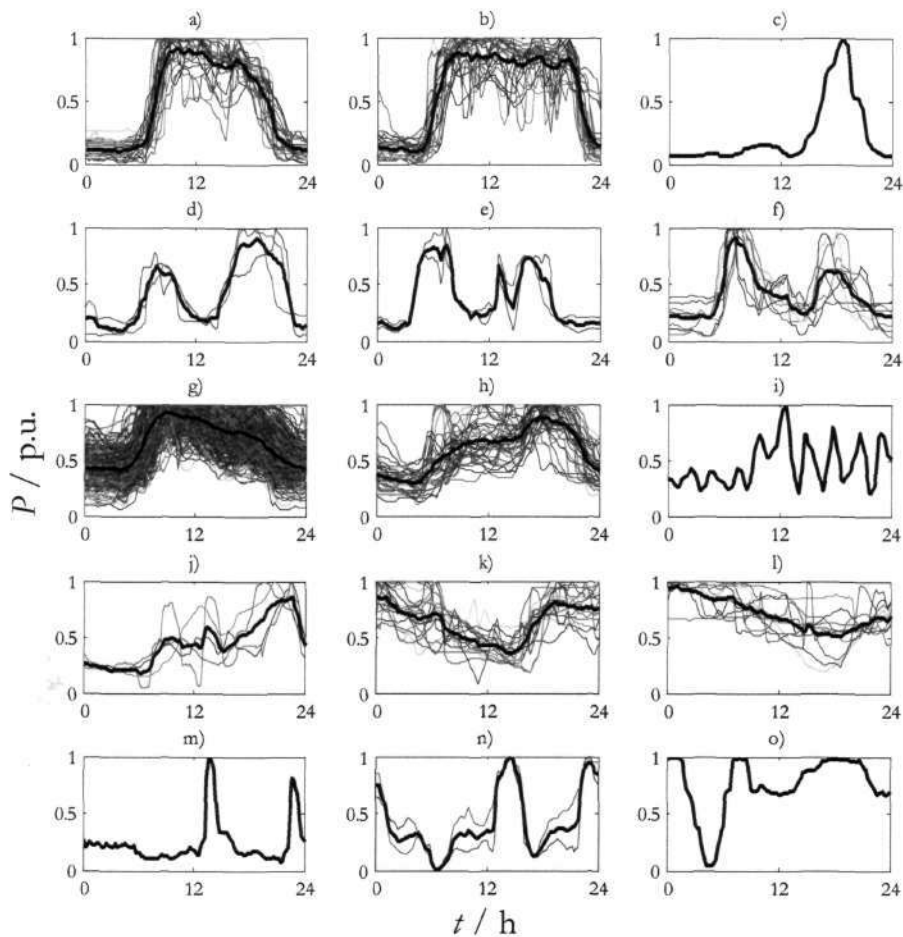
- [26] G. Chicco, R. Napoli, F. Piglione, P. Postolache, M. Scutariu, C. Toader, "Load Pattern Based Classification of Electricity Customers", *IEEE Transactions on Power Systems*, vol. 19, no. 2, pp. 31232-1239, May 2004.
- [27] R. Lamedica, L. Santolamazza, G. Fracassi, G. Martinelli, A. Prudenzi, "A novel methodology based on clustering techniques for automatic processing of MV feeder daily load profiles", *Power Engineering Society Summer Meeting*, vol. 1, pp. 96-101, 16-20 July, 2000.
- [28] F. Gubina, D. Gerbec, S. Gašperič, Krmiljenje napetostnega nivoja v distribucijskih omrežjih – Obremenitveni diagrami odjemalcev, Fakulteta za elektrotehniko v Ljubljani, maj 2001.
- [29] I. Šmon, "Oblikovanje nadomestnih obremenitvenih diagramov", diplomsko delo, Fakulteta za elektrotehniko, 2001.
- [30] D. Gerbec, S. Gašperič, I. Šmon, F. Gubina, "Hierarchic Clustering Method for Consumers Load profile Determination", *Proceedings of the 2<sup>nd</sup> Balkan Power Conference*, Belgrade, Yugoslavia, pp. 1-15, June 2002.
- [31] M. Misiti, Y. Misiti, G. Oppenheim, J.-M. Poggi, "Wavelet Toolbox for Use With Matlab", *User's Guide, Version 2*, Mathworks, 2002.
- [32] U. Lotrič, Uporaba valčne analize in nevronske mreže pri napovedovanju časovnih vrst, doktorska disertacija, Fakulteta za elektrotehniko v Ljubljani, april 2000.
- [33] J. C. Bezdek, "Pattern Recognition With Fuzzy Objective Function Algorithms", New York Plenum, New York, 1981.
- [34] R. N. Dave, "Characterization and detection of noise in clustering", *Pattern Recognition Letters*, Vol. 12, pp. 657-664, November 1991.
- [35] R. N. Dave, R. Krishnapuram, "Robust Clustering Methods: A Unified View", *IEEE Transactions on Fuzzy Systems*, Vol. 5, No. 2, pp. 270-293, May 1997.
- [36] Nikola Pavešić, "Razpoznavanje vzorcev", ZAFER, Ljubljana 1992.
- [37] A. K. Jain, P. W. Duin, J. Mao, "Statistical Pattern Recognition: A Review", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 2, No. 1, January 2000.
- [38] M. Y. Kiang, "Extending the Kohonen self-organizing map networks for clustering analysis", *Computational Statistics & Data Analysis*, 38, pp. 161-180, 2001.
- [39] R. A. Johnson, D. W. Wicheren, "Applied Multivariate Statistical Analysis", Fifth Edition, Practice Hall, New Jersey, 2002.
- [40] *Statistics Toolbox for Use with Matlab, User's Guide Version 4*, July 2002.
- [41] D. L. Davies, D. W. Bouldin, "A Cluster Separation Measure", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-1, No. 2, April 1979.
- [42] H. Demuth, M. Beale, "Neural Network Toolbox for Use with MATLAB", *Users Guide, version 4*, MathWorks, 2002.

- 
- [43] C. H. Chen, "Fuzzy Logic and Neural Networks Handbook", IEEE Press, McGraw-Hill Companies, pp. 3.1-3.13, 1996.
- [44] M. K. Celik, A. Abur, "Use of Scaling in WLAV Estimation of Power System States", IEEE Transactions on Power Systems, Vol. 7, No. 2, pp. 684-692, May 1992.
- [45] H. Snigh, F. L. Alvarado, "Weighted Least Absolute Value State Estimation Using Interior Point Methods", IEEE Transactions on Power Systems, Vol. 9, No. 3, pp. 1478-1484, August 1994.
- [46] M. R. Irving, R.C. Owen, M. J. H. Sterling, "Power-system state estimation using linear programming", IEE Proc. vol. 125, No. 9, pp. 879-885, September 1978.
- [47] D. M. Falcao, H. O. Henriques, "Load Estimation in Radial Distribution Systems Using Neural Networks and Fuzzy Set Techniques", Power Engineering Society Summer Meeting, 2001. IEEE, Vol. 2 , pp. 1002-1006, 15-19 July 2001.
- [48] S.S. Rao, "Engineering Optimization, 3rd edition", West Lafayette, Indiana, John Wiley & Sons, 1995.
- [49] I. Vidav, "Višja matematika 2", Državna založba Slovenije, Ljubljana 1975.
- [50] Optimization Toolbox for Use with Matlab, user's guide, version 2, The MathWorks.
- [51] H. J. Zimmermann, "Fuzzy Sets Decision Making and Expert Systems", Kluwer Academic Publishers, Boston/Dordrecht/London, 1986.
- [52] H. J. Zimmermann, "Fuzzy Set Theory – and Its Applications", Kluwer Academic Publishers, Boston/Dordrecht/London, 2001.

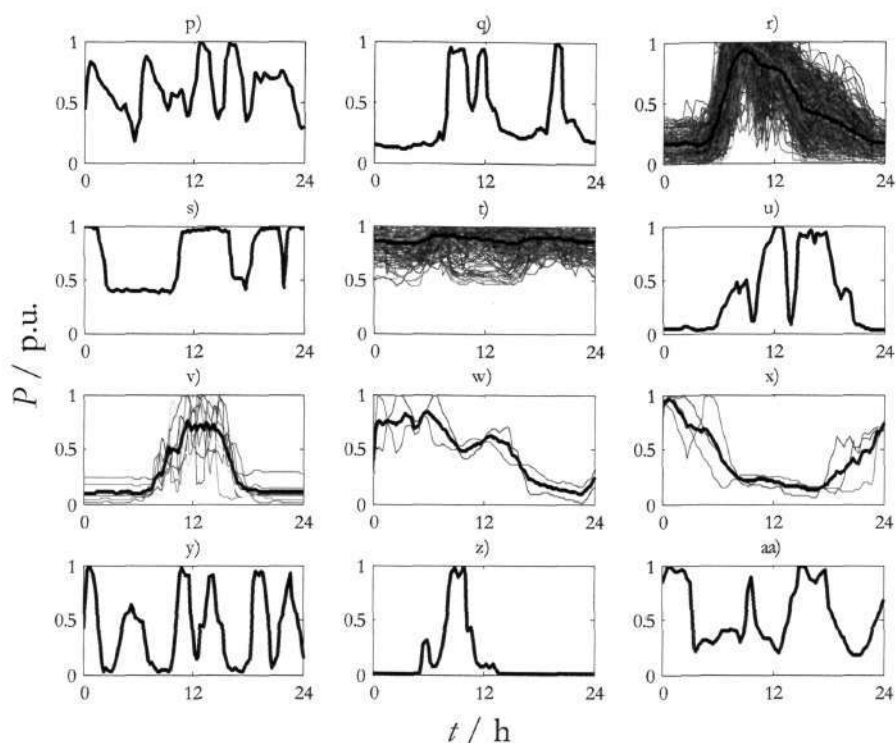
# Priloge

## Priloga A

Rezultat razvrščanja OD-jev v 27 skupin pri izračunu povprečne razdalje med skupinami.



Sl. 1: Rezultati razvrščanja prvih petnajst skupin.



Sl. 2: Rezultati razvrščanja preostalih dvanajst slupin.

## Priloga B

Tab. 1: Opis dejavnosti iz Standardnega klasifikata dejavnosti.

Šifra dejavnosti	Opis dejavnosti
01	KMETIJSTVO IN LOV TER Z NJIMA POVEZANE STORITVE
02	GOZDARSTVO IN GOZDARSKE STORITVE
10	PRIDOBIVANJE ČRNEGA PREMOGA, RJAVEGA PREMOGA IN LIGNITA, ŠOTE
13	PRIDOBIVANJE RUD
15	PROIZVODNJA HRANE, PIJAČ IN KRMIL
17	PROIZVODNJA TEKSTILIJ
18	PROIZVODNJA OBLAČIL; STROJENJE IN DODELAVA KRZNA; PROIZVODNJA KRZNENIH IZDELKOV
19	PROIZVODNJA USNJA, OBUTVE IN USNJENIH IZDELKOV, RAZEN OBLAČIL
20	OBDELAVA IN PREDELAVA LESA; PROIZVODNJA IZDELKOV IZ LESA, PLUTE, SLAME IN PROTJA, RAZEN POHIŠTVA
21	PROIZVODNJA VLAKNIN, PAPIRJA IN KARTONA TER IZDELKOV IZ PAPIRJA IN KARTONA
22	ZALOŽNIŠTVO, TISKARSTVO; RAZMNOŽEVANJE POSNETIH NOSILCEV ZAPISA
24	PROIZVODNJA KEMIKALIJ, KEMIČNIH IZDELKOV, UMETNIH VLAKEN
25	PROIZVODNJA IZDELKOV IZ GUME IN PLASTIČNIH MAS
26	PROIZVODNJA DRUGIH NEKOVINSKIH MINERALNIH IZDELKOV



27	PROIZVODNJA KOVIN
28	PROIZVODNJA KOVINSKIH IZDELKOV, RAZEN STROJEV IN NAPRAV
29	PROIZVODNJA STROJEV IN NAPRAV
30	PROIZVODNJA PISARNIŠKIH STROJEV IN RAČUNALNIKOV
31	PROIZVODNJA ELEKTRIČNIH STROJEV IN NAPRAV
32	PROIZVODNJA RADIJSKIH, TELEVIZIJSKIH IN KOMUNIKACIJSKIH NAPRAV IN OPREME
33	PROIZVODNJA MEDICINSKIH, FINOMEHANIČNIH IN OPTIČNIH INSTRUMENTOV TER UR
34	PROIZVODNJA MOTORNIH VOZIL, PRIKOLIC IN POLPRIKOLIC
35	PROIZVODNJA DRUGIH VOZIL IN PLOVIL
36	PROIZVODNJA POHIŠTVA IN DRUGE PREDELOVALNE DEJAVNOSTI
37	RECIKLAŽA
40	OSKRBA Z ELEKTRIČNO ENERGIJO, PLINOM, PARO IN TOPLO VODO
41	ZBIRANJE, ČIŠČENJE IN DISTRIBUCIJA VODE
45	GRADBENIŠTVO
50	PRODAJA, VZDRŽEVANJE IN POPRAVILA MOTORNIH VOZIL; TRGOVINA NA DROBNO Z MOTORNIMI GORIVI
51	POSREDNIŠTVO IN TRGOVINA NA DEBELO, RAZEN Z MOTORNIMI VOZILI
52	TRGOVINA NA DROBNO, RAZEN Z MOTORNIMI VOZILI; POPRAVILA IZDELKOV ŠIROKE PORABE
55	GOSTINSTVO
60	KOPENSKI PROMET; CEVOVODNI TRANSPORT
63	POMOŽNE PROMETNE DEJAVNOSTI; DEJAVNOST POTOVALNIH IN TURISTIČNIH ORGANIZACIJ
64	POŠTA IN TELEKOMUNIKACIJE
65	FINANČNO POSREDNIŠTVO, RAZEN ZAVAROVALNIŠTVA IN DEJAVNOSTI POKOJNINSKIH SKLADOV
66	ZAVAROVALNIŠTVO IN DEJAVNOST POKOJNINSKIH SKLADOV, RAZEN OBVEZNEGA SOCIALNEGA ZAVAROVANJA
70	POSLOVANJE Z NEPREMIČNINAMI
71	DAJANJE STROJEV IN OPREME BREZ UPRAVLJAVCEV V NAJEM; IZPOSOJANJE IZDELKOV ŠIROKE PORABE
72	OBDELAVA PODATKOV, PODATKOVNE BAZE IN S TEM POVEZANE DEJAVNOSTI
74	DRUGE POSLOVNE DEJAVNOSTI
75	DEJAVNOST JAVNE UPRAVE IN OBRAMBE, OBVEZNO SOCIALNO ZAVAROVANJE
80	IZOBRAŽEVANJE
85	ZDRAVSTVO IN SOCIALNO VARSTVO
90	DEJAVNOSTI JAVNE HIGIENE
91	DEJAVNOSTI ZDRUŽENJ, ORGANIZACIJ
92	REKREACIJSKE, KULTURNE IN ŠPORTNE DEJAVNOSTI
93	DRUGE STORITVENE DEJAVNOSTI
99	EKSTERITORIALNE ORGANIZACIJE IN ZDRUŽENJA

## IZJAVA

Izjavljam, da je doktorska disertacija z naslovom "Metode za ugotavljanje porabe električne energije odjemalcev brez sprotnih merite" izključno rezultat mojega lastnega raziskovalnega dela pod vodstvom mentorja prof. dr. Ferdinanda Gubine. Izkazano pomoč drugih sodelavcev sem v celoti navedel v zahvali.



David Gerbec