



УНИВЕРЗИТЕТ У НОВОМ САДУ  
ФАКУЛТЕТ ТЕХНИЧКИХ НАУКА У  
НОВОМ САДУ



---

**АДАПТИВНИ МОДЕЛ ПРОГНОЗЕ  
ВРЕМЕНСКИХ СЕРИЈА У  
ПАМЕТНИМ  
ИНФРАСТРУКТУРНИМ  
МРЕЖАМА**

ДОКТОРСКА ДИСЕРТАЦИЈА

Ментор:  
др Александар Селаков

Кандидат:  
мр Зоран Јанковић

Нови Сад, 2022. године

КЉУЧНА ДОКУМЕНТАЦИЈСКА ИНФОРМАЦИЈА<sup>1</sup>

Врста рада:	Докторска дисертација
Име и презиме аутора:	Зоран Јанковић
Ментор (титула, име, презиме, звање, институција)	др Александар Селаков, доцент, Универзитет у Новом Саду, Факултет техничких наука
Наслов рада:	Адаптивни модел прогнозе временских серија у паметним инфраструктурним мрежама
Језик публикације (писмо):	Српски (ћирилица)
Физички опис рада:	Унети број: Страница: 154 Поглавља: 9 Референци: 126 Табела: 41 Слика: 63 Графикона: 25 Прилога: 1
Научна област:	Електротехничко и рачунарско инжењерство
Ужа научна област (научна дисциплина):	Примењено софтверско инжењерство
Кључне речи / предметна одредница:	Прогноза временских серија / машинско учење / метахеуристичка оптимизација

---

<sup>1</sup> Аутор докторске дисертације потписао је и приложио следеће Обрасце:

5б – Изјава о ауторству;

5в – Изјава о истоветности штампане и електронске верзије и о личним подацима;

5г – Изјава о коришћењу.

Ове Изјаве се чувају на факултету у штампаном и електронском облику и не кориче се са тезом.

Резиме на језику рада:	<p>Дисертација се бави креирањем напредног адаптивног модела краткорочне прогнозе вредности временских серија у паметним инфраструктурним мрежама. Појаве које модел прогнозира односе се пре свега на потрошњу ресурса као што су електрична енергија, природни гас и вода. Модел користи постојеће методе машинског учења, које унапређује како у погледу тачности прогнозе тако и у погледу перформанси. Примарни допринос дисертације је унапређење постојећих техника машинског учења коришћењем предложеног модела.</p> <p>Основа модела је оптимизација избора узорка који користе методе машинског учења у поступку прогнозе. Узорак се односи на дане из историјске базе података. Узорак је репрезентативнији што су дани из узорка сличнији прогнозираном дану. Сличност дана се одређује на основу релевантних фактора сличности. Модел је адаптиван у смислу могућности прилагођавања утицаја и додавања нових фактора сличности. Утицај појединачних фактора сличности оптимизује се метахеуристички, коришћењем генетског алгоритма. У извршавању генетског алгоритма у адаптивном моделу прогнозе уочени су проблеми, који су описани а затим и решени увођењем иновативних техника који унапређују конвергенцију, а последично и тачност коначних резултата прогнозе.</p> <p>Модел поседује могућност континуалног унапређења кроз прилагођавање промењеним календарским и временским околностима. Ова адаптивност омогућена је кроз механизам извршавања генетског алгоритма на дневном и кварталном нивоу, чиме се краткорочно и дугорочно оптимизује утицај фактора сличности.</p> <p>Модел је такође прилагођен специјалним данима – празницима и другим данима чији профили потрошње одступају у односу на стандардне дане. Предложеним решењем у моделу специјални дани се препознају и за њих се извршава прогноза коришћењем посебног алгоритма којим се препознају и примењују специфична правила понашања потрошње.</p> <p>Верификација модела извршена је на две студије случаја. Једна студија случаја односи се на краткорочну прогнозу потрошње електричне енергије у преносној мрежи електроенергетског система за подручје Србије. Друга студија случаја односи се на краткорочну прогнозу потрошње природног гаса у дистрибутивној мрежи за један регион у Холандији.</p>
Датум прихватања теме од стране надлежног већа:	27.10.2021. године
Датум одбране: (Попуњава одговарајућа служба)	
Чланови комисије: (титула, име, презиме, звање, институција)	<p>Председник: др Александар Ердељан, редовни професор, Факултет техничких наука, Нови Сад</p> <p>Члан: др Имре Лендак, ванредни професор, Факултет техничких наука, Нови Сад</p> <p>Члан: др Душко Бекут, редовни професор, Факултет техничких наука, Нови Сад</p> <p>Члан: др Стеван Цветићанин, доцент, Факултет техничких наука, Нови Сад</p> <p>Члан: др Александар Савић, ванредни професор, Електротехнички факултет, Београд</p> <p>Члан, ментор: др Александар Селаков, доцент, Факултет техничких наука, Нови Сад</p>
Напомена:	

**KEY WORD DOCUMENTATION<sup>2</sup>**

Document type:	Doctoral dissertation
Author:	Zoran Janković
Supervisor (title, first name, last name, position, institution)	PhD Aleksandar Selakov, Assistant Professor, Faculty of technical sciences, Novi Sad
Thesis title:	Adaptive time series forecasting model in intelligent infrastructure networks
Language of text (script):	Serbian language (cyrillic)
Physical description:	Number of: Pages: 154 Chapters: 9 References: 126 Tables: 41 Illustrations: 63 Graphs: 25 Appendices: 1
Scientific field:	Electrical and computing engineering
Scientific subfield (scientific discipline):	Applied software engineering
Subject, Key words:	Time series forecasting / machine learning / metaheuristic optimization

---

<sup>2</sup> The author of doctoral dissertation has signed the following Statements:

56 – Statement on the authority,

5B – Statement that the printed and e-version of doctoral dissertation are identical and about personal data,

5r – Statement on copyright licenses.

The paper and e-versions of Statements are held at the faculty and are not included into the printed thesis.

<p>Abstract in English language:</p>	<p>A topic of this dissertation is creating an advanced adaptive model of a short-term forecast of time series values in smart infrastructure grids. The primal phenomena that the model predicts are the consumption of resources such as electricity, natural gas, and water. The model uses existing machine learning methods, improved both in terms of forecast accuracy and in terms of performance. The primary contribution of the dissertation is the improvement of existing machine learning techniques using the proposed model.</p> <p>The basis of the model is the optimization of the selection of samples used by machine learning methods in the forecasting process. The sample refers to days from a historical database. The sample optimization is improved by the increase of the similarity of days from the sample. The similarity of the day is determined by relevant similarity factors. The model is adaptive in terms of adjusting the impact of similarity factors and adding new similarity factors.</p> <p>The influence of individual similarity factors is optimized by the metaheuristics, using a genetic algorithm. The execution of the genetic algorithm in the adaptive model of the forecast deals with problems noticed in this work. The detected issues are described and then solved by introducing innovative techniques that improve convergence, and consequently, the accuracy of the final results of the forecast.</p> <p>The model has the possibility of continuous improvement through adaptation to changed calendar and weather circumstances. This adaptability is enabled through the mechanism of execution of the genetic algorithm on a daily and quarterly level, which optimizes the influence of similarity factors in the short and long term.</p> <p>The model is also adapted to special days - holidays and other consumption profiles that deviate from standard days. The proposed solution recognizes special days and executes a specific forecast by a special algorithm that identifies and applies specific rules of consumption behavior on special days.</p> <p>Model verification was performed on two case studies. The first one refers to the short-term load forecast in the transmission system utility in Serbia. The second case study concerns a short-term forecast of natural gas consumption in the distribution network for a region in the Netherlands.</p>
<p>Accepted on Scientific Board on:</p>	<p>27/10/2021</p>
<p>Defended: (Filled by the faculty service)</p>	
<p>Thesis Defend Board: (title, first name, last name, position, institution)</p>	<p>President: PhD Aleksandar Erdeljan, Professor, Faculty of technical sciences, Novi Sad  Member: PhD Imre Lendak, Associate Professor, Faculty of technical sciences, Novi Sad  Member: PhD Duško Bekut, Professor, Faculty of technical sciences, Novi Sad  Member: PhD Stevan Cvetičanin, Assistant Professor, Faculty of technical sciences, Novi Sad  Member: PhD Aleksandar Savić, Associate Professor, School of Electrical Engineering, Belgrade  Member, Supervisor: PhD Aleksandar Selakov, Assistant Professor, Faculty of technical sciences, Novi Sad</p>
<p>Note:</p>	

# Резиме

Дисертација се бави креирањем напредног адаптивног модела краткорочне прогнозе вредности временских серија у паметним инфраструктурним мрежама. Појаве које модел прогнозира односе се пре свега на потрошњу ресурса као што су електрична енергија, природни гас и вода. Модел користи постојеће методе машинског учења, које унапређује како у погледу тачности прогнозе тако и у погледу перформанси. Примарни допринос дисертације је унапређење постојећих техника машинског учења коришћењем предложеног модела.

Основа модела је оптимизација избора узорка који користе методе машинског учења у поступку прогнозе. Узорак се односи на дане из историјске базе података. Узорак је репрезентативнији што су дани из узорка сличнији прогнозираном дану. Сличност дана се одређује на основу релевантних фактора сличности. Модел је адаптиван у смислу могућности прилагођавања утицаја и додавања нових фактора сличности.

Утицај појединачних фактора сличности оптимизује се метахеуристички, коришћењем генетског алгорита. У извршавању генетског алгорита у адаптивном моделу прогнозе уочени су проблеми, који су описани а затим и решени увођењем иновативних техника који унапређују конвергенцију, а последично и тачност коначних резултата прогнозе.

Модел поседује могућност континуалног унапређења кроз прилагођавање промењеним календарским и временским околностима. Ова адаптивност омогућена је кроз механизам извршавања генетског алгорита на дневном и кварталном нивоу, чиме се краткорочно и дугорочно оптимизује утицај фактора сличности.

Модел је такође прилагођен специјалним данима – празницима и другим данима чији профили потрошње одступају у односу на стандардне дане. Предложеним решењем у моделу специјални дани се препознају и за њих се извршава прогноза коришћењем посебног алгорита којим се препознају и примењују специфична правила понашања потрошње.

Верификација модела извршена је на две студије случаја. Једна студија случаја односи се на краткорочну прогнозу потрошње електричне енергије у преносној мрежи електроенергетског система за подручје Србије. Друга студија случаја односи се на краткорочну прогнозу потрошње природног гаса у дистрибутивној мрежи за један регион у Холандији.

# Summary

A topic of this dissertation is creating an advanced adaptive model of a short-term forecast of time series values in smart infrastructure grids. The primal phenomena that the model predicts are the consumption of resources such as electricity, natural gas, and water. The model uses existing machine learning methods, improved both in terms of forecast accuracy and in terms of performance.

The primary contribution of the dissertation is the improvement of existing machine learning techniques. In this respect, an innovative model for developing software systems in intelligent infrastructure networks for predicting and forecasting phenomena that manifest in time series was introduced.

The basis of the model is the optimization of the selection of samples used by machine learning methods in the forecasting process. The sample refers to days from a historical database. The sample optimization is improved by the increase of the similarity of days from the sample. The similarity of the day is determined by relevant similarity factors. The model is adaptive in terms of adjusting the impact of similarity factors and adding new similarity factors.

The influence of individual similarity factors is optimized by the metaheuristics, using a genetic algorithm. The execution of the genetic algorithm in the adaptive model of the forecast deals with problems noticed in this work. The detected issues are described and then solved by introducing innovative techniques that improve convergence, and consequently, the accuracy of the final results of the forecast.

The model has the possibility of continuous improvement through adaptation to changed calendar and weather circumstances. This adaptability is enabled through the mechanism of execution of the genetic algorithm on a daily and quarterly level, which optimizes the influence of similarity factors in the short and long term.

The model is also adapted to special days – holidays and other consumption profiles that deviate from standard days. The proposed solution recognizes special days and executes a specific forecast by a special algorithm that identifies and applies specific rules of consumption behavior on special days.

Model verification was performed on two case studies. The first one refers to the short-term load forecast in the transmission system utility in Serbia. The second case study concerns a short-term forecast of natural gas consumption in the distribution network for a region in the Netherlands.

# Садржај

1.	Увод.....	1
2.	Концепт прогнозе временских серија у паметним инфраструктурним мрежама ....	5
2.1.	Прогноза потражње за ресурсима .....	5
2.1.1.	Краткорочна прогноза потрошње електричне енергије.....	7
2.1.1.1.	Методе краткорочног прогнозирања потрошње електричне енергије	7
2.1.1.2.	Фактори који утичу на потрошњу електричне енергије.....	8
2.1.2.	Краткорочна прогноза потражње за водом .....	9
2.1.2.1.	Методе краткорочног прогнозирања потражње за водом .....	9
2.1.2.2.	Фактори који утичу на потражњу за водом .....	9
2.1.3.	Краткорочна прогноза потрошње природног гаса .....	10
2.1.3.1.	Фактори који утичу на потражњу за природним гасом.....	10
2.1.3.2.	Методе краткорочног прогнозирања потрошње природног гаса .....	11
2.2.	Прогноза временских серија.....	11
2.2.1.	Машинско учење у прогнозирању потражње за ресурсима .....	11
2.2.2.	Методе машинског учења .....	12
2.2.3.	Линеарне методе машинског учења .....	12
2.2.3.1.	Мултиваријабилна линеарна регресија .....	12
2.2.3.2.	<i>ARMA</i> и <i>ARIMA</i> модели.....	13
2.2.4.	Нелинеарне методе машинског учења.....	15
2.2.4.1.	Неуронске мреже .....	15
2.2.4.2.	Метод потпорних вектора .....	16
2.2.4.3.	Расплинуте временске серије .....	17
2.2.5.	Избор узорка за методе машинског учења .....	18
3.	Унапређење метода прогнозе временских серија адаптивним моделом .....	19
4.	Преглед поступака прогнозе временских серија .....	23
4.1.	Избор узорка у прогнози потрошње ресурса .....	23



4.1.1.	Избор узорка у краткорочној потражњи потрошње електричне енергије	23
4.1.1.1.	Модел са једноставном селекцијом сличних дана	24
4.1.1.2.	Модел са комплексном селекцијом сличних дана	25
4.1.2.	Избор узорка у прогнози потражње за водом	27
4.1.3.	Избор узорка у прогнози потрошње природног гаса	28
4.2.	Метахеуристички методи	28
4.2.1.	Мрављи алгоритам	28
4.2.2.	Оптимизација ројева честица	29
4.2.3.	Генетски алгоритам	30
4.2.4.	Метахеуристички методи у прогнозирању краткорочне прогнозе потрошње електричне енергије	31
5.	Архитектура адаптивног модела прогнозе	32
5.1.	Филтер сличних дана	33
5.1.1.	Иницијални филтер	33
5.1.2.	Филтер сличности	34
5.1.2.1.	Коефицијент временских прилика	35
5.1.2.1.1.	Одређивање партиципационих коефицијената	35
5.1.2.2.	Инерциони коефицијент	36
5.1.2.2.1.	Коефицијент инерције потрошње	37
5.1.2.2.2.	Коефицијент термалне инерције	38
5.1.2.2.3.	Коефицијент инерције последњег сата	38
5.1.2.2.4.	Резултујући инерциони коефицијент	39
5.1.2.3.	Коефицијент трајања дневне светлости	39
5.1.2.4.	Коефицијент близине дана	39
5.1.2.5.	Резултујући коефицијент сличности	40
5.1.3.	Филтер девијације потрошње	40
5.2.	Метахеуристичка оптимизација	43
5.2.1.	Избор метахеуристичке методе	45
5.2.1.1.	Подешавање оптимизације ројева честица	46
5.2.1.2.	Подешавање стандардног генетског алгоритма	46

5.2.1.3.	Подешавање унапређеног генетског алгоритма.....	47
5.2.1.4.	Поређење резултата функције трошка .....	47
5.3.	Филтер специјалних дана.....	48
5.4.	Алгоритам прогнозе .....	51
5.5.	Финална архитектура адаптивног модела прогнозе временских серија .....	52
5.5.1.	Перформансе модела .....	53
6.	Подешавање, унапређење и примена генетског алгоритма .....	54
6.1.	Подешавање генетског алгоритма .....	54
6.2.	Унапређење генетског алгоритма .....	56
6.2.1.	Концепт насељавања .....	57
6.2.2.	Концепт мутације засноване на диверзитету .....	60
6.2.2.1.	Диверзитет генетских варијабли.....	60
6.2.2.2.	Диверзитет популације .....	60
6.2.3.	Мутација ка екстремима .....	62
6.3.	Тестирање унапређеног генетског алгоритма.....	63
6.4.	Примена генетског алгоритма у адаптивном моделу прогнозе .....	66
7.	Софтверска реализација модела .....	71
7.1.	Приказ функционалности .....	71
7.1.1.	Учитавање података из датотека.....	71
7.1.2.	Извршавање прорачуна .....	72
7.1.3.	Преглед прорачуна .....	72
7.1.4.	Преглед историјских података .....	73
7.1.5.	Преглед статистичких података .....	73
7.1.6.	Подешавање специјалних дана.....	74
7.1.7.	Поставка тежинских коефицијената .....	74
7.1.8.	Подешавање апликације .....	75
7.1.9.	Дневна и квартална оптимизација .....	75
7.2.	Архитектура апликације .....	76
7.2.1.	Компоненте система.....	77
7.3.	Модел података.....	79

7.4.	Апликативни модели.....	80
8.	Студије случаја .....	85
8.1.	Прогноза потрошње електричне енергије у преносној мрежи.....	85
8.1.1.	Прогноза за типичан дан.....	85
8.1.1.1.	Оптимизација генетским алгоритмом .....	85
8.1.1.1.1.	Дневна оптимизација током времена.....	87
8.1.1.2.	Селекција фактора од утицаја и партиципационих коефицијената ...	88
8.1.1.3.	Релевантни параметри прогнозираног дана.....	94
8.1.1.4.	Селекција података иницијалним филтером .....	96
8.1.1.5.	Пример прорачуна коефицијента сличности.....	98
8.1.1.5.1.	Прорачун коефицијента временских прилика .....	98
8.1.1.5.2.	Прорачун инерционог коефицијента .....	100
8.1.1.5.3.	Прорачун коефицијента трајања дневне светлости.....	101
8.1.1.5.4.	Прорачун коефицијента близине дана .....	101
8.1.1.5.5.	Прорачун резултујућег коефицијента сличности .....	101
8.1.1.6.	Филтер девијације потрошње.....	102
8.1.1.7.	Прорачун прогнозе потрошње електричне енергије неуронском мрежом	106
8.1.1.8.	Резултати прогнозе за изабрани дан.....	108
8.1.2.	Прогноза за специјални дан.....	110
8.1.3.	Анализа резултата прогнозе .....	116
8.1.4.	Поређење са резултатима прогнозе коришћењем саме неуронске мреже	116
8.1.5.	Поређење са другим техникама .....	121
8.1.5.1.	Упоредне прогнозе над истим сетом података.....	121
8.1.5.2.	Поређење са резултатима других техника краткорочне прогнозе потрошње електричне енергије.....	122
8.2.	Прогноза потрошње природног гаса у дистрибутивној мрежи .....	124
8.2.1.	Расположиви подаци и коришћени параметри у поступку прогнозе.....	125
8.2.2.	Оптимизација генетским алгоритмом .....	125
8.2.3.	Селекција фактора од утицаја и партиципационих коефицијената.....	126

8.2.4.	Релевантни параметри прогнозираног дана.....	129
8.2.5.	Селекција података иницијалним филтером.....	130
8.2.6.	Прорачун прогнозе потрошње природног гаса неуронском мрежом.....	133
8.2.7.	Резултати прогнозе за избрани дан .....	133
8.2.8.	Резултати прогнозе за један месец.....	135
8.2.9.	Поређење са резултатима прогнозе потрошње природног гаса коришћењем саме неуронске мреже .....	136
8.2.10.	Поређење са резултатима других техника краткорочне прогнозе потрошње природног гаса .....	139
9.	Закључак и правци даљег истраживања .....	140
9.1.	Правци даљег истраживања.....	142
	Литература.....	144

# Списак слика

Слика 1 – Типичан процес прогнозе потражње за ресурсима.....	6
Слика 2 – Типичан STLF процес.....	7
Слика 3 – Структура Feedforward неуронске мреже.....	16
Слика 4 – Модел технике модела потпорних вектора.....	17
Слика 5 – Глобална архитектура адаптивног модела прогнозе.....	32
Слика 6 – Прогнозирани и историјски низ дана.....	37
Слика 7 – Пример у коме профил потрошње једног сличног дана значајно одступа од осталих.....	41
Слика 8 – Алгоритам филтера девијације потрошње.....	42
Слика 9 – Једна итерација у метахеуристичкој оптимизацији.....	44
Слика 10 – Резултати функције трошка добијени тестираним оптимизационим методама.....	48
Слика 11 – Иницијалне поставке у филтеру специјалних дана.....	49
Слика 12 – Поступак прогнозе за специјалне дане.....	50
Слика 13 – Коначна архитектура адаптивног модела прогнозе временских серија.....	53
Слика 14 – Ток генетског алгоритма.....	55
Слика 15 – Преостали део популације у генетском алгоритму.....	59
Слика 16 – Пример конвергенције и стопе конвергенције у генетском алгоритму.....	61
Слика 17 – Мутација ка екстремима – кораци у генетском алгоритму.....	63
Слика 18 – Облици кривих за појединачне делове оптимизационог проблема.....	64
Слика 19 – Периоди метахеуристичке оптимизације.....	67
Слика 20 – Типови оптимизације параметара у адаптивном моделу прогнозе.....	68
Слика 21 – Усаглашавање дневних и кварталних параметара.....	70
Слика 22 – Дијаграм случајева коришћења.....	71
Слика 23 – Активност учитавања података из датотека.....	72
Слика 24 – Активност извршавања прорачуна.....	72

Слика 25 – Активност прегледа резултата прорачуна .....	73
Слика 26 – Активност прегледа историјских података .....	73
Слика 27 – Активност прегледа статистичких података .....	74
Слика 28 – Подешавање специјалних дана .....	74
Слика 29 – Активност поставке тежинских коефицијената .....	75
Слика 30 – Активност подешавања апликације .....	75
Слика 31 – Активност дневне/кварталне оптимизације .....	76
Слика 32 – Архитектура апликације.....	77
Слика 33 – Дијаграм компоненти сервисног дела апликативног система.....	78
Слика 34 – Дијаграм компоненти сервиса оптимизације генетским алгоритмом .....	78
Слика 35 – Дијаграм компоненти клијентске апликације .....	79
Слика 36 – Модел података апликације .....	80
Слика 37 – Дијаграм секвенци за опис припреме података .....	81
Слика 38 – Дијаграм секвенци за резидентно извршавање оптимизације.....	83
Слика 39 – Функционални дијаграм класа генетског алгоритма.....	84
Слика 40 – Кретање вредности дневних параметара током времена (a).....	88
Слика 41 – Кретање вредности дневних параметара током времена (b) .....	88
Слика 42 – Историјски тренд промене потрошње електричне енергије на основу промене температуре.....	90
Слика 43 – Историјски тренд промене потрошње електричне енергије на основу промене брзине ветра при температури до 20 °C.....	90
Слика 44 – Историјски тренд промене потрошње електричне енергије на основу промене брзине ветра при температури преко 20 °C .....	91
Слика 45 – Историјски тренд промене потрошње електричне енергије на основу промене облачности при температури до 20 °C .....	91
Слика 46 – Историјски тренд промене потрошње електричне енергије на основу промене облачности при температури преко 20 °C .....	92
Слика 47 – Историјски тренд промене потрошње електричне енергије на основу промене трајања дневног светла .....	93
Слика 48 – Историјски тренд промене потрошње електричне енергије по годинама...	94
Слика 49 – Криве сатних вредности потрошње електричне енергије за иницијални скуп изабраних дана.....	103

Слика 50 – Криве сатних вредности потрошње електричне енергије за нови скуп изабраних дана.....	106
Слика 51 – Криве учења и валидације за изабрану структуру неуронске мреже .....	107
Слика 52 – Архитектура неуронске мреже за прогнозу у адаптивном моделу .....	108
Слика 53 – Криве прогнозираних и остварених сатних вредности потрошње за дан 2021-01-25.....	109
Слика 54 – Криве сатних вредности потрошње електричне енергије за изабране сличне дане у односу на специјални дан.....	111
Слика 55 – Резултати симулираних прогноза за историјске специјалне дане.....	113
Слика 56 – Криве прогнозираних и остварених сатних вредности потрошње за специјални дан 2021-01-01.....	115
Слика 57 – Просечна сатна МАРЕ вредност по месецима .....	116
Слика 58 – Криве учења и валидације на примеру саме неуронске мреже .....	117
Слика 59 – Архитектура на примеру саме неуронске мреже .....	119
Слика 60 – Историјски тренд промене потрошње природног гаса на основу промене температуре.....	127
Слика 61 – Историјски тренд промене потрошње природног гаса на основу промене брзине ветра .....	128
Слика 62 – Историјски тренд промене потрошње природног гаса по годинама .....	128
Слика 63 – Криве прогнозираних и остварених вредности потрошње природног гаса .....	134

# Списак табела

Табела 1 – Особине тестних варијабли у оптимизацији генетским алгоритмом .....	63
Табела 2 – Резултати теста генетског алгоритма.....	65
Табела 3 – Параметри који се оптимизују генетским алгоритмом у прогнози потрошње електричне енергије.....	86
Табела 4 – Вредности параметара добијених кварталном оптимизацијом у прогнози потрошње електричне енергије .....	87
Табела 5 – Фактори од утицаја за прогнозу потрошње електричне енергије и њихови партиципациони коефицијенти .....	94
Табела 6 – Дневни подаци временске прогнозе за прогнозирани дан.....	95
Табела 7 – Сатни подаци временске прогнозе за прогнозирани дан .....	95
Табела 8 – Инерциони подаци у односу на прогнозирани дан .....	96
Табела 9 – Дани селектовани иницијалним филтером.....	97
Табела 10 – Сатни подаци о временским приликама за датум 2019-12-02.....	98
Табела 11 – Прорачун коефицијента сличности за температуру за датум 2019-12-02..	99
Табела 12 – Доступни подаци за факторе временских прилика за датум 2019-12-02 .	100
Табела 13 – Инерциони подаци у односу на датум 2019-12-02 .....	100
Табела 14 – Доступни подаци за инерционе факторе .....	101
Табела 15 – Вредности подкоефицијената и њихови тежински коефицијенти.....	102
Табела 16 – Израчунати коефицијенти за изабране датуме .....	104
Табела 17 – Стандардна девијација потрошње електричне енергије за иницијални скуп изабраних дана .....	105
Табела 18 – Стандардна девијација за нови скуп изабраних дана.....	105
Табела 19 – Оптимизоване вредности хипер параметара неуронске мреже за прогнозу потрошње електричне енергије у адаптивном моделу прогнозе .....	108
Табела 20 – Прогнозирана и остварена потрошња електричне енергије по сатима за прогнозирани дан.....	110
Табела 21 – Слични дани изабрани филтером сличних дана у прогнози за специјални дан 2021-01-01.....	111



Табела 22 – Међупрогноза потрошње електричне енергије за специјални дан .....	112
Табела 23 – Резултати линеарне регресије за специјални дан .....	114
Табела 24 – Прогнозирана и остварена потрошња по сатима за прогнозирани специјални дан .....	115
Табела 25 – Оптимизоване вредности хипер параметара на примеру саме неуронске мреже .....	118
Табела 26 – Резултати прогнозе неуронском мрежом и адаптивним моделом прогнозе .....	120
Табела 27 – Просечно време извршавања прогнозе неуронском мрежом и адаптивним моделом прогнозе .....	120
Табела 28 – Упоредни резултати прогнозе добијени Хибридном PSO-SVM методом и адаптивним моделом прогнозе .....	122
Табела 29 – поређење са резултатима прогнозе коришћењем таласасте неуронске мреже базиране на сличним данима .....	123
Табела 30 – Параметри који се оптимизују генетским алгоритмом у прогнози потрошње природног гаса .....	126
Табела 31 – Фактори од утицаја и њихови партиципациони коефицијенти.....	129
Табела 32 – Дневни подаци временске прогнозе за прогнозирани дан.....	129
Табела 33 – Сатни подаци временске прогнозе за прогнозирани дан.....	130
Табела 34 – Инерциони подаци у односу на прогнозирани дан .....	130
Табела 35 – Првих 50 дана селектованих иницијалним филтером.....	132
Табела 36 – Оптимизоване вредности хипер параметара неуронске мреже за прогнозу потрошње природног гаса.....	133
Табела 37 – Поређење прогнозиране и остварене потрошње природног гаса .....	135
Табела 38 – Резултати прогнозе потрошње природног гаса по данима, коришћењем адаптивног модела прогнозе и неуралне мреже .....	136
Табела 39 – Оптимизоване вредности хипер параметара на примеру саме неуронске мреже за прогнозу потрошње природног гаса.....	137
Табела 40 – Резултати прогнозе потрошње природног гаса по данима, коришћењем саме неуралне мреже.....	138
Табела 41 – Просечно време извршавања прогнозе потрошње природног гаса неуронском мрежом и адаптивним моделом прогнозе.....	139

# Листа скраћеница

ACO	Ant Colony Optimization
ANN	Artificial Neural Network
ARIMA	Autoregressive Integrated Moving Average
ARMA	Autoregressive Moving Average
AS	Ant System
avg_cost	Average Cost
avgC	Average Consumption+
avgT	Average Temperature
ccn	Convergence Check Number
Cf	Consumption (forecast)
con_rate	Convergence Rate
Cr	Consumption (realized)
CRISP	Cross-Industry Standard Process
cspd	(C)onsumption – for (sp)ecial (d)ays
Ctspd	(C)onsumption – (t)emporary, for (sp)ecial (d)ays
C $\sigma$	Contribution for deviation
DAO	Data Access Objects
DBN	Deep Belief Network
dd	Daylight Duration
ddwc	Day Decreasing Weighting Coefficient
dp	Date Proximity
EF	Entity Framework
GA	Genetic Algorithm
HDs	Historical Days (Similar)
in	Inertia
LDF	Load Distribution Factor
lhC	Last Hour Consumption
lhin	Last Hour Inertia
LTLF	Long Term Forecasting
MA	Moving Average
MAPE	Mean Absolute Percentage Error
MTLF	Medium Term Forecasting
ND	Number of Days
NDMAPE	(Number of Days) MAPE

nfw	(n)umber of (f)actors for (w)eather
pc	(p)articipation (c)oefficient
PSF-NN	Pattern Sequence Similarity — Neural Networks
PSO	Particle Swarm Optimization
sc	Similarity Coefficient
sc	Sky Cover
sccin	(s)imilarity (c)oefficient for (c)onsumption (in)ertia
schin	(s)imilarity (c)oefficient for (l)ast (h)our (in)ertia
sctin	(s)imilarity (c)oefficient for (t)hermal (in)ertia
scw	(s)imilarity (c)oefficient for (w)eather
STLF	Short Term Forecasting
SVM	Support Vector Machine
t	Temperature
UI	User Interface
UML	Unified Modeling Language
w	(w)eather factor value in day (d) for hour (i)
wc	(w)eighting (c)oefficient
wi	Wind
$\sigma$	Standard deviation

# 1. Увод

У данашње време, снабдевање становништва и привреде ресурсима као што су електрична енергија, природни гас или вода не сме да трпи последице неадекватног планирања и управљања инфраструктурним системима. Савремене инфраструктурне мреже све више се ослањају на софтверске системе који су намењени „паметном“ управљању и планирању ресурса. Ови софтверски системи обухватају функције као што су прорачуни токова ресурса, управљање испадима у мрежи, оптимално географско и инфраструктурно распоређивање постројења и читав низ других процеса којима се оптимизује поуздано снабдевање.

Један од ових процеса је и прогноза појава (феномена), као што су потрошња електричне енергије, потрошња природног гаса, нафте или воде. Према временском периоду који се прогнозира, прогноза се може поделити на дугорочну и краткорочну. Док дугорочна прогноза представља грубље планирање кумулативних вредности потрошње, краткорочна прогноза најчешће се односи на временске серије на сатном нивоу, у периоду од неколико сати до неколико дана. Поуздана краткорочна прогноза представља основу за оперативно планирање и поузданост свих других функција у паметним мрежама. Прогнозирање ресурса разликује се у зависности од врсте инфраструктурне мреже, односно од тога да ли се ради о потрошњи електричне енергије, природног гаса, нафте или воде. Свака од ових врста прогнозе има своје специфичности и правила у прорачунима.

Постоје бројне савремене технике које се могу користити за прорачуне прогнозе. Међу постојећим техникама за краткорочну прогнозу издвајају се оне које спадају у област машинског учења: регресиона анализа [1, 2, 3, 4], неуронске мреже [5, 6, 7, 8, 9, 10], расплинута логика [11, 12, 13, 14], метода потпорних вектора [15, 16] и многе друге. У циљу што боље прогнозе, често се користе и комбиноване методе [17, 18]. Ове технике обављају прогнозу на основу различитих параметара, који се разликују у зависности од врсте ресурса, периода који се прогнозира, географског подручја итд. Међутим, оно што је заједничко за све наведене врсте прогнозе јесте то да се прогноза не може обавити без адекватног узорка.

Предмет истраживања у дисертацији јесте унапређење поступака краткорочне прогнозе временских серија у паметним инфраструктурним мрежама. Иако су методе машинског учења за праћење законитости наведених појава адекватно покривене у доступној литератури, то се не може тврдити и за технике избора узорка на који се ове методе ослањају. Евидентно је да у овој области постоји доста простора за унапређење. Предлаже се иновативни адаптивни модел за развој софтверских система намењених

прогнози вредности временских серија, који се ослања на оптимизацију избора узорка из историјских података. Појаве за чију је прогнозу овај модел намењен имају следеће особине:

- Ради се о периодичним појавама које се могу проверавати (мерити) у правилним временским размацама у току дана.
- Имају (ауто)корелацију са вредностима у прошлости.
- Постоје фактори који су мерљиви и који утичу на манифестацију појава (феномена).

Предложени модел може равноправно да се користи за различите врсте ресурса, као што су електрична енергија, природни гас и вода. Моделом се стварају поставке за предвиђање на основу сличних дана у прошлости. Слични дани се бирају на основу детектованих фактора од утицаја, а сваки детектовани фактор има одговарајући тежински коефицијент. Предложеним моделом оптимизују се вредности тежинских коефицијената тако да њихова примена омогућава избор оптималног скупа сличних дана, који ће служити као узорак у поступку прогнозе.

У избору оптималног узорка, односно оптималног скупа сличних дана, посебан проблем представља променљивост окружења. Окружење се мења, како на сезонском нивоу (у току године) тако и дугорочно (промена капацитета, трендови и навике потрошача итд.). Ово намеће потребу за адаптивним моделом, који ће се самостално мењати и прилагођавати променама окружења. Према предложеном моделу вредност тежинских коефицијената аутоматски се прилагођава календарском добу у години и другим утицајима. Односно, апликације које имплементирају предложени модел имају могућност адаптивности, тако да се у сваком моменту бира узорак сличних дана на основу оптимално прорачунатих фактора од утицаја, те се прорачун прилагођава датим условима.

Предложени модел поседује и могућност идентификовања специјалних дана. Специјални дани су најчешће празници и други дани који садрже специфичности захваљујући којима профили потрошње одступају од стандардних и очекиваних вредности. Адаптивни модел прогнозе је у стању да креира адекватну прогнозу за овакве дане, примењујући уочена правила на резултате алгоритма прогнозе.

Циљ истраживања и примарни допринос ове докторске дисертације је унапређење постојећих техника машинског учења. То је постигнуто увођењем иновативног модела за развој софтверских система у паметним инфраструктурним мрежама за предвиђање и прогнозу феномена који се манифестују у временским серијама на дневном/сатном нивоу. Модел који се уводи заснива се на оптимизацији избора узорка за предвиђање и унапређује постојеће алгоритме машинског учења у погледу резултата прогнозе и перформанси.

На основу главног циља произилазе и пратећи циљеви:

- Ефикасна оптимизација избора сличних дана коришћењем одговарајуће метахеуристичке методе. Ово подразумева адекватан избор метахеуристичке методе и унапређење постојећих поступака оптимизације у циљу што тачније крајње прогнозе.
- Унапређење перформанси процеса краткорочне прогнозе потрошње кроз оптималан избор узорка.
- Оптималан избор узорка током целе године, коришћењем дневне и кварталне оптимизације.
- Препознавање оптималних вредности тежинских коефицијената фактора од утицаја, и у складу са тим повећање тачности резултата у паметним мрежама. Повећање тачности резултата се референцира на тренутно стање у области, а доказује се на примеру преносне електроенергетске мреже за подручје Србије, и на примеру дистрибутивне мреже природног гаса за један регион у Холандији.
- Успешна примена иновативног концепта прогнозе потрошње у специјалним данима, што подразумева резултате прогнозе који по тачности не одступају од стандардних дана.
- Могућност примене модела за прогнозу потрошње различитих врста ресурса.

Из аспекта прогнозе временских серија, равноправно се користе појмови потрошње ресурса и потражње за ресурсима. Наиме, за неке врсте ресурса се прогнозира потрошња (на пример електрична енергија и природни гас) јер је претпоставка да ће се сва потражња претворити у потрошњу. Код ових ресурса постоји могућност надокнаде мањка ресурса кроз додатни увоз, активирање резервних капацитета итд. У случају других ресурса, као што је вода, у зависности од инфраструктуре, географског подручја и метеоролошких прилика, постоји могућност да потражња буде већа од потрошње, тако да се у тим случајевима говори о прогнози потражње. Међутим, свака потрошња произилази из потражње, и свака потражња може да се исказе као потенцијална потрошња, тако да ће се, са аспекта краткорочне прогнозе временских серија, ова два појма користити равноправно.

У наставку је наведена структура дисертације.

У поглављу 2 представљен је концепт прогнозе временских серија у паметним инфраструктурним мрежама. Такође, објашњени су појмови који су неопходни за сагледавање проблема: прогноза потражње за ресурсима и прогноза временских серија. У оквиру прогнозе потражње за ресурсима анализирана је прогноза за потражњом различитих врста ресурса, као што су електрична енергија, вода и природни гас. У разматрању прогнозе временских серија представљено је машинско учење као јединствена научна област која се користи за потребе прогнозе потражње за ресурсима. Посебно су анализирани линеарне и нелинеарне методе машинског учења, као и могућност њихове примене у поступцима прогнозе. Коначно, уведена је могућност избора сличних дана као узорка у алгоритме машинског учења.

У поглављу 3 постављена је хипотеза дисертације, која се односи на креирање јединственог адаптивног модела прогнозе који ће омогућити побољшање крајњих резултата прогнозе. У овом поглављу представљене су и појединачне ставке које се односе на начин остварења основне хипотезе.

У поглављу 4 направљен је преглед литературе која је у вези са проблематиком дисертације. Описани су начини избора узорка у краткорочној прогнози потрошње ресурса који су до сада обрађени у литератури. Наведени су радови који се баве избором узорка у прогнози потрошње краткорочне енергије, воде и природног гаса. Посебно је обрађена област метахеуристичких метода, који представљају основу оптимизације узорка у дисертацији.

У поглављу 5 представљена је архитектура предложеног модела: појединачни елементи и везе између њих. Посебно су обрађени филтери који служе за селекцију сличних дана и метахеуристичка оптимизација. У овом поглављу представљени су могући избори метахеуристичке методе, и на основу урађених анализа изабран је унапређени генетски алгоритам као метод оптимизације.

Поглавље 6 бави се унапређењем концепта генетског алгоритма ради што ефикасније оптимизације у предложеном моделу. Представљени су концепти насељавања, мутације засноване на диверзитету и мутације ка екстремима.

У поглављу 7 представљена је имплементација адаптивног модела прогнозе временских серија кроз дизајн апликације која примењује концепте модела. Приказане су функционалности апликације, архитектура, модел података и апликативни модели. За приказ дизајна коришћен је UML језик моделовања.

Поглавље 8 бави се студијама случаја. Једна студија случаја обавља се на реалном моделу и приказује примену концепата описаних у моделу. Студија се бави прогнозом потрошње електричне енергије у преносној мрежи електроенергетског преносног система „Електромрежа Србије“. Анализирају се кораци прогнозе за један обичан и један специјални дан. Посебно се анализира оптимизација параметара генетским алгоритмом. Анализа резултата се бави периодом од једне године. Друга студија случаја представља симулацију процеса прогнозе потрошње природног гаса за један регион у Холандији. Анализа резултата бави се периодом од једног месеца. Да би се утврдио допринос предложеног модела, у обе студије случаја пореде се резултати прогнозе коришћењем саме неуронске мреже и коришћењем адаптивног модела прогнозе временских серија.

Поглавље 9 садржи закључна разматрања и правце даљег истраживања.

## 2. Концепт прогнозе временских серија у паметним инфраструктурним мрежама

Прогноза временских серија често се употребљава у поступцима машинског учења. Прогноза потражње за ресурсима може да се обавља унутар и изван временских серија. У овом поглављу биће ближе описани појмови прогнозе потражње за ресурсима и прогнозе временских серија.

### 2.1. Прогноза потражње за ресурсима

Прогнозирање потражње за ресурсима у инфраструктурним мрежама нераскидиво је повезано са процесом планирања. Планирање омогућава стратегију и оперативне акције, на основу добијених прогноза. Са друге стране, прогнозе омогућавају анализу над резултатима који су добијени планирањем [19]. Може се рећи да потреба за прогнозом потражње постоји свуда где постоји потреба за ресурсима.

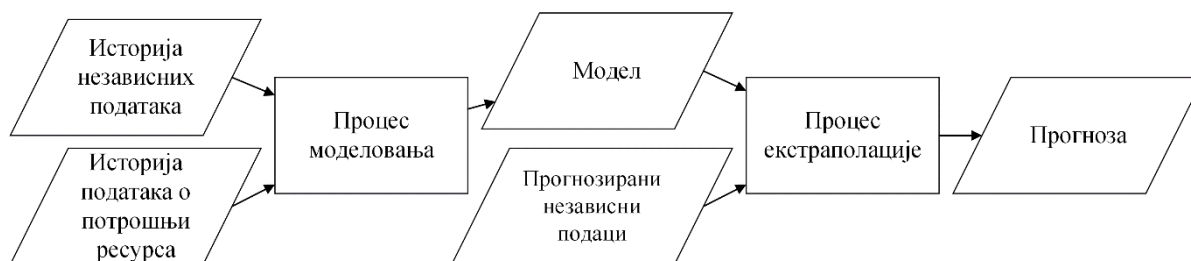
Потреба за адекватним планирањем се посебно истиче у инфраструктурним мрежама које служе за снабдевање становништва и привреде виталним ресурсима, као што су струја, природни гас и вода. Стога се у компанијама које се баве снабдевањем ових ресурса све више користе системи за паметно управљање ресурсима. То подразумева интеграцију свих чинилаца, и то на такав начин да интелигентно комуницирају у циљу обезбеђења ефикасне, одрживе, економичне и сигурне испоруке ресурса [20]. Овакви системи се зову „паметне мреже“ (енгл. *smart grids*). Према концепту паметних мрежа, прогноза потражње за ресурсима игра кључну улогу, јер се на њу, осим планирања, ослања више других повезаних процеса. Ови процеси обухватају управљање потражњом, интеграцију обновљивих извора, складиштење, логистику, планирање ремонта, економски диспечинг, прорачун оптималног ангажовања агрегата и друго [21].

Не постоји јединствен критеријум за класификацију прогнозирања потреба за ресурсима. Класификација може да зависи од врсте ресурса за које се користи прогноза, географског и инфраструктурног подручја, броја потрошача итд. Класификација која се у литератури најчешће спомиње јесте она која се тиче временског периода који се прогнозира. На основу тога може да се изврши подела на краткорочну (енгл. *Short Term Forecasting* – *STLF* – прогноза за 1-2 дана), средњорочну (енгл. *Medium Term Forecasting* – *MTLF* – до једног месеца) и дугорочну (енгл. *Long Term Load Forecasting* – *LTLF* – од



једног месеца до једне године) [19]. Према неким ауторима, краткорочна прогноза обухвата прогнозу за период до седам дана [22, 23]. Очекивано, прецизност прогнозе опада са дужим временским периодом. Док се за средњорочну и дугорочну прогнозу најчешће прогнозирају кумулативни дневни/недељни/месечни периоди, краткорочна прогноза најчешће обухвата прогнозу за временске серије у дефинисаним размацама: 15 минута, пола сата или сат.

Слика 1 приказује уопштен модел прогнозе потражње за ресурсима. Претпоставка је да постоје подаци о факторима који утичу на потрошњу ресурса. У том смислу, подаци о потрошњи ресурса називају се зависни подаци, док се подаци о факторима који утичу на потрошњу називају независни подаци. Историја података о потрошњи и историја независних података користе се као улаз за моделовање процеса. Након што се у процесу моделовања израчунају одговарајући параметри, креира се одговарајући модел. Заједно са моделом, улаз у процес екстраполације представљају независни подаци за временски период прогнозирања, за које се унапред знају вредности потпуно тачно или са великом вероватноћом тачности. Процес екстраполације доводи до коначне прогнозе потрошње.



Слика 1 – Типичан процес прогнозе потражње за ресурсима

Независни подаци разликују се према врсти ресурса за који се прогнозира потрошња. Међутим, и у оквиру прогнозе за исту врсту ресурса постоје бројне варијације. За прогнозу потрошње електричне енергије уобичајено се користе метеоролошки фактори: температура, ветар, облачност итд. Утицај на прогнозу сваког од њих може да варира током времена. Слично је и код прогнозе потражње за природним гасом, где су такође доминантни метеоролошки фактори попут температуре, влажности ваздуха, облачности и падавина. Зависни параметри у прогнози потражње за водом обухватају количину кише, дужину сушних периода, рестрикције воде и др.

За процес моделовања на располагању су бројне методе, које се у највећем броју случајева могу сврстати у процесе машинског учења. Која год метода да је изабрана, потребно је да буде прилагођена прогнози временских серија, најчешће на сатном, полусатном или 15-минутном нивоу.

Прогноза потрошње ресурса у временским серијама најчешће се односи на електричну енергију. Међутим, ефикасно управљање ресурсима све чешће доноси потребу за прогнозом потрошње и других типова ресурса, као што су вода и природни гас [24, 25].

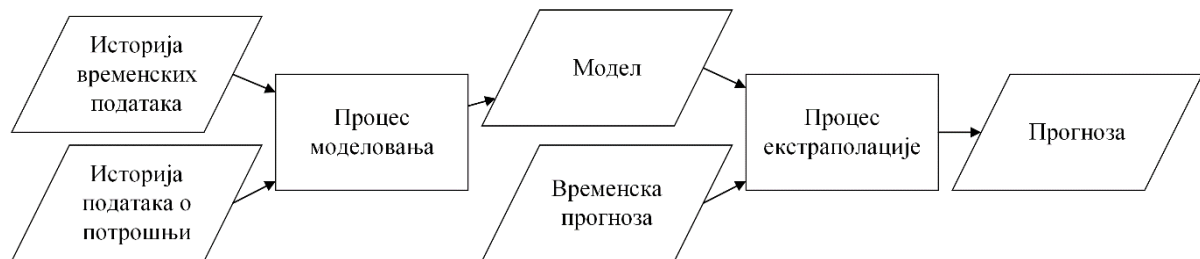
Концепт представљен у дисертацији може равноправно да се примени на све ресурсе чија се потрошња манифестује у временским серијама, а посебно се наглашава прогноза потрошње електричне енергије, природног гаса и воде.

### 2.1.1. Краткорочна прогноза потрошње електричне енергије

У савременим електричним постројењима, краткорочна прогноза потражње електричне енергије је есенцијална за поуздано извршавање операција, а значајно утиче и на економске аспекте система напајања. *STLF* представља један од критичних корака електричних постројења и преносних система за доношење одлука које се тичу продаје и куповине електричне енергије, балансирања система, планирања одржавања [26], планирања преноса и дистрибуције и финансијског планирања [19].

Специфичност електричне енергије у односу на друге ресурсе састоји се у томе што не постоји могућност складиштења. Према томе, сваки вишак у планирању преноса и дистрибуције електричне енергије представља трошак без приноса, док сваки мањак може да утиче на поуздан рад или може захтевати хитну куповину енергије на берзи која је по правилу скупа.

Слика 2, као адаптација типичног процеса прогнозе потражње за ресурсима приказаног на слици 1, приказује стандардни *STLF* процес који се извршава у постројењу које се ослања на временске информације и историјске податке о потрошњи [19]. Историјске информације о времену и потрошњи користе се као улаз за моделовање процеса.



Слика 2 – Типичан *STLF* процес

#### 2.1.1.1. Методе краткорочног прогнозирања потрошње електричне енергије

Процес моделовања обухвата обраду историјских података о потрошњи и историјских временских података, кроз утврђивање правила зависности између ове две групе података. Кроз ова правила утврђује се како временски подаци, као независни подаци, утичу на податке о потрошњи. На модел се примењују подаци о временској прогнози као независни подаци и кроз процес екстраполације добијају се подаци о прогнозираној потрошњи електричне енергије.

Процес моделовања може се одвијати кроз више различитих метода, као што су линеарна регресија [27], неуронске мреже [5, 6, 7] или расплнута логика [11, 12, 13].

Историјски временски подаци обухватају серијске вредности метеоролошких параметара, у одговарајућој резолуцији (један сат, пола сата или 15 минута), за претходни историјски период, на пример за претходних пет година. Међутим, зависност између података о потрошњи није линеарна и мења се током времена, током једног дана, током недеље или током године. Ову зависност директно одређује сличност историјског дана са даном који се прогнозира. Да би прогноза била што прецизнија, могуће решење је креирање посебног модела који би се бавио што адекватнијим избором историјских података, односно дана из прошлости, који ће бити што сличнији са даном који се прогнозира.

### **2.1.1.2. Фактори који утичу на потрошњу електричне енергије**

На слици 2, као независни подаци у моделовању зависности приказани су историјски временски подаци. Познато је да временски подаци значајно утичу на потрошњу електричне енергије [19]. Разлози за ово су бројни и обухватају коришћење електричних уређаја за регулацију температуре у домовима, електричну енергију која је потребна за функционисање других средстава за грејање (на пример природног гаса) и боравак људи у домовима за време ниских температура. Поред температуре, временски параметри који утичу на потрошњу су ветар, влажност ваздуха и облачност. Поред ових, постоје и изведени временски подаци који могу да утичу на потрошњу. Неки од њих су разлика у температури у односу на претходни сат, субјективни осећај температуре и комбинација температуре са другим факторима, као што су ветар или важност.

Независни подаци у моделовању зависности треба да имају резолуцију исту као и период који се прогнозира (један сат, пола сата или 15 минута). Међутим, за моделовање зависности могу да се користе и подаци са другачијом резолуцијом, на пример подаци на дневном нивоу. Ови подаци обухватају просечну, максималну или минималну температуру, просечан ветар или просечну облачност.

Иако се за краткорочну прогнозу потрошње електричне енергије најчешће спомињу временски подаци, постоје и друге врсте података који могу значајно да утичу на потрошњу. Међу њима се издвајају календарски подаци и инерциони подаци. Календарски подаци који утичу на потрошњу нису често презентовани у доступној литератури, а у овом раду биће речи о утицају дужине дневног светла у току дана и о утицају временске удаљености дана од дана који се прогнозира. Инерциони подаци односе се на потрошњу у претходном периоду и на временске податке у претходном периоду, и њихов утицај на потрошњу. У овом раду ће бити уведен феномен утицаја потрошње у нивовима дана из прошлости на низ дана у будућности.

Поставља се питање: како одредити које је од споменутих података потребно узети у обзир приликом прогнозе потрошње. Ако се за историјске податке узимају само дани који су у довољној мери слични са прогнозираним даном, потребно је да се утврди на основу којих параметара да се сличност одреди и какав је утицај на сличност сваког

појединачног параметра. У дисертацији ће бити речи о томе како да се утврди утицај независних параметара на оптимизацију узорка за прогнозу, односно на оптималан избор сличних дана.

## **2.1.2. Краткорочна прогноза потражње за водом**

Када се говори о прогнози потражње за ресурсима, прогноза потрошње електричне енергије је евидентно најзаступљенија у литератури и постојећим техничким решењима. Међутим, све чешће се помиње и прогноза потражње за другим типовима ресурса, као што је прогноза потражње за водом. Под потражњом за водом подразумева се потражња за пијаћом и техничком водом. Према извештајима Уједињених нација, током двадесетог века глобална потрошња воде је порасла шестоструко, дупло брже од раста броја становништва. Смањење извора пијаће воде, проблеми са квалитетом узроковани загађењем, застарела инфраструктура и потребе за новим капацитетима узрокују проблеме са водоснабдевањем широм света [28]. Међу осталим начинима за решавање проблема, значајно место заузима и планирање потражње за водом, које је тесно повезано са адекватном прогнозом. Док је дугорочно планирање потражње за водом корисно у планирању, пројектовању и креирању планова проширења за постојеће системе, краткорочно планирање је корисно за ефикасне операције и управљање постојећим системима водоснабдевања [10].

### **2.1.2.1. Методе краткорочног прогнозирања потражње за водом**

У прошлости се за прогнозу потражње за водом у насељеним подручјима најчешће користио метод „по глави становника“ (енгл. *Per-capita method*). Овај метод подразумева да једна независна променљива (број становника) обезбеђује адекватно објашњење за потрошњу воде. Током времена овакав начин прогнозе се показао поуздан само за веома грубе процене [29], тако да су га заменили математички модели и модели машинског учења. Процес моделовања зависности укључује методе као што су неуронске мреже [10, 30, 16], методе потпорних вектора [16], хибридне методе [17] и друге.

### **2.1.2.2. Фактори који утичу на потражњу за водом**

Фактори који утичу на потражњу за водом, односно независни подаци у моделовању зависности, могу у значајној мери да се разликују у зависности од географског подручја за које се прогнозира потражња, економских карактеристика подручја и врсте инфраструктуре за водоснабдевање. Међу метеоролошким факторима, један од најчешће коришћених јесте количина падавина [31]. Овај фактор универзално утиче на потрошњу воде, а нарочито у сушним подручјима и у подручјима са развијеним системима вештачког наводњавања. Повезани са овим фактором су број сушних дана и број кишних дана. Један од често коришћених метеоролошких фактора је и температура

ваздуха, која, заједно са другим метеоролошким факторима, утиче на повећано конзумирање воде и коришћење климатских уређаја који троше воду [29].

Поред метеоролошких, постоје и други фактори, као што је број становника [32], чији број може да варира и на дневном нивоу, нарочито у време годишњих одмора, и нарочито у туристички развијеним подручјима.

### **2.1.3. Краткорочна прогноза потрошње природног гаса**

Производња природног гаса састоји се од процеса производње, транспорта и дистрибуције. Као и друга фосилна горива, гас се проналази испод земље, у налазиштима која су блиско повезана са налазиштима нафте. Често је природни гас нуспроизвод нафтних бушотина. Након што се сирови природни гас преради, он се складишти или се цевима транспортује локалним дистрибуционим компанијама. Дистрибуционе компаније дистрибуирају гас становништву и комерцијалним и индустријским потрошачима. Дистрибуција се најчешће извршава преко локалних дистрибутера који се зову *оперативна подручја* [33, 8]. Оперативни и комерцијални поступци и одлуке у дистрибуционим компанијама и оперативним подручјима зависе од поуздане прогнозе потрошње природног гаса. Непрецизна краткорочна прогноза изискује високе трошкове за дистрибутере, оперативна подручја и потрошаче. Потцењена прогноза може да доведе до потребе за куповином гаса на тржишту по већим ценама. Прецењена прогноза може да доведе до потребе за непредвиђеним складиштењем гаса, што такође може да доведе до високих трошкова.

#### **2.1.3.1. Фактори који утичу на потражњу за природним гасом**

Дистрибуирани гас може имати више намена, али је најчешћа намена грејање домова и пословних зграда. Потрошња гаса у сврху загревања зове се топлотно оптерећење [8]. Промене топлотног оптерећења највише зависе од метеоролошких фактора, пре свега од спољне температуре ваздуха. Поред температуре, на топлотно оптерећење могу да утичу и други метеоролошки фактори, као што су влажност ваздуха, облачност и падавине. Топлотно оптерећење је на највишем нивоу током зиме, при ниским температурама, док је током лета приближно нула. Друге врсте употребе природног гаса, као што је кување, коришћење апарата за домаћинство, загревање воде и сл. заједно се зову базно оптерећење. На базно оптерећење у већој мери утиче бројност становништва, тако да су у краткорочној прогнози значајни фактори који утичу на дневно кретање становништва, на пример празници, годишњи одмори и нерадни дани. Базно оптерећење је у мањој мери узроковано метеоролошким факторима, пре свега због утицаја на дневно кретање становништва и излазак из домова.

Природни гас је фактор који је специфичан по томе што значајан број потрошача поседује могућност коришћења алтернативног ресурса за потребе грејања. То се пре

свега односи на електричну енергију, али и на друге ресурсе (дрва, угаљ, итд.). Због тога је цена гаса значајан фактор у краткорочној прогнози потрошње природног гаса [34].

### **2.1.3.2. Методе краткорочног прогнозирања потрошње природног гаса**

Уз краткорочну прогнозу потрошње гаса везују се проблеми типични за прогнозу потрошње у инфраструктурним системима: комплексне и нелинеарне зависности различитих фактора на потрошњу [9]. Први коришћени методи прогнозе били су засновани на принципима апроксимације. Након тога, популарне су постале статистичке методе [34], док се у новије време користе неуронске мреже [8, 9] и расплнута логика [14].

## **2.2. Прогноза временских серија**

Временске серије представљају приказ података вредности жељених параметара у правилним временским размацима. Главни циљ анализе временских серија јесте да се прогнозирају будуће вредности које још увек нису измерене. Један од најчешће коришћених метода за решавање проблема прогнозе временских серија јесте машинско учење. Машинско учење може се описати као научна област којом се тежи да рачунари добију могућност учења, без експлицитног програмирања [35]. Другим речима, у областима машинског учења тежи се решавању задатих проблема не тако што се правила за решење програмирају, него тако што их рачунар (односно програм) „научи“, на основу доступних корисних података. У основи машинског учења је знање и коришћење података на адекватан начин. То укључује прикупљање одговарајућих података, прочишћавање података и процесирање података итеративним коришћењем алгоритама учења. Подаци се користе да би се саградили модели коришћењем кључних одлика података и да би се креирало предвиђање на основу хипотеза добијених из изграђених модела [36].

Као што је раније приказано, краткорочна прогноза за ресурсима као што су електрична енергија, вода и гас, претпоставља прогнозу у временским серијама.

### **2.2.1. Машинско учење у прогнозирању потражње за ресурсима**

Потражња ресурса у инфраструктурним системима односи се на количину потребног ресурса у периодима у временској серији. То може да буде количина потрошње електричне енергије, природног гаса или воде на сатном нивоу. Машинским учењем могу се прогнозирати ове вредности, на основу доступних података који се односе на измерене податке и на предвиђене податке. Измерени подаци су вредности сатне количине потрошње ресурса из прошлости за дужи временски период. Поред тога, измерени подаци се односе и на измерене вредности оних фактора који су утицали на

потрошњу ресурса. На пример, у случају потрошње електричне енергије, фактори који утичу на потрошњу су температура, ветар, облачност, трајање дневног светла итд. *Предвиђени подаци* су будући подаци о факторима који утичу на потрошњу ресурса. Предвиђени подаци могу да буду прогнозирани (на пример у случају метеоролошких података) или израчунати (на пример у случају календарских података, попут дужине трајања дневног светла). За предвиђене податке постоје вредности за период у коме ће предвиђање бити извршено. У случају потрошње електричне енергије то би била прогнозирана температура, ветар, облачност, трајање дневног светла итд. Поузданост тачности предвиђених података разликује се за сваки фактор појединачно. На пример, за трајање дневног светла поузданост је потпуна, а за прогнозу температуре у сату поузданост је мања и смањује се са повећањем временског интервала за који се прогнозира потражња. Степен утицаја фактора на потражњу такође може да варира. Утицај неког фактора може да се мења у зависности читавог низа околности: доба године, времена у току дана, изражености других фактора од утицаја итд. На пример, облачност може да утиче на потрошњу електричне енергије тако што појачава или снижава утицај температуре.

У прогнози потражње за ресурсима, методе машинског учења врше „учење“ утицаја измерених вредности фактора из прошлости на измерене вредности потрошње ресурса из прошлости. Научена правила се примењују на предвиђене вредности фактора од утицаја у будућности и тиме се предвиђа потражња за ресурсима.

## **2.2.2. Методе машинског учења**

У доступној литератури [37, 38, 39, 40] постоје многе методе машинског учења које се односе на прогнозу временских серија. Неке од метода су линеарне, односно односе се на временске серије у којима је задовољен услов линеарности. Један од најједноставнијих метода је мултиваријабилна линеарна регресија. Од ове методе потичу и линеарне методе прогнозе временских серија, као што су *ARMA* (енгл. *Autoregressive Moving Average*) и *ARIMA* (енгл. *Autoregressive Integrated Moving Average*). *ARIMA* је најчешће коришћени метод анализе временских серија у којима је задовољен услов линеарности. Међутим, прогнозе временских серија потрошње у инфраструктурним мрежама ретко могу да задовољавају овај услов [41]. Међу нелинеарним методама истичу се неуронске мреже (енгл. *Artificial Neural Network – ANN*), расплинуте временске серије (енгл. *Fuzzy Time Series*) и хибридне методе [18].

## **2.2.3. Линеарне методе машинског учења**

### **2.2.3.1. Мултиваријабилна линеарна регресија**

Регресиона анализа представља поље статистике које се користи ради испитивања међузависности између зависне променљиве и независних променљивих. Кључни елемент у регресионој анализи је естимација регресионе једначине, којом се описује

колико у просеку зависна варијабла зависи од независних варијабли [42]. Линеарна регресија је такав модел према којој је зависна варијабла у линеарној зависности од независних варијабли. Случај када постоји само једна независна варијабла зове се проста линеарна регресија, док се случај када постоји више независних варијабли зове **мултиваријабилна линеарна регресија**. Формално, за  $n$ -димензионални узорак, модел мултиваријабилне линеарне регресије може се представити на следећи начин [43]:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}, \quad (1)$$

где је:

$y_i$  – зависна варијабла  $i$ -те опсервације у узорку,

$x_i$  – независна варијабла  $i$ -те опсервације у узорку,

$\beta_0$  – одсечак,

$\beta_p$  – нагиб за варијаблу  $p$ .

Када се примене претпостављене вредности за одсечак  $\beta_0$  и нагибе  $\beta_1 \dots \beta_p$ , добијају се предвиђене (фитоване) вредности  $\hat{y}_i$ . Вредности за одсечак  $\beta_0$  и нагибе  $\beta_1 \dots \beta_p$  одређују се методом најмањих квадрата. То значи да се вредности за  $\beta_0$  и за  $\beta_1 \dots \beta_p$  одређују тако да је минимализована сума квадрата разлика између  $y_i$  и  $\hat{y}_i$ . Постоје математичка решења за израчунавање одсечка и нагиба и та решења су уграђена у статистичке софтверске апликације и системе машинског учења.

У случају прогнозе потражње за ресурсима у инфраструктурним системима, прогнозира се свака тачка у временској серији (на пример за сваки сат). За сваки сат у току дана креира се и  $n$ -димензионални узорак који би се састојао од измерених вредности потрошње ( $y_i$ ) и измерених вредности фактора од утицаја ( $x_{i1}, x_{i2}, \dots, x_{ip}$ ). Мултиваријабилном регресионом анализом одређивали би се тежински коефицијенти сваког од фактора од утицаја ( $\beta_1 \dots \beta_p$ ) и одсечак ( $\beta_0$ ) који би представљао фиксну потрошњу ресурса која би постојала и без утицаја сваког од фактора од утицаја.

### 2.2.3.2. ARMA и ARIMA модели

ARMA и ARIMA користе се за прогнозу временских серија које испуњавају услов линеарности. ARMA модел се заснива на процесу ауторегресије и процесу покретних просека (енгл. *Moving Average*).

Према [44], ауторегресивни модели се користе да опишу процесе који се мењају у времену. Док се мултиваријабилни регресиони модел користи за прогнозирање вредности зависне променљиве коришћењем линеарне комбинације зависних и независних променљивих, ауторегресивним моделом се предвиђају зависне променљиве коришћењем линеарне комбинације прошлих вредности зависне варијабле.

Формално, ауторегресивни модел реда  $p$ ,  $AR(p)$  представља се на следећи начин:



$$y_t = c + \sum_{i=1}^p \varphi_i y_{t-i} + \varepsilon_t, \quad (2)$$

где је:

$y_t$  – вредност зависне променљиве у тренутку  $t$ ,  
 $y_{t-1}$  – вредност зависне променљиве у тренутку  $t-1$ ,  
 $\varphi_1, \dots, \varphi_p$  – параметри модела у временским тачкама,  
 $c$  – константа,  
 $\varepsilon_t$  – бели шум.

Модел покретних просека у предвиђању вредности зависне променљиве користи просек серије у тачки у којој се посматра и грешке у претходним тачкама временске серије. Модел покретних просека реда  $q$ ,  $MA(q)$ , формално се може представити као:

$$y_t = \mu + \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i}, \quad (3)$$

где је:

$y_t$  – вредност зависне променљиве у тренутку  $t$ ,  
 $\mu$  – просек серије,  
 $\theta_1, \dots, \theta_q$  – параметри модела у временским тачкама,  
 $\varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-q}$  – бели шум.

Дакле, вредност зависне променљиве у једној тачки у временској серији зависи од просека вредности у претходним тачкама, модификованих за вредност зависних параметара и вредност грешке.

$ARMA(p,q)$  модел представља комбинацију ауторегресивног модела и модела покретних просека:

$$y_t = c + \varepsilon_t + \sum_{i=1}^p \varphi_i y_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i}. \quad (4)$$

Према овом моделу, независне варијабле са десне стране формуле садрже и претходне вредности зависне променљиве и претходне грешке. Ако се свака претходна вредност  $y_{t-1}$  за свако  $t > 1$ , односно,  $Lag$  оператор означи са  $L$ , иста формула се може представити са:

$$(1 - \sum_{i=1}^p \varphi_i L^i) y_t = (1 + \sum_{i=1}^q \theta_i L^i) \varepsilon_t. \quad (5)$$

$ARIMA$  модел представља генерализацију  $ARMA$  модела и користи се у ситуацијама када подаци показују тенденцију нестационарности, односно када се аритметичка средина мења током времена. Термин *Integrated* означава да се вредности података замењују разликама вредности између тачака у временској серији. Формално се овај модел означава  $ARIMA(p,d,q)$ , где  $p$  представља ред ауторегресије,  $d$  представља степен разлике (означава колико пута се вредност смањила од почетка временске серија), а  $q$  представља ред покретних средина:

$$(1 - \sum_{i=1}^p \varphi_i L^i)(1 - L)^d y_t = (1 + \sum_{i=1}^q \theta_i L^i) \varepsilon_t. \quad (6)$$

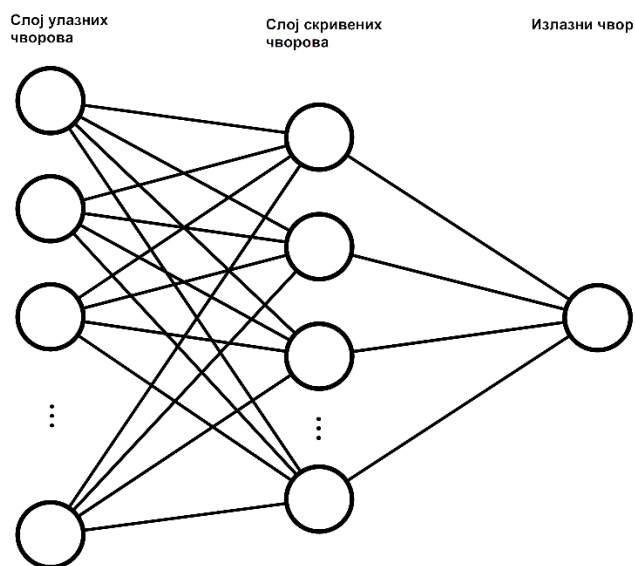
## 2.2.4. Нелинеарне методе машинског учења

### 2.2.4.1. Неуронске мреже

Како зависности између фактора од утицаја и прогнозиране потражње за ресурсима најчешће не задовољавају услов линеарности, јавља се потреба за сложенијим моделима. *Неуронске мреже* су један од најчешће коришћених алгоритама за нелинеарно моделовање временских серија у многим областима као што су енергетика, финансије, хидрологија и мрежне комуникације [45]. Неуронске мреже представљају математички алат који је инспирисан начином на који људски мозак обрађује информације. У основи, неуронска мрежа може да израчуна сваку функцију подложну рачунању [46].

Основна јединица је вештачки неурон. Неурон прима информације преко већег броја улазних чворова, интерно их обрађује и обезбеђује одговор. Обрада информације се обично обавља у два корака: прво се улазне вредности линеарно комбинују, а затим се резултат користи као аргумент за нелинеарну активациону функцију. Свака комбинација користи тежински коефицијент (енгл. *weight*) и константу нагиб (енгл. *bias*) [47]. Неуронске мреже у којима везе између чворова не формирају затворен круг зову се *Feedforward* неуронске мреже (слика 3). У њима су неурони организовани у неколико слојева, и сваки неурон може да се преноси само до неурона у следећем слоју [46]. Учење у неуронским мрежама представља прилагођавање тежинских коефицијената и одсечака, уз помоћ минимизације уочених грешака. Грешке се посматрају за сваку употребљену комбинацију вредности тежинских коефицијената и нагиба.

Структура неуронске мреже подешава се уз помоћ хипер параметара. Неки од хипер параметара су стопа учења, оптимизатор, активациона функција, број епоха и број пакета. Оптималне вредности хипер параметара разликују се за сваки појединачни случај имплементације неуронске мреже. Подешавање ових вредности може да се изврши испробавањем различитих комбинација, праћењем криве учења или коришћењем неке од оптимизационих техника.



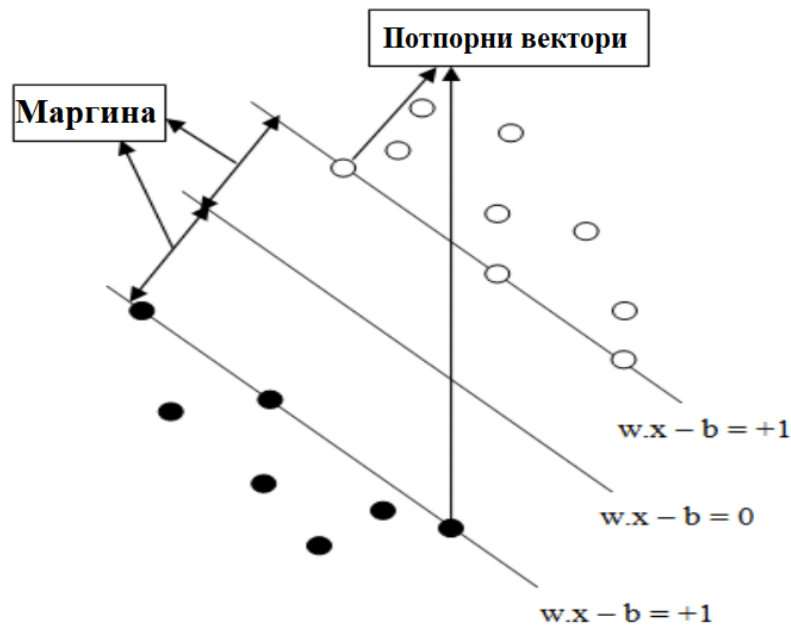
Слика 3 – Структура Feedforward неуронске мреже

#### 2.2.4.2. Метод потпорних вектора

Метод потпорних вектора (енгл. *Support Vector Machine – SVM*) представља релативно нову технику за класификационе и регресионе проблеме [15] и подржава и линеарне и нелинеарне зависности.

Слика 4 [48] приказује једноставан модел технике потпорних вектора. Модел се састоји од два различита шаблона. Циљ методе потпорних вектора је да раздвоји та два шаблона. Модел садржи три линије. Линија  $w \cdot x - b = 0$  зове се маргина раздвајања. Линије  $w \cdot x - b = 1$  и  $w \cdot x - b = -1$  налазе се на свакој од страна маргине раздвајања. Ове три линије креирају такозвану хипер-раван, која раздваја дате шаблоне. Параметар  $w$  представља нормализован вектор хипер-равни. Свака тачка  $x$  је мулти-димензиони реални вектор, док  $b$  представља граничну вредност.

Шаблони који се налазе на ивицама хипер-равни зову се потпорни вектори. Вертикална дистанца између маргине раздвајања и ивица хипер-равни зове се маргина. Један од циљева методе потпорних вектора јесте да максимизује маргину ради боље класификације. Већа вредност маргине значи бољи процес класификације и самим тим минимизације појаве грешке у предвиђању.



Слика 4 – Модел технике модела потпорних вектора

Алгоритам који се налази у основи *SVM* користи се за класификацију различитих типова шаблона. Шаблиони којима управља овај алгоритам могу бити линеарни и нелинеарни. За линеарне шаблоне, основна идеја је конструкција идеалне хипер-равни, која ће се користити за класификацију линеарно раздвојивих шаблона. Оптимална хипер-раван је она која максимизује маргину, тј. дистанцу између хипер-равни и најближе тачке сваког од шаблона. Главни циљ методе потпорних вектора је да максимизује маргину тако да шаблони могу да буду адекватно класификовани.

За нелинеарни модел потпорних вектора, шаблони се мапирају у вишедимензионални простор, користећи тзв. *kernel* функцију. Постоје три типичне *kernel* функције: Гаусова, полиномска и линеарна [49].

### 2.2.4.3. Расплинуте временске серије

Према [50], расплинута логика представља генерализацију стандардне логике, према којој постоји концепт који може да има степен истине у распону од 0 до 1. Према стандардној логици, концепти могу да буду потпуно тачни или потпуно нетачни. Расплинута логика претпоставља могућност рационализације концепата који нису толико јасни.

Појам расплнутих временских серија ослања се на појам универзума дискурса. Универзум дискурса представља универзални скуп ентитета који су од интереса за прорачун, и које се могу рангирати према варијаблама ентитета. Идеја расплнутих временских серија састоји се у томе да се универзум дискурса подели на партиције, односно временске серије. За сваку временску серију учи се понашање, односно одређују се шаблони понашања за сваку временску серију. Ова правила означавају какве везе партиције граде међу собом током времена. Понашање партиција условљено

је креираним „речником“, где сваки израз описује стање у коме се налазе ентитети у партиципи. То стање произилази из саме расплнуте логике, према којој сваки ентитет може да има придружен степен истине за неку тврдњу. Креирани шаблони се на крају примењују на ентитете за које још не постоје измерене вредности и на тај начин се посматране вредности прогнозирају [51].

### **2.2.5. Избор узорка за методе машинског учења**

Проблем прогнозе потрошње ресурса линеарним методама састоји се у томе што утицај независних на зависне променљиве није подразумевано линеаран. Утицај посматраних фактора на потражњу мења се у зависности од времена у току дана, дана у недељи, доба у години, сличности метеоролошких услова са прогнозираним даном и од многих других услова. Неки од ових проблема могу се решити применом *ARIMA* модела. Њиме се могу пратити промене у временској серији у току дана (на пример на сатном, полусатном или 15-минутном нивоу). Међутим, нелинеарност може да се јавља у зависности потрошње од било ког фактора, и у том случају самосталне нелинеарне методе не поседују решење. Једно решење може да буде једноставно употреба нелинеарних метода за прогнозу. Нелинеарне методе машинског учења поседују неке могућности оптимизације узорка. Међутим, адекватнији избор узорка унапређује и нелинеарне методе, како у погледу тачности резултата прогнозе, тако и у погледу перформанси.

У инфраструктурним мрежама, оптимизација избора узорка најчешће се своди на селекцију дана из прошлости који су по својим карактеристикама довољно слични дану за који се прогнозира потрошња. Оваква оптимизација има своју употребу у прогнозама и најчешће се зове „селекција сличних дана“. Овај метод може значајно да унапреди прогнозу, како у линеарним, тако и у нелинеарним методама машинског учења. У линеарним методама, адекватан избор сличних дана може да приближи линеарности утицај фактора на потражњу. У нелинеарним методама, адекватан избор сличних дана може додатно да побољша тачност резултата и перформансе.

# 3. Унапређење метода прогнозе временских серија адаптивним моделом

У претходном поглављу размотрени су аспекти краткорочне прогнозе потражње за различитим врстама ресурса: електричном енергијом, водом и природним гасом. Очито је да све описане врсте прогнозе за потражњом ресурса имају заједничке карактеристике:

- прогноза потрошње, односно потражње за ресурсима, извршава се у оквиру временских серија, са резолуцијом која може да варира (15 минута, пола сата, сат итд.).
- постоји потреба за моделовањем зависности између фактора који утичу на потражњу и саме потражње.
- фактори који утичу на потражњу и који се користе у моделовању зависности имају различите нивое утицаја на потражњу. Тај утицај може да се мења током времена и у зависности од географског подручја за које се прогнозира потрошња.
- постоји могућност оптимизације узорка избором сличних дана.

У доступној литератури [52, 53, 54, 55, 56, 57, 58] широко је обрађен утицај појединачних фактора на потражњу за ресурсима. Такође, постоје модели који се баве факторима који утичу на избор сличних дана. Међутим, евидентно је да није у довољној мери обрађено питање оптимизације фактора који утичу на избор сличних дана. Није наглашено ни питање адаптивности избора узорка, односно, сличних дана, у односу на промењене временске и календарске околности. Коначно, не постоји јединствен модел прогнозе који се ослања на оптимизацију узорка, независно од алгорита који ће за прогнозу бити коришћен.

Утицај појединачних фактора на адекватан избор сличних дана може да се одреди експерименталним методама, сталним праћењем утицаја и упоређивањем резултата. Међутим, то је веома захтеван процес са становишта потребног времена. Поред тога, утицај фактора се мења током времена, тако да не постоји сигурност да ће једном одређене вредности остати оптималне током времена.

На основу свега наведеног, постављају се следеће дилеме:

- Да ли је могуће креирати јединствени модел прогнозе који ће бавити свим аспектима оптимизације узорка?
- Како се могу оптимизовати варијабле које се користе у избору сличних дана за методе прогнозе потражње за ресурсима у инфраструктурним мрежама?
- Да ли је могуће креирати оптимизацију избора сличних дана без губитка перформанси?
- Да ли се може обезбедити стална адаптивност избора сличних дана, у односу на променљиве околности током времена?

Овај рад се бави креирањем јединственог модела краткорочне прогнозе ресурса, који се базира на оптимизацији избора сличних дана, независно од алгоритма прогнозе. Модел треба да буде проширив, што значи да се слични дани бирају на основу различитих фактора, њиховим додавањем и уклањањем и променом утицаја сваког фактора.

Оптimalан скуп сличних дана који ће се користити у алгоритмима машинског учења представља такав скуп којим ће се максимизирати тачност прогнозе потрошње ресурса. Према томе, оптимизациони проблем је минимизовање грешке прогнозе алгоритмом машинског учења.

За оптimalан избор сличних дана могу да се користе технике којима се одређује вредност варијабле које утичу на минимизовање грешке у прогнози. Варијабле које се оптимизују могу да буду вредности фактора који директно утичу на оптимизациони проблем. Пример за овакве варијабле су вредности метеоролошких и календарских параметара у претрази сличних дана. У овом случају опсег вредности варијабле зависи од природних карактеристика феномена на које се варијабле односе. Са друге стране, варијабле које се оптимизују могу да буду изведене, односно дефинисане у алгоритмима. Пример за ово су варијабле које се односе на утицај појединачних феномена на коначну грешку у прогнози. У овом случају опсег вредности варијабле може да буде вештачки генерисан (на пример од 1 до 100).

Избор добре оптимизационе технике зависи од природе оптимизационог проблема. На пример, линеарно програмирање је оптимизациона техника која обухвата критеријумске функције које линеарно зависе од нумеричких променљивих. У овој техници скуп ограничења се изражава линеарним једнакостима [59, 60, 61]. У оптимизационом проблему избора сличних дана, критеријумска функција не зависи од променљивих линеарно. Предлог решења јесте креирање модела којим ће се применити метахеуристички метод за избор сличних дана. Под метахеуристичким методом подразумева се алгоритам за решавање комплексног оптимизационог проблема, претрагом простора могућих комбинација решења, и минимизацијом дефинисане функције трошка [62]. У овом раду, функција трошка подразумева грешку приликом прогнозе – однос прогнозиране и остварене вредности потрошње. Комбинација решења односи се на комбинацију вредности фактора који утичу на оптimalан избор сличних

дана. Дакле, метахеуристичком методом долази се до комбинације вредности варијабли које утичу на избор сличних дана, у циљу минимизовања грешке у прогнози.

Препознавање оптималних вредности за факторе који утичу на оптималан избор сличних дана не подразумева нужно да се користе исти фактори који се користе и у случају метода машинског учења за прогнозирање потрошње. Предложеним моделом уводи се могућност уочавања утицаја фактора који су се већ спомињали у досадашњим истраживањима и литератури, али и увођења нових фактора.

Међу бројним метахеуристичким методама, потребно је изабрати ону која ће давати оптималне резултате у погледу могућности техничке имплементације, тачности прогнозе и перформанси.

Предложени модел треба да испуњава услов адаптивности у односу на промењене околности. Изабрани метахеуристички метод треба да извршава оптимизацију вредности фактора од утицаја током времена. Такође, предложени модел треба да поседује могућност адаптивности и у смислу могућности додавања нових фактора од утицаја.

Оптимизација у моделу треба да обухвати све дане у години, па и оне који се разликују од осталих по очекиваним дневним профилима потрошње. Ово су такозвани специјални дани, и модел треба да има могућност ефикасне прогнозе за овакве дане.

На основу свега наведеног, основна хипотеза у дисертацији може да се сведе на следеће ставке:

- Предложени адаптивни модел омогућиће оптимизацију узорка за прогнозу временских серија коришћењем система машинског учења, тако да се побољшају крајњи резултати прогнозе.
- Предложени адаптивни модел биће конципиран тако да буде отворен за коришћење постојећих фактора који утичу на избор сличних дана, али и за додавање нових фактора.
- Оптимизација избора узорка у моделу биће успешно извршена коришћењем метахеуристичког метода.
- Изабрани метахеуристички метод може се адаптирати у погледу оптималне имплементације, перформанси и резултата прогнозирања временских серија у инфраструктурним мрежама.
- Предложени модел омогућиће препознавање и оптимизацију утицаја фактора за избор сличних дана.
- Предложени модел омогућиће адаптивност, односно стално прилагођавање промењеним околностима током времена.
- Предложени модел поседује могућност препознавања специјалних дана и прилагођавања прогнозе потрошње оваквим данима.



Избор метахеуристичког метода извршиће се кроз анализу постојећих решења и могућност њихове примене у моделу.

Предложени модел биће описан кроз архитектуру, дизајн и имплементацију софтверске апликације за краткорочну прогнозу потрошње ресурса. Верификација модела извршиће се над прогнозираним и оствареним вредностима потрошње електричне енергије у електроенергетском преносном систему за подручје Србије без територије Косова и Метохије „Електромрежа Србије“, као и над прогнозираним и оствареним вредностима потрошње природног гаса у дистрибутивном систему за један регион у Холандији. Подаци у овим примерима су скалирани скривеним вредностима, ради заштите осетљивих података.

## **4. Преглед поступака прогнозе временских серија**

Прогнозирање временских серија је процес предвиђања будућности уз помоћ историјских података. Претпоставка предвиђања временских серија је да ће се информације понављати у будућности [63]. Прогнозирање потрошње ресурса у временским серијама у већем делу литературе представљено је независно за електричну енергију [64, 65, 66, 67], природни гас [68, 69, 70] и воду [71, 72]. Технике избора узорка у прогнозама и технике које користе метахеуристичку оптимизацију у доступној литератури коришћене су најчешће у комбинацији са другим техникама прогнозе. У наставку ће бити наведени постојећи поступци избора сличних дана у прогнози, као и постојеће технике метахеуристичке оптимизације.

### **4.1. Избор узорка у прогнози потрошње ресурса**

Идеја оптимизације узорка за потребе ефикасне прогнозе потрошње није нова. У доступној литератури постоји широк спектар модела који се користе ради што репрезентативнијег избора узорка који ће се користити као улаз у прорачуну прогнозе. Већина ових модела презентована је за потребе краткорочне прогнозе за потражњом појединачних ресурса, а најчешће се ради о прогнози потрошње електричне енергије.

За употребу метахеуристичких метода у прогнози такође постоји мноштво модела. Међутим, евидентно је да се ове методе у најчешћем случају користе за потребе оптимизације саме прогнозе, док недостају модели који се метахеуристичким методама користе ради избора сличних дана.

#### **4.1.1. Избор узорка у краткорочној потражњи потрошње електричне енергије**

Избор узорка у краткорочној прогнози потрошње електричне енергије најчешће се своди на неки од модела избора сличних дана. Модели у доступним радовима могу се грубо поделити на две групе:

1. Модели са једноставном селекцијом сличних дана, код којих се фокус ставља на поступак прорачуна прогнозе, док је избор сличних дана помоћно средство за прорачун прогнозе

2. Модели са комплексном селекцијом сличних дана, где се предлажу иновативни методи којима се избор сличних дана поставља као есенцијалан услов ефикасне прогнозе

#### **4.1.1.1. Модели са једноставном селекцијом сличних дана**

Често коришћен принцип моделовања селекције сличних дана подразумева фокусирање аутора на методологију саме прогнозе, а слични дани су коришћени на једноставан начин, без намере да се дубље уђе у анализу селекције сличних дана [26]. Коришћени фактори у оваквом приступу подразумевају дан у недељи, сат у дану и основне метеоролошке и временске податке, што се најчешће односи на температуру. Оваквим приступом могу да се добију задовољавајући резултати, пре свега у случајевима када изабрани фактори имају очигледан примат у односу на друге.

Прогноза таласастом неуронском мрежом базираном на сличним данима представљена је у раду [73], у коме аутори користе субјективан осећај температуре као фактор за селекцију сличних дана.

У радовима [74, 75, 76], фактори који се користе за детекцију сличних дана јесу девијација између оптерећења у прошлом прогнозираном дану и историјским данима, девијација температуре и девијација облика криве оптерећења између прогнозираног дана из прошлости и историјских дана. Селекција сличних дана је базирана на Еуклидској норми, а тежински фактори се користе за процену сличности између прогнозираног дана и дана из историјских података. Сваки фактор има придружени тежински коефицијент, који се одређује регресионим моделом и методом најмањих квадрата, коришћењем историјских података о температури и оптерећењу.

У раду [77], аутори су предложили минималну температуру, максималну температуру и тип дана као факторе за селекцију сличних дана.

У раду [78], аутори су користили влажност ваздуха и субјективни осећај хладноће у сваком сату да би одредили сличност дана.

Хибридни интелигентни алгоритам је представљен у раду [79]. У овом чланку, за прогнозирање оптерећења користи се комбинација таласне трансформације, неуронске мреже, методе сличних сати и адаптивне расплинуте логике. Што се тиче методе сличних сати у овом раду, разлика између прогнозираног оптерећења и оптерећења у сличном сату добија се поређењем дана у недељи, временских услова и годишњег раста оптерећења.

У раду [80], аутори предлажу адаптивни алгоритам учења у неуронској мрежи базиран на сличности дана, уважавајући рапидне промене температуре у прогнозираном дану. За процену сличности између прогнозираног и дана из историјских података је коришћена Еуклидска норма са тежинским факторима, уз употребу максималне и минималне температуре и дана у недељи.

Рад [81], предлаже метод за прогнозу оптерећења током празника. Овај метод је базиран на сличним данима и експертном систему распинуте логике. Главна идеја је селекција три слична профила оптерећења која кореспондирају са три есенцијалне карактеристике оптерећења, а то су: средња вредност 24-часовне криве оптерећења, стандардна девијација и сличност између одговарајућих криви оптерећења.

У раду [82] презентован је *PSF-NN* (Сличност секвенци шаблона – Неуронске мреже, енгл. *Pattern Sequence Similarity — Neural Networks*), приступ за прогнозу временских серија, који комбинује сличност секвенци шаблона са неуронским мрежама. За сличност секвенци, овај метод користи информације о оптерећењу у претходном дану, претходној недељи и заједно претходном дану и недељи.

У радовима [83, 84], аутори су користили Еуклидску норму са тежинским факторима за евалуацију између прогнозираног и историјског дана. Еуклидска норма користи максималну и минималну температуру, влажност ваздуха и дан у недељи.

У раду [85], коришћена је имплементација неуронске мреже *Deep Belief Network* за краткорочну прогнозу оптерећења, унапређена са аспекте оптимизације података, модела и перформанси. Оптимизација података састоји се у успостављању Хенкелове матрице електричне енергије и корелационе анализе. Овај чланак предлаже метахеуристички метод за оптимизацију припреме података.

#### **4.1.1.2. Модели са комплексном селекцијом сличних дана**

У другој групи радова аутори се фокусирају на унапређење селекције сличних дана, дајући овом процесу већи значај у процесу прогнозе.

У чланку [86], у сврху одређивања коефицијента сличности, аутори користе карактеристике као што су тип дана, дан у недељи, максимална температура, минимална температура и датумска разлика. Кроз коришћен метод уводи се прорачун тежинских коефицијената за варијабле које одређују сличност дана.

Међусобна сличност оптерећења, представљена у раду [87], омогућава представљање карактеристика скупа података оптерећења, као што је правилност криве. За дугорочне податке, међусобна сличност добија се методом таласне анализе (енгл. *Wavelet Analysis*).

У радовима [88, 89], аутор описује принципе метода сличности засноване на шаблонима. Шаблон чува информације о облику криве дневног оптерећења. Поред тога, овај чланак уводи модел који садржи филтер девијације оптерећења, који елиминише девијантне дане.

Температура и датумска близина дана користе се за одређивање индекса сличности дана у раду [90].

Рад [91] уводи тежинске факторе ради препознавања одговарајућих атрибута (температура, тип дана, брзина ветра). За одређивање тежинских фактора користи се *Xgboost* алгоритам.

У раду [92], аутори су изабрали 22 фактора који утичу на оптерећење, и увели су корелациону функцију да би истражили повезаност историјског оптерећења са тренутним оптерећењем. Предложени су следећи скупови фактора: историјско оптерећење, температура, временски подаци и тип дана.

У раду [93], аутори користе међуиндустријски стандардни процес (енгл. *Cross-Industry Standard Process – CRISP*) за истраживање података, да би пронашли узрочно-последичну везу у смислу удаљености времена међу улазним варијаблама, да би се пронашла одговарајућа корелација између улазних и излазних варијабли. У ту сврху, користе се подаци о енергији и времену. У кораку припреме, користи се прочишћавање података, укључујући попуњавање недостајућих и некомплетних вредности одговарајућим вредностима, уочавање и уклањање вредности које одступају, уклањање дубликата и вођење рачуна о неодговарајућим типовима података. Да би се пронашао тренд и правилности за потрошњу електричне енергије, аутори су препознали шаблоне оптерећења на недељном, месечном и сезонском нивоу. Након припреме података и препознавања шаблона, прогноза оптерећења се извршава увођењем приступа заснованог на кластерима, користећи неуронску мрежу и метод потпорних вектора. Процес истраживања података у овом раду садржи кораке који се могу користити у селекцији сличних дана, као што су прочишћавање података и препознавање шаблона оптерећења.

У раду [94], аутори предлажу *STLF* модел за енергетске трансформаторе који се састоји из два корака. Први корак предвиђа агрегирано оптерећење на нивоу трансформаторске станице, коришћењем историјског оптерећења и временских и календарских података као улазних података. У другом кораку, дефинише се фактор дистрибуције оптерећења (енгл. *Load Distribution Factor – LDF*) као однос оптерећења трансформатора и оптерећења трансформаторске станица. Проблем неочекиваних промена у радним условима и нерегуларних података о оптерећењу решава се функцијом нелинеарне регресије, која се примењује на предвиђено оптерећење паралелних трансформатора, уместо коришћењем неуронске мреже.

У раду [95], аутори предлажу метод заснован на сличним данима, *PSO* и *Deep Belief Network – DBN*. Сличност дана се израчунава у складу са датумском удаљеношћу, типом дана и метеоролошким карактеристикама. Да би се превазишли недостаци *DBN*, који лако може да падне у локални минимум у процесу учења и тренинга, користи се *PSO* алгоритам за оптимизацију тежинских фактора мреже.

У раду [96] аутори су презентовали нови систем кооперативне сарадње, који усавршава прогнозе засноване на вероватноћама индивидуалних оптерећења. За процедуру

усавршавања користи се мултиваријабилна квантилна регресија. У процедури пре-процесирања података, оптерећења су груписана према географским подручјима, што уобичајено води до повећања перформанси система за прогнозу.

#### **4.1.2. Избор узорка у прогнози потражње за водом**

У већини радова који се баве прогнозом потражње за водом мала пажња је посвећена избору узорка. Најчешће се избор узорка своди на једноставну селекцију за претходни временски период, док мањи број радова више пажње посвећује претпроцесирању расположивих података. У неким случајевима постоји проблем и са доступношћу одговарајућих података. Тако, у раду [97] у коме се аутори баве естимацијом детерминанти које утичу на потражњу становништва за водом на територији Италије, наглашава се да подаци о потрошњи на нивоу домаћинства нису доступни иако би били пожељни за естимацију детерминанти који утичу на потражњу за водом.

У раду [17] аутори се баве прогнозом потражње за водом за период од једног сата, 8 сати, 24 сата и једне недеље унапред, коришћењем надгледаних и ненадгледаних техника машинског учења, на територији југозападни Онтарио, Канада. Као улазни подаци за поступак прогнозе користе се за последња четири месеца 2017. године.

У раду [98], аутори се баве прогнозом потражње за водом коришћењем техника машинског учења. Као критеријуми за избор улазних података наглашени су сезоналност и спорадични сезонски догађаји као што су празници. Сезоналност је имплементирана идентификацијом различитих шаблона у скупу података. На основу шаблона подаци су кластерисани и прослеђени алгоритму за прогнозу.

Рад [99] бави се техником шаблона сличности у прогнози потражње за водом. Подаци су подељени на основу локације података, а затим претпроцесирани чишћењем и поправком података. Чишћење и поправка података ослањају се на рад [100], и заснивају се на декомпозицији сезонских временских серија на компоненте које се односе на тренд, сезоналност и на све остале податке. Шапони података се креирају на основу дана у недељи.

Рад [54] предлаже нови приступ прогнози потражње за водом, базиран на *Random Forests* методи. Ова метода припада моделима вођеним подацима и састоји се од стабла одлучивања које се може користити за класификацију података. У овом раду подаци се деле по потрошачким групама, а након тога и на основу фактора као што су дужина сунчаних дана, радијација, кишовитост, влажност и температура ваздуха.

У раду [101], урађена је компарација класичних и адаптираних алгоритама за прогнозу потражње за водом. У компарацију су укључени методи *ARIMA*, метода потпорних вектора и *Random Forests*. Претпроцесирање података подразумева одстрањивање података са грешком и података са вредностима које одступају од вредности груписаних по кварталима.

### 4.1.3. Избор узорка у прогнози потрошње природног гаса

Специфичност избора узорка за краткорочну прогнозу потрошње природног гаса представља чињеница да се потрошња овог ресурса динамично мења у смислу технологије и обима коришћења, тако да постоји могућност да старији историјски подаци не одговарају новијим трендовима потрошње. Ове промене могу бити дугорочне, у смислу тренда развоја инфраструктуре и потрошачких навика. Међутим промене могу да буду и тренутне. На пример, у раду [9] напомиње се да за разматрани сет података, постоје значајне варијације у потрошњи природног гаса у априлу и октобру, без видљивог разлога. Пошто се ови месеци налазе на почетку и на крају грејне сезоне, претпоставка је да се разлози за повећану потрошњу налазе у тестирању великих потрошача и нерегуларних стања мреже.

У раду [102] предлаже се детрендизација података (енгл. *detrending*) у циљу побољшања краткорочне прогнозе потрошње природног гаса. Детрендизација података прилагођава историјске серије података тако да се они приближавају стационарној бази потрошача. То практично значи да се тренутни трендови понашања потрошње примењују на историјске податке, и тиме се смањује могућност грешке у прогнози.

Рад [8], који се бави краткорочном прогнозом потрошње природног гаса неуронском мрежом, као улаз у неуронску мрежу користи податке из 62 оперативна подручја дистрибутивних компанија са подручја Сједињених Америчких Држава. Нормализација података се извршава методом предложеном у раду [102].

Још једна специфичност избора узорка за краткорочну прогнозу потрошње природног гаса представља недостатак валидних података који би се могли користити у циљу истраживања. Овом проблематиком се бави рад [34] и у њему се предлаже сет података за шест узастопних година, који садржи податке о потрошњи, као и додатне варијабле које су се показале као корисне у процесу прогнозе потрошње. Ове варијабле обухватају метеоролошке податке и податке о празницима, али и цену гаса која је такође значајан фактор у прогнози потрошње.

## 4.2. Метакхеуристички методи

Метакхеуристички методи могу међу собом значајно да се разликују према томе на којим основама су успостављени. Неке метакхеуристички методи опонашају процесе који наизглед немају директне везе са оптимизацијом, као што су природна еволуција, или понашање група животиња [103].

### 4.2.1. Мрављи алгоритам

Мрављи алгоритам (енгл. *Ant Colony Optimization – ACO*) инспирисан је понашањем тражења хране неких врста мрва. Ови мрави складиште феромон на земљи у циљу означавања омиљених путева којима треба да се крећу остали чланови колоније.

Мрављи алгоритам користи сличан механизам за решавање оптимизационих проблема [104].

Мрављи алгоритам је итеративан. У свакој итерацији, узима се у обзир одређени број вештачких мрава. Сваки од њих креира решење крећући се у простору могућих решења.

Оптимизациона техника мрављег алгоритма полази од броја вештачких мрава који креирају решење за оптимизациони проблем који се разматра, и који размењују информације о квалитету решења. Размена информације извршава се преко комуникационе шеме, која подсећа на шему комуникације мрава у природи. У решавању оптимизационог проблема, вредност феромона је повезана са могућим решењем, тј. доделом вредности оптимизованих варијабли.

У литератури је представљено неколико решења мрављег алгоритма. Према алгоритму *систем мрава* (енгл. *Ant System – AS*), у свакој итерацији вредности феромона се ажурирају за све мраве који су креирали решење у самој итерацији. [105]. *Макс-мин систем мрава* (енгл. *Max-Min Ant System*) представља унапређење оригиналног система мрава. За овај систем карактеристично је да само најбољи мрав ажурира вредности феромона [106]. *Систем колонија мрава* (енгл. *Ant Colony System – ACS*) уводи концепт локалног ажурирања феромона, са циљем спречавања мрава да креирају идентична решења [104, 107].

#### 4.2.2. Оптимизација ројева честица

Оптимизација ројева честица (енгл. *Particle Swarm Optimization – PSO*) инспирисана је јатима птица и понашањима јата риба. Концепт оптимизације ројева честица први пут је описан у раду [108]. Предложени алгоритам полази од претпоставке да свака јединка у јату има специфичну локацију и брзину, и следи следеће циљеве: 1) **Поравнавање**: свака јединка тежи да се усклађује са другим јединкама регулишући своју брзину. 2) **Кохезија**: Појединачне јединке управљају својим положајем тежећи центру јата. 3) **Одвојеност**: Свака јединка покушава да избегне да се судари са најближим суседом тако што држи разумну удаљеност. Ова удаљеност је довољна да избегне судар, а да се истовремено избегне излазак из јата [109].

Имплементација овог алгоритма почиње креирањем иницијалног скупа локација (енгл. *locations*) и брзина (енгл. *velocities*) за сваку јединку. Вредности локације и брзине се ажурирају релативно једноставном формулом која одражава особине поравнавања, кохезије и одвојености. То резултира у симулацији јединки у јату над дефинисаним подручјем. Јато може да претражи подручје на ефикасан и кохезиван начин.

Концепт оптимизације ројева честица може да се употреби у решавању комплексних оптимизационих проблема, где постоји већи број локалних оптимума и потреба да се пронађе глобални оптимум. Алгоритам почиње са скупом кандидата за решење, односно локацијама за сваку јединку. У реалном алгоритму то су вредности параметара



који се оптимизују. Ови кандидати се ажурирају и крећу се претражујући простор на случајан начин, и трагајући за унапређеним решењем, односно вредношћу критеријумске функције. Ажурирање се врши изменом брзине за сваку јединку, чиме се ажурира и локација сваке јединке.

### 4.2.3. Генетски алгоритам

Генетски алгоритам (енгл. *Genetic Algorithm – GA*) је оптимизациона техника базирана на принципима генетске и природне селекције. Генетски алгоритам омогућава популацији која се састоји од великог броја јединки да еволуира под специфичним правилима селекције, до стања које максимизује еволутивну способност, односно минимизује функцију трошка [110].

Кључни елементи генетског алгоритма су концепти гена, хромозома, популације, селекције, укрштања, мутације и функције трошка [111].

Сваки хромозом представља јединку која се састоји од гена, односно варијабли чије вредности је потребно оптимизовати. Популација представља скуп који се састоји од фиксног броја хромозома. Гени сваког хромозома представљају улаз у критеријумску функцију, чију вредност је потребно минимизовати.

Иницијална популација се креира од хромозома чији су гени креирани по случајном принципу. Хромозоми се рангирају на основу вредности критеријумске функције, тако што се хромозоми са нижом вредношћу критеријумске функције („бољи хромозоми“) налазе на врху листе. Одређени број хромозома са вишом вредношћу критеријумске функције („лошији“ хромозоми) се одбацују, док се преостали хромозоми селекутују, укрштају се међу собом и креирају нове хромозоме. Хромозоми „родитељи“ и хромозоми „наследници“ представљају нову генерацију популације хромозома [110].

Селекција представља избор парова хромозома који ће се укрстити међу собом. Постоји више начина за селекцију хромозома. Један од најуспешнијих је пондерисана случајна селекција. Према овом концепту, хромозоми се селекутују на основу свог ранга, који се креира на основу вредности функције трошка [112].

Укрштање представља креирање нових потомака хромозома од упарених хромозома – родитеља, при чему је генетски материјал близак генетском материјалу хромозома – родитеља. У случају оптимизовања генетских варијабли које су континуалне нумеричке вредности, нова вредност гена ће представљати нову континуалну нумеричку вредност чија вредност ће се налазити између вредности генетских варијабли које припадају хромозомима – родитељима [111].

Мутација се генетском алгоритму користи ради одржавања генетске разноврсности популације. У конкретној имплементацији, мутације служе за удаљавање популације од локалног оптимума критеријумске функције. Мутација подразумева промену вредности на изабраним генима изабраних хромозома.

#### **4.2.4. Метакхеуристички методи у прогнозирању краткорочне прогнозе потрошње електричне енергије**

У раду [113] аутори користе расплинуту логику у циљу прорачуна прогнозе краткорочне потрошње електричне енергије. За генерисање скупа правила у алгоритму расплинуте логике, паралелно се користе генетски алгоритам и мрављи алгоритам. Генетски алгоритам је коришћен за потребе глобалне претраге правила, док је мрављи алгоритам коришћен за локалну претрагу.

У раду [114] користи се мрављи алгоритам ради унапређења кластерисања историјских података о оптерећењу. Кластерисани подаци се користе као улаз у алгоритам методе потопорних вектора. Оптимизација кластерисања односи се на минимизацију оптимизационих центара и степена учествовања, што значи да се избацују узорци који сувише одступају од осталих. На тај начин се побољшавају перформансе извршавања алгоритма потпорних вектора.

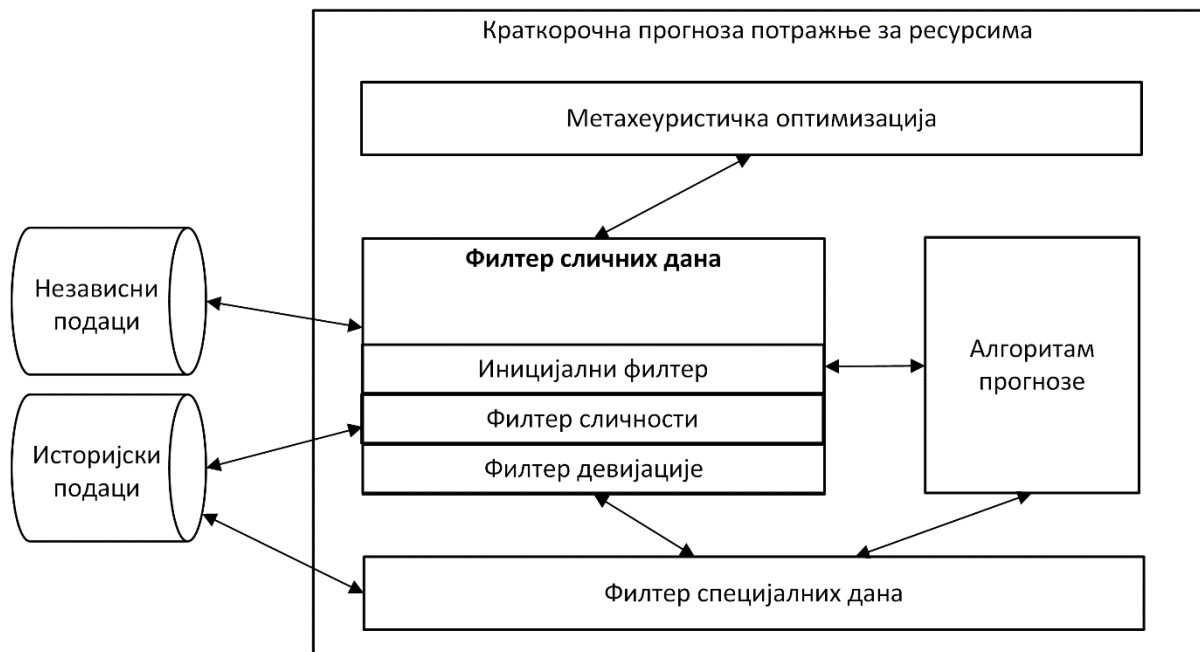
Модели представљени у радовима [115, 116, 117] за прогнозу користе неуронску мрежу. Оптимизација ројева честица се користи у сврху оптимизације параметара као што су број узорака за тренинг, број скривених неурона, тип активационе функције, број епоха и максимално време извршавања.

У раду [118], краткорочна прогноза потрошње електричне енергије прорачунава се неуронском мрежом подржаном генетским алгоритмом. Генетским алгоритмом се оптимизују тежински фактори неуронске мреже, кроз операције селекције, укрштања и мутације.

У раду [119] предлаже се интеграција генетског алгоритма са неуронском мрежом ради краткорочне прогнозе потрошње електричне енергије, тако што генетски алгоритам оптимизује топологију неуронске мреже. Генетским алгоритмом одређује се број улазних неурона и број скривених слојева неуронске мреже.

# 5. Архитектура адаптивног модела прогнозе

Глобална архитектура адаптивног модела прогнозе временских серија приказана је на слици 5. Улазни подаци за процес прогнозе су историјски подаци о потражњи за ресурсима, историјски независни подаци и независни подаци за период прорачуна. Ови подаци се прослеђују филтеру сличних дана. Филтер сличних дана састоји се од иницијалног филтера, филтера сличности и филтера девијације. Оптимизација филтера сличних дана извршава се кроз метахеуристички метод оптимизације. Излазни подаци из филтера сличних дана су улазни подаци у алгоритам прогнозе. Алгоритам прогнозе ће у коначној архитектури бити имплементиран кроз методологију машинског учења.



Слика 5 – Глобална архитектура адаптивног модела прогнозе

Очигледно је да је приказани модел представљен на апстрактном нивоу, још увек без конкретног метода метахеуристичке оптимизације, и без конкретног метода машинског учења. Главна идеја модела јесте оптимизација избора узорка. Ова оптимизација потпомогнута је метахеуристичким методом, а корисник методе је алгоритам прогнозе. Без обзира који метахеуристички метод и алгоритам машинског учења да се изабере, модел ће имати позитиван утицај на резултате прогнозе и перформансе процеса. Суштина модела на овом нивоу је у филтеру сличних дана, и он ће у наставку бити

детаљно описан. Са филтером сличних дана повезан је филтер специјалних дана, који служи за управљање прогнозом дана за које се унапред зна да имају специфичности које ванредно утичу на профил потрошње.

## 5.1. Филтер сличних дана

Сличност дана који се прогнозира и дана из историјске базе података огледа се у процени блискости две криве потражње за ресурсима у току дана. Циљ филтера сличних дана је процена сличности и издвајање дана за које је процењено да постоји довољан степен сличности за оптималну прогнозу.

Као улаз за филтер сличних дана користе се историјски подаци о потражњи за ресурсима, историјски независни подаци и независни подаци за период прогнозе. Овај филтер се састоји од три секвенцијална филтера за селекцију сличних дана:

1. Иницијални филтер се користи за иницијалну селекцију сличних дана.
2. Филтер сличности одређује коефицијент сличности за сваки историјски дан. Сличност у овом контексту подразумева сличност сваког од дана селектованих у претходном кораку са даном који се прогнозира. Нижи коефицијент сличности значи већу сличност.
3. Филтер девијације се примењује на рангиране дане након филтера сличности. Циљ овог филтера је да се одреде најбољи кандидати за примену у алгоритму прогнозе, тако што ће бити искључени дани чија потражња за ресурсима значајно одступа од потражње осталих дана. Излаз из овог филтера је финални скуп дана који ће се користити као улаз у алгоритам машинског учења за прогнозу потражње за ресурсима.

### 5.1.1. Иницијални филтер

Иницијални филтер има сврху иницијалног издвајања дана за које ће се одређивати сличност у односу на прогнозирани дан и одбацивање оних дана за које се утврди да немају основне карактеристике сличности са прогнозираним даном. Овакво иницијално издвајање служи пре свега за побољшање перформанси, пошто се сужава број дана који ће се даље обрађивати.

Иницијални филтер идентификује скуп историјских дана код којих сваки дан има исте основне карактеристике као и дан који се прогнозира. Ове основне карактеристике односе се на календарске, метеоролошке и техничке податке: тип дана, грејна сезона и температурни опсег.

Начин на који се одређује тип дана је конфигурабилан и може да се одређује на следеће начине:

- Сваки тип дана је један дан у недељи. Селектују се дани који припадају истом дану у недељи као и дан за који се потрошња прогнозира. Овај начин се

примењује у случајевима када очигледно сваки дан у недељи значи и посебан оквирни шаблон потражње за ресурсима, и када у историјској бази података постоји довољно података.

- Посебни типови дана су викенд и радни дани. Овај начин се примењује када су оквирни шаблони потражње такви да се разликују између радних дана и викенда.
- Посебни типови дана су: субота, недеља, понедељак и остали дани. Овај начин се примењује када се оквирни шаблони потражње разликују током викенда за суботу и недељу и када потражња понедељком има шаблон који је под утицајем викенда.

Грејна сезона представља логички (енгл. *boolean*) податак који означава да ли је у току грејна сезона или не, односно да ли су укључени системи даљинског грејања на већем делу територије за који се потрошња прогнозира. Овај податак може пресудно да утиче на прогнозу, пре свега када се ради о прогнози потрошње гаса или електричне енергије. Ако се ради о прогнози потрошње електричне енергије, грејна сезона која је у току може да утиче на мању потрошњу електричне енергије која се користи за грејање. Ако се ради о потрошњи гаса, грејна сезона која је у току значи већу потрошњу гаса.

Ако се ради о прогнози потрошње воде, ова карактеристика у иницијалном филтеру може да буде искључена.

Температура представља један од главних параметара за одређивање сличних дана независно од тога о ком ресурсу се ради. Карактеристика температурног опсега се користи због претпоставке да нема смисла поредити дане који имају велику разлику у температури. Број степени којим се одређује температурни опсег је конфигурабилан, али не може бити мањи од 5 степени целзијуса. Искусствено је закључено да би у случају мањег опсега температуре иницијални скуп дана могао да буде недовољан за ефикасну примену наредних филтера. Број степени постављен на 5 значио би да се у иницијалном филтеру бирају дани који имају просечну дневну температуру  $\pm 5$  степени у односу на прогнозирани дан.

### **5.1.2. Филтер сличности**

Филтер сличности рангира селековане сличне дане на основу коефицијента сличности. Коефицијент сличности креира се на основу израчунатих подкоефицијената. Подкоефицијенти сличности могу да се мењају, проширују, да се додају нови или да се избацују постојећи. За потребе прогнозе потражње за електричном енергијом, водом и природним гасом идентификована су четири основна подкоефицијента:

1. Коефицијент временских прилика.
2. Инерциони коефицијент.
3. Коефицијент трајања дневне светлости.
4. Коефицијент близине дана.

На основу четири предложена подкоефицијента формираће се коначни коефицијент сличности, на основу кога ће се дани сортирати и на основу кога ће бити изабрани дани који ће служити као улаз у алгоритам прогнозе.

### 5.1.2.1. Коефицијент временских прилика

Коефицијент временских прилика ( $sc_w$ ) се израчунава за изабране временске факторе (температура, брзина ветра, облачност, влажност ваздуха, падавине и други), користећи следећу формулу:

$$sc_w = \frac{\sum_{j=1}^{nfw} wc_j pc_j \sqrt{\frac{\sum_{i=1}^{24} (w_{di} - w_{(d-n)_i})^2}{24}}}{\sum_{j=1}^{nfw} wc_j}, \quad (7)$$

где је:

$nfw$  – број фактора временских прилика,

$i$  – сат у дану,

$w_{di}$  – вредност фактора у дану  $d$  који се прогнозира, у сату  $i$ ,

$w_{(d-n)_i}$  – вредност фактора за историјски, потенцијално сличан дан  $d-n$  у сату  $i$ ,

$wc_j$  – индивидуални тежински коефицијент за  $j$ -ти фактор временских прилика.

Представља меру учешћа  $j$ -тог фактора временских прилика у коефицијенту  $sc_w$ ,

$pc_j$  – партиципациони коефицијент за  $j$ -ти фактор временских прилика.

Формула представља пондерисану квадратну девијацију између прогнозираних и остварених временских параметара. Коефицијент сличности  $sc_w$  означава у којој мери су слични дан  $d$  који се прогнозира и дан  $d-n$  из историјске базе података, када се узму у обзир разлике у вредностима фактора временских прилика (температура, брзина ветра, влажност ваздуха и др.) и процењени утицај сваког од фактора на сличност дана. Утицај сваког фактора на сличност дана означена је тежинским коефицијентом  $wc_j$ . Одређивање тежинских коефицијената за сваки од фактора временских прилика представља један од кључних изазова за квалитетну имплементацију приказаног модела, и одређиваће се метахеуристичком оптимизацијом.

Сумирањем вредности тежинских коефицијената у делиоцу предложене формуле одржава се упросечавање утицаја појединачних фактора, пондерисано тежинским коефицијентима.

Коефицијент сличности се односи на цео дан и зато се у предложеној формули израчунава сатна разлика у вредностима фактора временских прилика.

#### 5.1.2.1.1. Одређивање партиципационих коефицијената

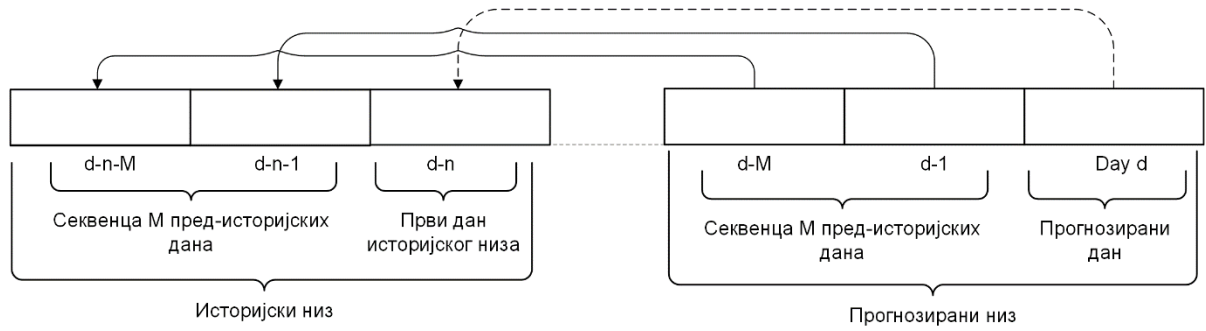
Партиципациони коефицијент  $pc_j$  омогућава сумаризацију различитих фактора временских прилика. Овај фактор је фиксна вредност за сваки фактор временских

прилика. Његов циљ није утицај на сличност дана или на тачност прогнозе, него сумаризација вредности различитих врста фактора, уз интуитивност анализе додељених коефицијената сличности. Овај коефицијент се одређује једнократно. Да би се одредио партиципациони коефицијент, бира се један референтни фактор за који ће партиципациони коефицијент имати вредност 1, док се сви остали партиципациони коефицијенти одређују на основу партиципационог коефицијента референтног фактора. У представљеном моделу референтни фактор је температура, и он има партиципациони коефицијент 1. После детаљне анализе података, партиципациони коефицијенти за остале факторе одређени су на следећи начин, на примеру партиципационог коефицијента за ветар: промена температуре за вредност од  $1\text{ }^{\circ}\text{C}$ , када се остали фактори занемаре, изазива просечну промену вредности потрошње ресурса за вредност  $A1$ . Промена брзине ветра за  $1\text{ m/s}$  изазива промену потрошње за вредност  $A2$ , када се сви остали фактори занемаре. Партиципациони коефицијент за брзину ветра ће, према томе, бити  $A2/A1$ . Вредност партиципационог коефицијента зависи од многих фактора (време дана, период у години итд.) и може се користити у циљу сумаризације вредности и интуитивне употребе коефицијената.

### 5.1.2.2. Инерциони коефицијент

Инерциони коефицијент односи се на концепт поређења низова дана. Низ једног дана обухвата посматрани дан и дане који претходе посматраном дану. Дане који претходе посматраном дану називамо „предисторија“ дана. Идеја инерционог коефицијента је да се сличност два дана повећава ако је и предисторија два дана такође слична. Употреба предисторије у одређивању сличности може у крајњој инстанци да повећа тачност прогнозе.

Поређење историјског низа дана ( $HA$ ) и прогнозираног низа дана ( $FA$ ) приказано је на слици 6.  $FA$  садржи прогнозирани дан  $d$  као први дан овог низа и секвенцу од  $M$  дана пре прогнозираног дана (предисторију дана  $d$ ). Типична вредност за  $M$  је 2, али генерално може да има нижу или вишу вредност. Оптимална вредност за  $M$  одређује се метахеуристичком оптимизацијом. Историјски низ  $HA$  садржи потенцијално сличан дан  $d-n$  из историјске базе података као први члан низа и секвенцу од  $M$  претходних дана (предисторију дана  $d-n$ ). Оба низа имају укупан број дана  $M+1$ .



Слика 6 – Прогнозирани и историјски низ дана

Инерциони коефицијент означава меру сличности између секвенце  $M$  предисторијских дана прогнозираног низа и секвенце  $M$  предисторијских дана историјског низа. Поређење се извршава за сваки дан предисторије историјског низа и прогнозираног низа. Инерциони коефицијент састоји се од коефицијента инерције потрошње, коефицијента термалне инерције и коефицијента инерције последњег сата.

#### 5.1.2.2.1. Коефицијент инерције потрошње

Коефицијент инерције потрошње означава меру сличности потрошње ресурса између секвенце  $M$  предисторијских дана прогнозираног низа и секвенце  $M$  предисторијских дана историјског низа:

$$sc_{cin} = \frac{\sum_{j=1}^M |avgC_{d-j} - avgC_{d-n-j}| ddwc^{j-1}}{\sum_{j=1}^M ddwc^{j-1}} pc_{cin} \quad (8)$$

где је:

$sc_{cin}$  – коефицијент инерције потрошње за дан  $d$ ,

$d$  – дан за који се прорачунава коефицијент,

$M$  – број предисторијских дана,

$n$  – број дана између прогнозираног дана  $d$  и сличног дана из историјског низа  $HA$ ,

$avgC_{d-j}$  – просечна сатна потрошња у  $d-j$ -ом дану из предисторијске секвенце прогнозираног низа дана;  $j=1, \dots, M$ ,

$avgC_{d-n-j}$  – просечна сатна потрошња у  $d-n-j$ -ом дану из предисторијске секвенце историјског низа дана,

$ddwc$  – дневни умањујући коефицијент за сваки од  $M$  дана (0-1),

$pc_{cin}$  – партиципациони коефицијент за инерцију потрошње представља инверзну просечну вредност потрошње која се изазове променом температуре за  $1^\circ C$ . Овај коефицијент представља инверзну вредност  $AI$  из поглавља 5.1.2.1.1.

Коефицијент  $ddwc$  је уведен ради узимања у обзир колико се смањује утицај инерције са сваким претходним даном из предисторијске секвенце дана. Вредност овог коефицијента може да буде у распону од 0 до 1. Утицај првог предисторијског дана ( $j=1$ ) на сличност има вредност 1, у складу са  $ddwc^{j-1} = ddwc^0$ . Консеквентно, утицај



претходних предисторијских дана се смањује. Вредност за  $M$  број предисторијских дана и за коефицијент  $ddwc$  се одређује кроз метахеуристичку оптимизацију.

Сумаризацијом дневних умањујућих коефицијената у делиоцу формуле упросечавају се утицаји дана из низа  $M$ , пондерисано дневним умањујућим коефицијентима.

#### 5.1.2.2.2. Коефицијент термалне инерције

Коефицијент термалне инерције уведен је ради одређивања мере ефекта температуре у претходним данима на сличност између прогнозираног и историјског дана [120]. Означаваче се као просек апсолутних вредности разлика дневних просека температуре између чланова секвенце  $M$  предисторијских дана прогнозираног низа и секвенце  $M$  предисторијских дана историјског низа:

$$sc_{tin} = \frac{\sum_{j=1}^M |avgT_{d-j} - avgT_{d-n-j}| ddwc^{j-1}}{\sum_{j=1}^M ddwc^{j-1}}, \quad (9)$$

где је:

$sc_{tin}$  – коефицијент термалне инерције за дан  $d$ ,

$d$  – дан за који се рачуна прогноза,

$M$  – број предисторијских дана,

$n$  – број дана између прогнозираног дана  $d$  и сличног дана из историјског низа  $HA$ ,

$avgT_{d-j}$  – просечна сатна температура у  $d-j$ -ом дану из предисторијске секвенце прогнозираног низа дана;  $j=1, \dots, M$ ,

$avgT_{d-n-j}$  – просечна сатна температура у  $d-n-j$ -ом дану из предисторијске секвенце историјског низа дана,

$ddwc$  – дневни умањујући коефицијент за сваки од  $M$  дана (0-1).

#### 5.1.2.2.3. Коефицијент инерције последњег сата

Инерција последњег сата циља почетну тачку у дану који се прогнозира. Крива потрошње у дану за који се потрошња прогнозира има више шансе да добије прогнозирани облик ако има одговарајућу стартну позицију. Ако је дата потрошња последњег сата претходног дана у односу на прогнозирани дан, претрага сличних дана из историјске базе података може да се извршава и на основу потрошње у последњем сату претходног дана.

Коефицијент инерције последњег сата означава се као апсолутна вредност разлике између потрошње у последњем сату дана  $d-1$  и у последњем сату дана  $d-n-1$ :

$$sc_{lin} = |lhC_{d-1} - lhC_{d-n-1}| pc_{cin}, \quad (10)$$

где је:

$sc_{lin}$  – коефицијент инерције последњег сата,

$lhC_{d-j}$  – потрошња у последњем сату дана који претходи прогнозираном дану,

$lhC_{d-n-1}$  – потрошња у последњем сату дана који претходи историјском дану,

$pc_{cin}$  – партиципациони коефицијент за инерцију потрошње.

#### 5.1.2.2.4. Резултујући инерциони коефицијент

Коефицијент инерције потрошње, коефицијент термалне инерције и коефицијент инерције последњег сата формирају резултујући инерциони коефицијент, кроз аритметичку средину пондерисану тежинским коефицијентима:

$$SC_{in} = \frac{wC_{cin}SC_{cin}wC_{tin}SC_{tin}wC_{lhin}SC_{lhin}}{wC_{cin}+wC_{tin}+wC_{lhin}}, \quad (11)$$

где је:

$sc_{in}$  – резултујући инерциони коефицијент,

$wC_{cin}$  – тежински коефицијент за инерцију потрошње,

$sc_{cin}$  – коефицијент инерције потрошње,

$wC_{tin}$  – тежински коефицијент за термалну инерцију,

$sc_{tin}$  – коефицијент термалне инерције,

$wC_{lhin}$  – тежински коефицијент инерције последњег сата,

$sc_{lhin}$  – коефицијент инерције последњег сата.

#### 5.1.2.3. Коефицијент трајања дневне светлости

Коефицијентом трајања дневне светлости дефинише се промена у профилу потрошње која је резултат разлике у трајању дневног светла у току дана. Овај коефицијент израчунава се као апсолутна вредност разлике (у минутима) трајања дневног светла између прогнозираног дана ( $d$ ) и историјског дана ( $d-n$ ):

$$sc_{dd} = |dd_d - dd_{d-n}|pc_{dd}, \quad (12)$$

где је:

$sc_{dd}$  – коефицијент трајања дневне светлости

$dd_d$  – трајање дневне светлости за дан за који се рачуна прогноза

$dd_{d-n}$  – трајање дневне светлости за историјски дан

$pc_{dd}$  – партиципациони коефицијент за трајање дневне светлости

Партиципациони коефицијент за трајање дневне светлости одређује се на сличан начин као у случају коефицијента временских прилика: промена трајања дана за вредност  $A1$ , када се остали фактори занемаре, изазива просечну промену вредности потрошње ресурса за вредност  $A2$ . Партиципациони коефицијент за трајање дневне светлости ће, према томе, бити  $A2/A1$ .

#### 5.1.2.4. Коефицијент близине дана

Коефицијент близине дана дефинише релевантност дана из историје у одређивању сличности, на основу тога колико су историјски дани далеко у прошлости у односу на прогнозирани дан. Може се очекивати да су релевантнији скорији дани од оних који су даље у историји. Разлог за ово је у томе што се током времена изграђују капацитети и

мењају се инфраструктура, клима и навике потрошача. Овај коефицијент се израчунава као разлика у данима између дана  $d$  за који се рачуна прогноза и историјског дана  $d-n$ :

$$sc_{dp} = |d_d - d_{d-n}|pc_{dp}, \quad (13)$$

где је:

$sc_{dp}$  – коефицијент близине дана,

$d_d$  – дан за који се рачуна прогноза,

$d_{d-n}$  – дан из историјске базе података,

$pc_{dp}$  – партиципациони коефицијент за близину дана.

Партиципациони коефицијент за близину дана одређује се на основу два феномена: 1) на основу просечне разлике у потрошњи која се догађа са протоком времена, када су сви остали фактори приближно исти; 2) на основу промене понашања кривих потрошње током времена.

### 5.1.2.5. Резултујући коефицијент сличности

Када су израчунати сви подкоефицијенти који се узимају у обзир, коначни коефицијент сличности израчунава се као пондерисана аритметичка средина подкоефицијената:

$$s_c = \frac{wc_w sc_w wc_{in} sc_{in} wc_{dd} sc_{dd} wc_{dp} sc_{dp}}{wc_w + wc_{in} + wc_{dd} + wc_{dp}}, \quad (14)$$

где је:

$s_c$  – коефицијент сличности,

$sc_w$  – коефицијент временских прилика,

$wc_w$  – тежински коефицијент за временске прилике,

$sc_{in}$  – инерциони коефицијент,

$wc_{in}$  – тежински коефицијент за инерцију,

$sc_{dd}$  – коефицијент трајања дневне светлости,

$wc_{dd}$  – тежински коефицијент за дневну светлост,

$sc_{dp}$  – коефицијент близине дана,

$wc_{dp}$  – тежински коефицијент за близину дана.

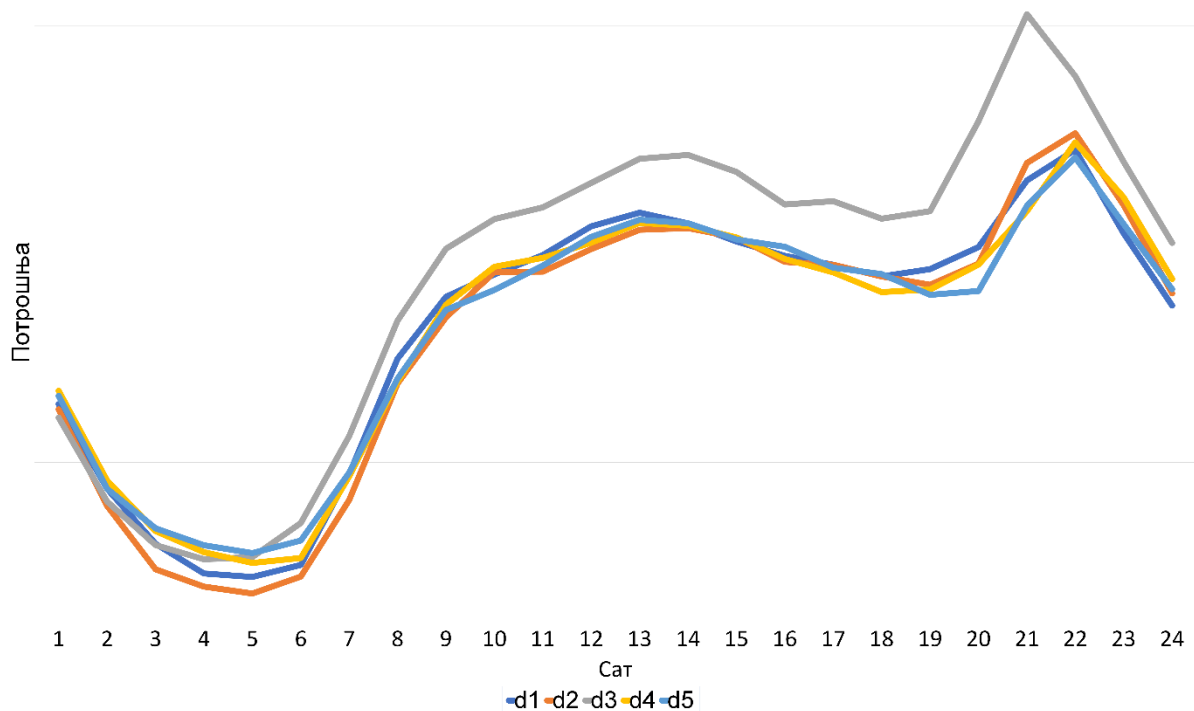
Тежински фактори за коефицијент временских прилика, инерциони коефицијент, трајање дневне светлости и близину дана одређују се уз помоћ метахеуристичке оптимизације.

### 5.1.3. Филтер девијације потрошње

Дневно понашање потрошње на сатном нивоу може да се прикаже профилем (кривом) потрошње. Претпоставка је да изабрани слични дани из историјске базе података имају сличне профиле потрошње. Ако нека крива значајно одступа од осталих, то би могао да буде знак да дан на који се та крива односи није одговарајући сличан дан. Разлози погрешног избора могу бити у специфичним околностима које алгоритам није

препознао: спортски догађаји националног значаја, телевизијске емисије неочекивано велике гледаности, политички догађаји, кризни догађаји, ретке метеоролошке појаве, хаварије на инфраструктурним системима или специфични друштвени догађаји/феномени. Циљ филтера девијације је да из узорка изабраних сличних дана елиминише оне чији профили потрошње значајно одступају од осталих.

Филтер девијације селекује  $z$  историјских дана са најмањим коефицијентом сличности  $sc$ , односно највећом сличности у односу на дан за који се потрошња прогнозира. Након тога, проверава се да ли постоје дани чији профили потрошње значајно одступају од осталих у овој групи. Ради појашњења, слика 7 приказује пример од  $z=5$  изабраних сличних дана. Очигледно је да су профили потрошње четири дана међусобно груписани, док профил потрошње једног дана значајно одступа од групе.

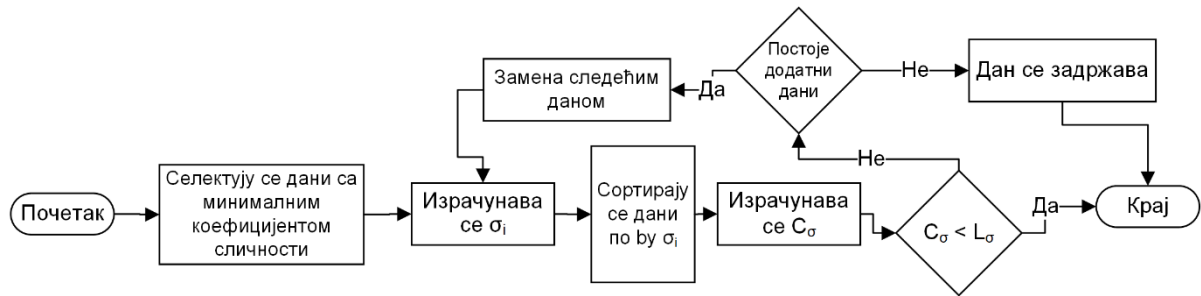


Слика 7 – Пример у коме профил потрошње једног сличног дана значајно одступа од осталих

Да би се елиминисали овакви дани из узорка, примењује се алгоритам филтера девијације потрошње (слика 8):

1. Селекује се  $z$  дана са минималном вредности коефицијента сличности  $Sc$ .
2. Израчунава се стандардна девијација сатне потрошње  $\sigma_i$  за сваки изабрани дан  $i=1, \dots, z$ , користећи формулу (15).
3. Дани се сортирају по растућој вредности девијације сатне потрошње  $\sigma_i$ .
4. Израчунава се допринос девијацији потрошње  $C_\sigma$  за дан са највећом вредности  $\sigma_i$ , користећи формулу (16).
5. Проверава се да ли је вредност  $C_\sigma$  нижа од изабране граничне вредности  $L_\sigma$ :

- ако јесте, дан је прихваћен, и изабрани узорак представља и финални узорак
- ако није, постоје две могућности:
  - Бира се следећи дан са најнижом вредности  $sc$ , али само ако  $sc$  има вредност до 200% најмање вредности  $sc$  изабраног скупа дана; у овом случају узорак од  $z$  сличних дана је ажуриран и кораци 2-5 се понављају.
  - Уколико није могуће да се пронађе следећи сличан дан ( $sc$  има вредност преко 200% у односу на најмању  $sc$  вредност изабраног скупа дана), дан се задржава (не замењује се) у скупу сличних дана.



Слика 8 – Алгоритам филтера девијације потрошње

Број  $z$  изабраних сличних дана је конфигурабилан и зависи од подешавања алгоритма прогнозе, односно имплементације машинског учења.

Стандардна девијација потрошње за  $i$ -ти дан  $\sigma_i$  историјског низа дана израчунава се формулом:

$$\sigma_i = \sqrt{\frac{\sum_{i=1}^{24} (C_i - \bar{C}_i)^2}{24}}, \quad (15)$$

где је:

$\sigma_i$  – стандардна девијација потрошње,

$C_i$  – потрошња у  $i$ -том сату дана који се разматра,

$\bar{C}_i$  – просечна потрошња у  $i$ -том сату за све  $z$  дане.

Допринос девијацији потрошње ( $C_\sigma$ ) израчунава се као релативни допринос у укупној девијацији за дан са највишом вредности  $\sigma_i$ :

$$C_\sigma = \frac{z\sigma_{max}}{\sum_{i=1}^z \sigma_i}, \quad (16)$$

где је:

$\sigma_{max}$  – највиша вредност стандардне девијације потрошње; то је девијација последњег дана после сортирања у трећем кораку алгоритма,

$z$  – број изабраних дана,

$\sigma_i$  – сатна девијација потрошње  $i$ -тог дана,  $i=1, \dots, z$ .

Вредност  $L_{\sigma}$  се израчунава метахеуристичком оптимизацијом, заједно са осталим параметрима.

## 5.2. Метахеуристичка оптимизација

Ефикасност филтера сличних дана зависи од тога како су оптимизоване вредности варијабли филтера. Оптимизациони проблем који се поставља дефинисан је објективном функцијом, варијаблама које се оптимизују, ограничењима за варијабле и начином претраге могућих решења.

Варијабле чије вредности је потребно оптимизовати су следеће:

- Индивидуални тежински коефицијенти фактора временских прилика  $w_{c_j}$  ( $j \in 1..nfw$ ).
- Тежински коефицијент за временске прилике  $w_{c_w}$ .
- Тежински коефицијенти за инерцију потрошње  $w_{c_{cin}}$ , термалну инерцију  $w_{c_{tin}}$  и инерцију последњег сата  $w_{c_{thin}}$ .
- Тежински коефицијент за инерцију  $w_{c_{in}}$ .
- Тежински коефицијент трајања дневне светлости  $w_{c_{dd}}$ .
- Тежински коефицијент за близину дана  $w_{c_{dp}}$ .
- Гранична вредност  $L_{\sigma}$ .
- Број предисторијских дана  $M$ .
- Дневни умањујући коефицијент  $ddw$ .

Ограничења за варијабле дефинисана су посебно за сваку варијаблу, на основу карактеристика самих варијабли.

Тежински коефицијенти фактора временских прилика могу да имају вредности између 0 и 100, са кораком 1. Ове граничне вредности дефинисане су тако да могу да прикажу разлику у утицају између различитих фактора који утичу на прогнозу.

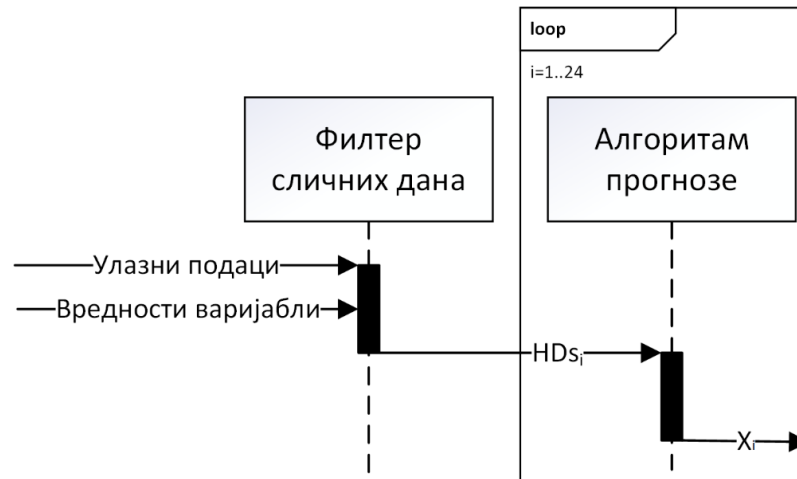
Гранична вредност  $L_{\sigma}$  може да има вредности у распону од 1 до 3, са кораком 0,1. Ове граничне вредности и корак одређени су тако да обухвате могуће вредности доприноса девијацији потрошње  $S_{\sigma}$ .

Варијабла  $M$  може да обухвати период до 5 предисторијских дана. Праћењем већег броја експеримената прогнозе потрошње, као и самом оптимизацијом, утврђено је да предисторијски дани који превазилазе овај опсег немају утицаја на резултате прогнозе потрошње.

Дневни умањујући коефицијент  $ddw$  може да има вредност између 0 и 1, са кораком 0,01. На овај начин се процентуално даје тежина сваком од предисторијских дана из опсега  $M$ .

Према адаптивном моделу прогнозе, претрага области могућих решења извршава се метахеуристичким методом.

Објективна функција (функција трошка) представља просечну апсолутну процентуалну грешку у процесу прогнозе за изабране дане из прошлости. Објективна функција се израчунава итеративно, понављањем прорачуна филтера сличних дана и алгоритма прогнозе са различитим вредностима варијабли које се оптимизују (слика 9).



Слика 9 – Једна итерација у метахеуристичкој оптимизацији

Једна секвенца од  $ND$  дана из историјске базе података проглашава се за „вештачку“ секвенцу дана за које се прогнозира потрошња. Вредност  $ND$  – број дана у секвенци – може да има различите вредности, у зависности од начина примене метахеуристичке методе. Уколико се оптимизација извршава ретко (на пример квартално или годишње) и циљ су стабилне вредности оптимизованих варијабли,  $ND$  може да има вредност 365 (једна година). Ако се оптимизација извршава на дневном нивоу и има циљ да прати текуће промене које утичу на оптимизоване параметре,  $ND$  може да има вредност 14 (две недеље).

За изабрани период од  $ND$  дана постоје и независни подаци који се користе као улаз у филтер сличних дана (на пример измерене метеоролошке вредности се проглашавају за „вештачке“ прогнозиране метеоролошке вредности).

Историјски подаци који претходе секвенци од  $ND$  дана представљају вештачку историјску базу података у метахеуристичкој оптимизацији.

У једној итерацији, улазни подаци у филтер сличних дана су вештачке историјске вредности потрошње и вештачки историјски независни подаци. Поред тога, креирана је комбинација вредности варијабли које се оптимизују. Филтер сличних дана се извршава за сваки сат у дану. На основу улазних података и вредности варијабли креира се низ сличних дана из вештачке историјске базе података ( $HDs$ ).

Главни циљ адаптивног модела прогнозе је што прецизније прогнозирање вредности потрошње ресурса за сваки од 24 сата у дану. Алгоритам прогнозе извршава прогнозу за сваки сат. Прогнозиране вредности  $x_i$  представљају излаз из алгоритма прогнозе. Метаксеуристички метод треба да пронађе такву комбинацију улазних варијабли да алгоритам прогнозе израчуна сатне вредности прогнозе са минималном просечном апсолутном процентуалном грешком (енгл. *Mean Absolute Percentage Error – MAPE*). Ова вредност се за један дан израчунава на следећи начин:

$$MAPE = \sum_{i=1}^{24} \left( \frac{|x_i - y_i|}{x_i} \right) 100\% \quad (17)$$

где:

$x_i$  – представља прогнозирану потрошњу у  $i$ -том сату вештачког дана који се прогнозира,

$y_i$  – представља остварену потрошњу у  $i$ -том сату вештачког дана који се прогнозира.

Коначна објективна функција метаксеуристичке оптимизације представља просечна *MAPE* (*NDMAPE*) вредност за секвенцу од  $ND$  дана:

$$NDMAPE = \frac{\sum_{j=1}^{ND} MAPE}{ND}. \quad (18)$$

### 5.2.1. Избор метаксеуристичке методе

Приликом разматрања избора метаксеуристичке методе за оптимизацију параметара, у обзир је узета чињеница да је оптимизациони проблем селекције сличних дана комплексан и може да има већи број локалних оптимума. На основу доступних извора, закључено је да су за решавање оваквих проблема најадекватнији избор оптимизација ројева честица и генетски алгоритам. Коначан избор направљен је на основу експеримента провере сличности изабраних дана, на подацима потрошње електричне енергије (оптерећења) на територији Србије без Косова. Експеримент је обухватао проверу ефикасности две наведене методе, кроз проверу сличности дана изабраних филтером сличности. Генетски алгоритам је тестиран у две варијанте: 1) са стандардним подешавањима која обухватају концепте као што су гени, хромозоми, популација, селекција, укрштање, мутација и функција трошка, и 2) са стандардним и унапређеним подешавањима, која обухватају и новоуведене концепте: концепт *мутације засноване на девијацији* и концепт *насељавања*.

Експериментом је тестирана просечна сличност изабраних сличних дана, у односу на прогнозирани дан. Због тога, и из разлога перформанси, није коришћен алгоритам машинског учења за поступак прогнозе, већ је извршен једноставан прорачун сличности: прогноза сатних вредности вештачког дана представља аритметичку



средину сатних вредности  $z$  изабраних сличних дана. Број  $z$  изабраних сличних дана је 5.

Вештачка секвенца дана за које се потрошња прогнозира је период од једне године, почев од дана 04.08.2018. године ( $ND=365$ ). Као узорак (вештачка историјска база) коришћени су подаци из претходне четири године у односу на први дан прогнозе. Вештачка историјска база је динамичка, и у прогнози сваког дана у низу  $ND$  завршава се непосредно пре дана за који се потрошња прогнозира.

#### **5.2.1.1. Подешавање оптимизације ројева честица**

Оптимизација ројева честица је програмирана у *Python* програмском језику и састоји се од 50 честица. Честице врше претрагу за решењем у 50-димензионалном простору решења, кроз 1000 итерација или до истека времена од 12 сати.

За сваку честицу, покренут је прорачун адаптивног модела прогнозе временских серија, за  $ND=365$ , од 04.08.2018. године до 03.08.2019. године.

Свака оптимизујућа варијабла има свој опсег – минималну и максималну вредност, и корак промене. Оптимизација је стартована са случајним вредностима у оквиру опсега за сваку оптимизујућу варијаблу. У свакој итерацији честице се групишу према честицама са најнижом вредности критеријумске функције ( $NDMAPE$ ).

Прекид калкулације је имплементиран за случајеве када нема напретка у конвергенцији за пет узастопних итерација или када истекне време од 12 сати.

#### **5.2.1.2. Подешавање стандардног генетског алгоритма**

Популација се у генетском алгоритму састоји од 64 хромозома са припадајућим генима. Сваки ген репрезентује једну оптимизујућу варијаблу. Комбинација гена у сваком хромозому се користи као комбинација варијабли за филтер сличних дана. Резултат критеријумске функције ( $NDMAPE$ ) придружује се сваком хромозому. Хромозоми са нижом вредности придруженог резултата критеријумске функције се задржавају, док се остали хромозоми одбацују. 32 задржана хромозома се међусобно укрштају и креирају нова 32 хромозома, са вредностима гена између вредности гена хромозома – родитеља. Свака нова популација садржи хромозоме са одређеним бројем мутираних вредности. Број мутираних гена у сваком хромозому је случајан, у опсегу од 1 до 4. Број мутираних хромозома зависи од вероватноће мутације. Вероватноћа мутације је у првој итерацији 3%, и линеарно се повећава у свакој следећој итерацији за 0,05%. Максимална вероватноћа мутације је 50%.

Оптимизација генетског алгоритма се завршава ако се вредност функције трошка не мења у 8 узастопних итерација, након истека од 10000 итерација или након истека времена од 12 сати.

### 5.2.1.3. Подешавање унапређеног генетског алгоритма

У унапређеном генетском алгоритму хромозоми, гени, критеријумска функција и услови конвергенција подешени су на исти начин као у подешавањима стандардног генетског алгоритма. Поред ових стандардних подешавања, имплементирани су новоуведени концепти: концепт насељавања, концепт мутације засноване на диверзитету и концепт мутације ка екстремима. Концепт насељавања полази од претпоставке да ефикасност оптимизације може да се повећава увођењем нових хромозома, са насумичним вредностима гена, у задржане („боље“) хромозоме. Концепт мутације засноване на диверзитету полази од претпоставке да учесталост мутације треба да зависи од диверзитета генетских варијабли и читаве популације. Концепт мутације ка екстремима има за циљ да повећа вероватноћу конвергенције варијабли код којих су оптималне вредности блиске екстремним (минималним или максималним) вредностима. Имплементирани концепти ће бити детаљније разрађени у поглављу 6.2.

### 5.2.1.4. Поређење резултата функције трошка

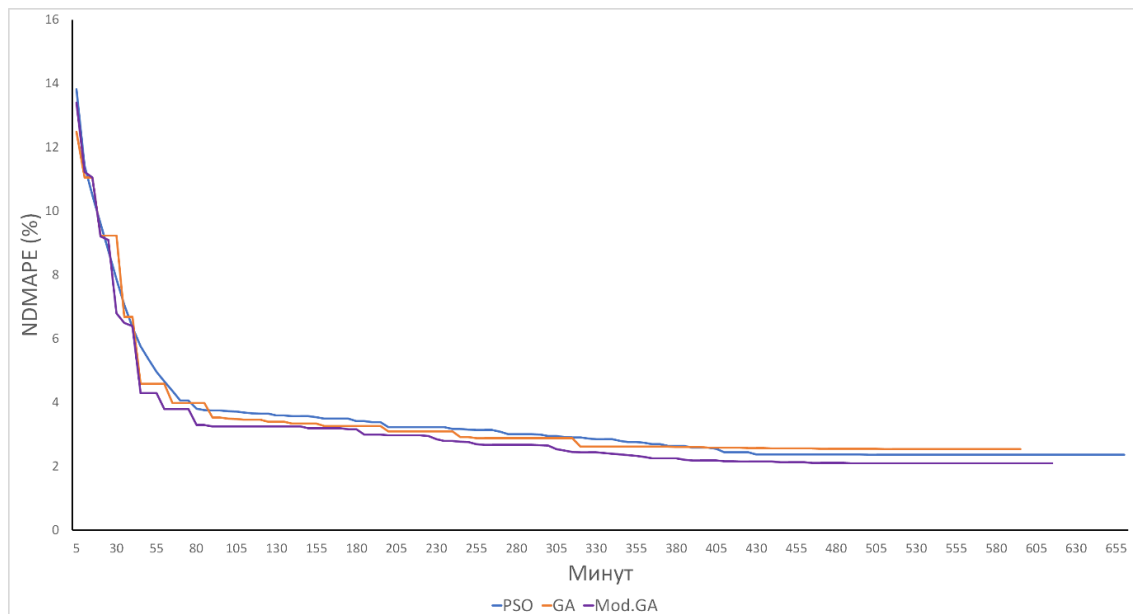
Све три упоређене технике – оптимизација ројева честица, стандардни генетски алгоритам и унапређени генетски алгоритам – успешно су завршили оптимизацију и вратили оптимизоване варијабле и минималну функцију трошка – *NDMAPE* вредност за вештачки период прогнозе од  $ND=365$  дана. Оптимизација је завршена после 595, 660 и 615 минута, респективно.

Вредности критеријумске функције по упоређиваним техникама су следеће:

Оптимизација ројева честица: 2,36%.

Стандардни генетски алгоритам: 2,55%.

Унапређени генетски алгоритам: 2,11%.



Слика 10 – Резултати функције трошка добијени тестираним оптимизационим методама

Слика 10 приказује криве конвергенције – вредности функције трошка по минутима, за сваки тестирани метод оптимизације. Евидентно је да је стандардни генетски алгоритам најраније конвергирао, али унапређени генетски алгоритам је конвергирао ка најнижој вредности критеријумске функције. Унапређени генетски алгоритам има раније скокове у конвергенцији у ранијим периодима извршавања, и више мањих скокова у каснијим периодима.

На основу извршеног експеримента, донета је одлука да се као метод оптимизације користи унапређени генетски алгоритам. Концепти унапређења биће детаљније описани у следећем делу.

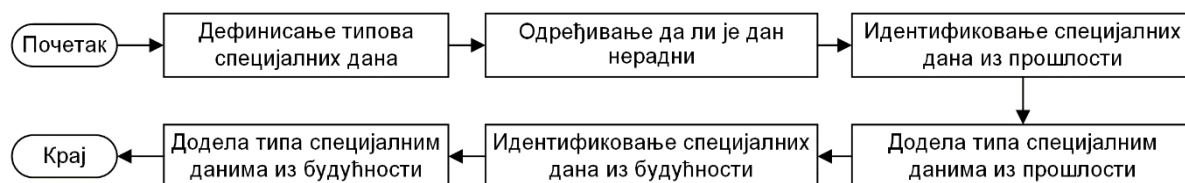
### 5.3. Филтер специјалних дана

Управљање специјалним данима најчешће је обавезан елемент система за управљање прогнозама у инфраструктурним мрежама. „Учење“ алгоритама прогнозе функционише на бази праћења трендова зависних и независних варијабли, у очекиваним околностима. Филтер сличних дана је задужен за то да околности у узорцима заиста буду на очекиваном нивоу у односу на прогнозирани дан. Међутим, постоје дани са специфичним околностима, у којима се очекује да крива потрошње значајно одступа од профила очекиваног за факторе које се узимају у обзир у поступку прогнозе. Овакви дани су најчешће празници, али то могу бити и значајни политички, спортски, верски, друштвени и други догађаји због којих се очекује промена у профилу потрошње.

У адаптивном моделу прогнозе временских серија, филтер специјалних дана извршава корекцију прогнозираног дана трендом грешке за специјалне дане из прошлости. Претпоставка је да је тренд грешке линеаран за дефинисани тип специјалног дана. Ова

претпоставка полази од чињенице да се у предложеном моделу прогноза за специјални дан у сваком сату израчунава симулацијом прогнозе, користећи стандардни алгоритам прогнозе који може да буде линеаран или нелинеаран, у зависности од избора поставке архитектуре. Ако су специјални дани груписани на одговарајући начин, очекује се да постоји линеаран тренд грешке у односу на стандардну прогнозу у свакој појави специјалног дана. На пример, у случају прогнозе потрошње електричне енергије, очекивање је да се крива оптерећења за дочек Нове године (31. децембра) понаша тако да потрошња у касним сатима расте уместо да опада (што би било прогнозирано стандардним алгоритмом прогнозе). Очекивање је да тренд грешке (разлике између стандардне прогнозе и остварених вредности) у сваком сату специјалног дана има линеарну правилност. Грешка може у мањој или већој мери да зависи и од разлике у другим параметрима међу специјалним данима, на пример од временских параметара. Према томе, у одређивању линеарног тренда грешке могу да се узму и вредности додатних фактора који се користе и у стандардном прорачуну прогнозе.

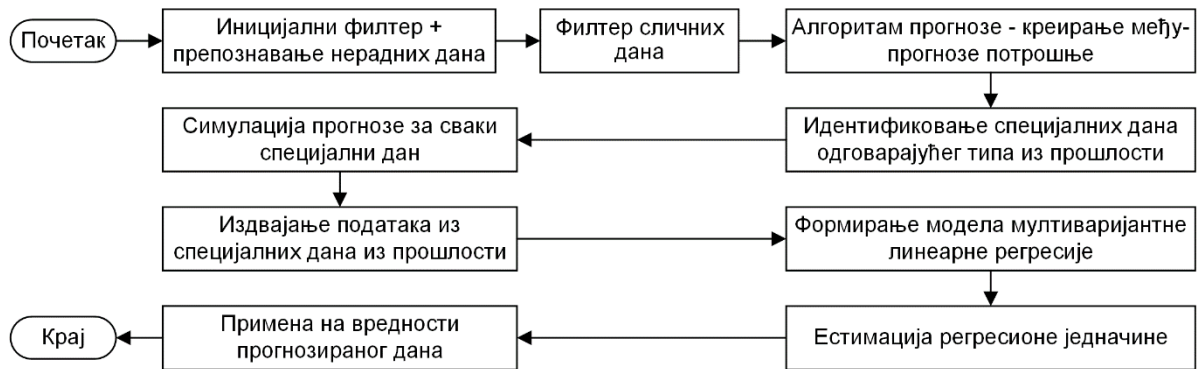
Иницијалне поставке које су неопходне за функционалност филтера специјалних дана приказане су на слици 11. Услов за прогнозу потрошње за специјалне дане је дефинисање типова специјалних дана. Сваки тип специјалних дана односи се на одређени дефинисани догађај који је имао више појава у прошлости. Сваки празник који утиче на профил потрошње може се дефинисати као тип специјалног дана (на пример Нова година, Први мај и др.). Сваки специјални дан има обележје које означава да ли се ради о нерадном дану. Нерадни дан подразумева да је у питању празник који се према регулативама држава за које се креира прогноза празнује нерадно.



Слика 11 – Иницијалне поставке у филтеру специјалних дана

У историјској бази података идентификују се дани који припадају дефинисаним специјалним данима и идентификованим данима се додељује припадајући тип специјалног дана. Након тога се идентификују специјални дани из будућности и извршава се додела типа специјалног дана за изабране дане, за које ће се тек прогнозирати потрошња. Доделом типа, ови дани постају специјални дани.

Када адаптивни модел прогнозе препозна да је дан за који се потрошња прогнозира специјални дан, поступак прогнозе се одвија на начин као што је приказано на слици 12.



Слика 12 – Поступак прогнозе за специјалне дане

Иницијалним филтером се прорачунава прогноза на стандардни, већ описани начин. Једина специфичност је у томе што се у иницијалном филтеру као дан у недељи узима недеља (или еквивалентни нерадни дан у зависности од државе), у случају да је тип специјалног дана нерадан. Разлог за ово налази се у претпоставци да ће профил потрошње у нерадном специјалном дану бити упоредив са стандардним нерадним данима. Филтером сличних дана селекују се слични дана и изабрани слични дани служе као узорак за алгоритам прогнозе.

Алгоритам прогнозе креира стандардну прогнозу у изабраном дану за сваки сат, на основу добијеног узорка. Ова прогноза представља сатну међупрогнозу потрошње за прогнозирани дан, и означава се са  $C_{tspd}$ .

Следећи корак је идентификација специјалних дана из историјске базе података. За сваки од њих извршава се симулација прогнозе, тако што дан за који се потрошња прогнозира постаје симулирани прогнозирани дан, а дани који претходе историјском специјалном дану и припадајући подаци постају симулирана историјска база података. Услов да би неки историјски дан могао да буде симулирани прогнозирани дан је да пре њега постоје историјски подаци од најмање две године.

Резултати прогнозе за сваки сат симулираних прогнозираних дана представљају међупрогнозу потрошње  $C_{tspd}$  за историјски специјални дан.

За сваки сат сваког специјалног дана из историјске базе података издвајају се независни и зависни подаци. Независне податке представља подскуп фактора који се користе у алгоритму прогнозе, и вредност  $C_{tspd}$ . Зависни подаци су подаци о оствареној потрошњи. На основу ових података формира се модел мултиваријабилне линеарне регресије, за сваки сат:

$$c_{spd} = \beta_0 + \beta_{tspd} C_{tspd} + \sum_{i=1}^P \beta_i x_i, \quad (19)$$

где је:

$c_{spd}$  – прогнозирана вредност потрошње,

$\beta_0$  – одсечак,

$\beta_{tspd}$  – нагиб за резултат симулације прогнозе за међупрогнозу потрошње,

$C_{tspd}$  – међупрогноза потрошње за специјални дан,

$P$  – број преосталих фактора,

$x_i$  – вредност преосталог фактора  $i$ ,

$\beta_i$  – нагиб за преостали фактор  $x_i$ .

Узорак за модел линеарне регресије добија се из специјалних дана из историјске базе података. Вредност потрошње ( $c_{spd}$ ) за специјалне дане из историјске базе података је измерена вредност потрошње. Преостали фактори из подскупа фактора који се користе у алгоритму прогнозе су опциони, односно вредност  $P$  може да буде 0.

Естимацијом регресионе једначине добијају се вредности  $\beta_0$ ,  $\beta_{tspd}$  и  $\beta_i$  за  $i \in 1..P$ . У одређивању вредности потрошње  $c_{spd}$  за прогнозирани дан, вредности добијене естимацијом регресионе једначине у сваком сату примењују се на међупрогнозу  $C_{tspd}$  одговарајућег сата прогнозираног дана и на вредности  $x_i$  за  $i \in 1, \dots, P$  одговарајућег сата прогнозираног дана. Добијена вредност представља коначну прогнозу потрошње за специјални дан.

## 5.4. Алгоритам прогнозе

Коначни прорачун за неспецијалне дане извршава се кроз алгоритам прогнозе. Алгоритам прогнозе се користи и за специјалне дане, али је у том случају резултат коригован филтером специјалних дана. Улаз у алгоритам прогнозе је листа сличних дана изабраних филтером сличних дана. Ефикасност алгоритма прогнозе директно зависи од изабраног узорка, али квалитет изабраног узорка може да одреди и коју имплементацију алгоритма је потребно изабрати за прогнозу. Квалитетан узорак изабраних сличних дана може да значи да су зависности независних и зависних варијабли линеарне, тако да и алгоритам прогнозе може да буде линеарни метод машинског учења. Међутим, готово немогуће је гарантовати довољно квалитетан узорак у сваком поступку прогнозе, тако да најчешће постоји потреба за нелинеарним моделом машинског учења.

Поступак оптимизације метахеуристичким методом подразумева велики број узастопних извршавања алгоритма прогнозе. С обзиром да је у метахеуристичкој оптимизацији важна сличност изабраних дана, у овом случају може да се користи проста аритметичка средина вредности потрошње за сваки сат сличних дана као алгоритам прогнозе.

У имплементацији адаптивног модела прогнозе временских серија користиће се могућност конфигурабилног избора алгоритма прогнозе. Постоји могућност избора једне од четири имплементације алгоритма прогнозе:

1. Линеарна регресија.

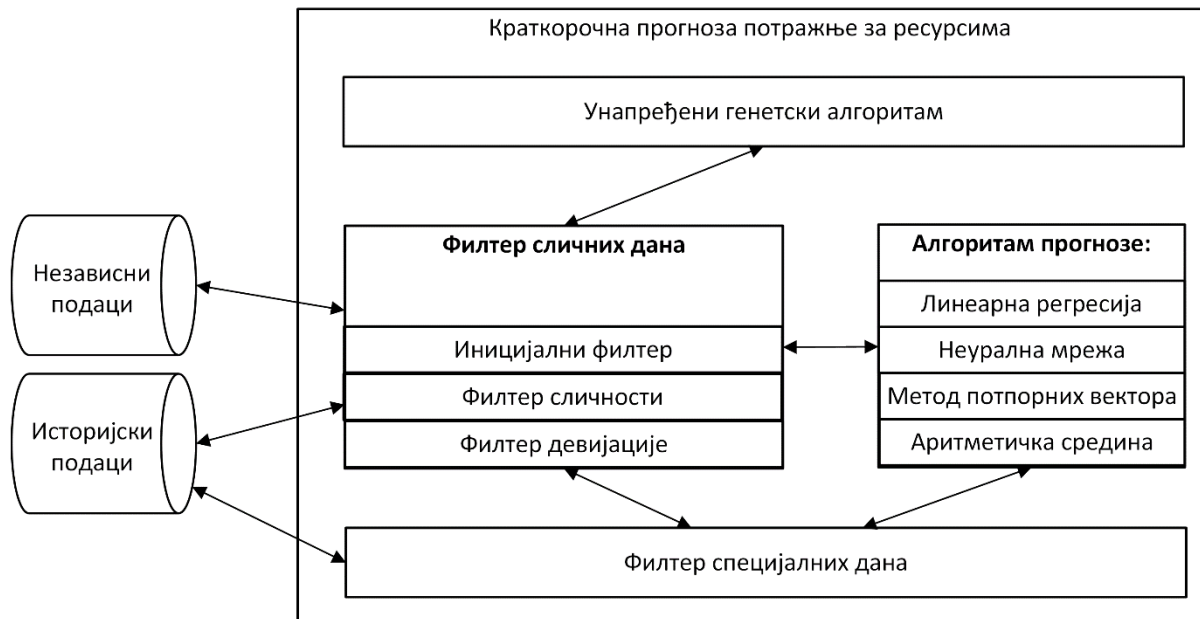
2. Неуронска мрежа.
3. Метод потпорних вектора.
4. Проста аритметичка средина (користи се само у метахеуристичкој оптимизацији).

Број дана који ће се користити као узорак, односно улаз за алгоритам прогнозе, такође, може да буде конфигурабилан и одређује се за сваки случај посебно. Овај број може да се бира на неки од следећих начина:

1. Фиксни број дана.
2. Број дана који испуњавају критеријум квалитета дана. Ово значи да се узимају у обзир сви дани који имају коефицијент сличности мањи или једнак од унапред постављене вредности. Највећи дозвољени коефицијент сличности се такође бира за сваки случај појединачно, итеративним испробавањем.
3. Бира се број дана у дефинисаном опсегу, уз поштовање критеријума коефицијента сличности. У овом случају постављају се минимални број дана, максимални број дана и максимални дозвољени коефицијент сличности. Број дана у узорку не може да има нижу вредност од максималног, без обзира на коефицијент сличности. Број дана не може да има већу вредност од максималног, без обзира на коефицијент сличности. Број дана између минималне и максималне вредности зависи од максималног дозвољеног коефицијента сличности.

## **5.5. Финална архитектура адаптивног модела прогнозе временских серија**

На слици 13 приказана је коначна архитектура адаптивног модела прогнозе временских серија у паметним инфраструктурним мрежама. Независни и историјски подаци користе се као улаз у филтер сличних дана. Филтер сличних дана састоји се од иницијалног филтера, филтера сличности и филтера девијације. Филтер сличних дана је оптимизован унапређеним генетским алгоритмом. Излаз из филтера сличних дана је листа сличних дана чији подаци се користе као улаз у алгоритам прогнозе. Алгоритам прогнозе може конфигурацијом да се подеси на једну од четири имплементације: линеарна регресија, неуронска мрежа, метод потпорних вектора и аритметичка средина која се користи само у процесу метахеуристичке оптимизације. Филтер специјалних дана препознаје специјалне дане и извршава модификацију прогнозе на основу линеарног тренда грешке из прошлости.



Слика 13 – Коначна архитектура адаптивног модела прогнозе временских серија

### 5.5.1. Перформансе модела

Перформансе у поступку прогнозе односе се пре свега на време корисничког извршавања процеса прогнозе. Алгоритми прогнозе конзумирају одређено време током извршавања, а то време између осталог зависи од величине броја узорака. Филтер сличних дана ограничава број узорака, тако да позитивно утиче на перформансе алгоритма прогнозе. Са друге стране, и филтер сличних дана захтева одређено време за своје извршавање. Међутим, ово време, које обухвата приступ бази података и филтрирање података на основу већ израчунатих коефицијената, требало би да траје далеко краће од уштеде времена добијене смањењем броја узорака за алгоритам прогнозе.

Метахеуристичка оптимизација, која подразумева континуелно позивање целог прорачуна, има релативно дугачко време извршавања. Међутим, ова оптимизација може да се извршава временски и архитектурално независно од филтера сличних дана и алгоритма прогнозе, тако да не утиче на корисничко време извршавања прогнозе.



# 6. Подешавање, унапређење и примена генетског алгоритма

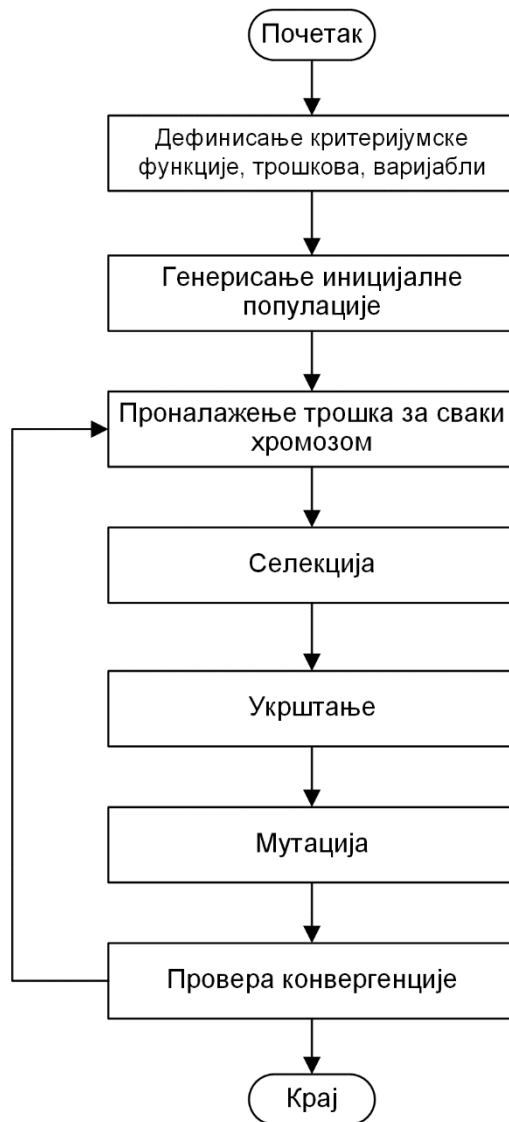
Генетски алгоритам је изабран као метод метахеуристичке оптимизације у адаптивном моделу прогнозе временских серија пре свега због могућности изласка из локалних оптимума током оптимизације. Важна особина генетског алгоритма је конфигурабилност, што значи да је у зависности од врсте оптимизационог проблема могуће подешавати параметре као што су величина популације, врста селекције, учесталост мутације и др. Поред тога, генетски алгоритам је конципиран тако да је могуће модификовати његове елементе у циљу додатног побољшања перформанси [111].

## 6.1. Подешавање генетског алгоритма

Ток генетског алгоритма приказан је на слици 14. Први корак је *избор критеријумске функције*. У случају адаптивног модела прогнозе временских серија, критеријумска функција је просечна *MAPE* вредност (*NDMAPE*) резултата прогнозе добијеног алгоритмом прогнозе. Поступак добијања *NDMAPE* вредности описан је у поглављу 5.2. Трошак за сваки хромозом представља *NDMAPE* вредност која је добијена филтером сличних дана и алгоритмом прогнозе. Алгоритам прогнозе у функцији трошка за сваки сат представља просту аритметичку средину вредности потрошње у одговарајућем сату за изабране сличне дане. Варијабле које се оптимизују набројане су у поглављу 5.2. Свака од варијабли представља један ген (генетску варијаблу) у генетском алгоритму.

*Генерисање иницијалне популације* подразумева креирање одређеног броја хромозома. Сваки хромозом се састоји од гена, при чему сваки ген представља варијаблу чија вредност се оптимизује генетским алгоритмом. Вредност генетских варијабли у иницијалној популацији представља случајно изабране вредности у распону од 0 до 100. Број хромозома у популацији је 64. Овај број је добијен итеративним тестирањем и узимањем у обзир брзине и успешности конвергенције.

*Проналажење трошка* за сваки хромозом подразумева извршавање критеријумске функције за сваки хромозом, односно комбинацију генетских варијабли.



Слика 14 – Ток генетског алгоритма

**Процес селекције**, који одређује улазак хромозома у процес репродукције, може да буде имплементиран на више различитих начина. Селекцијом се један део хромозома одбацује, док остали хромозоми улазе у процес репродукције. Хромозоми који улазе у процес репродукције у даљем тексту ће се звати *преостали хромозоми*.

Селекција методом „од врха на доле“ подразумева укрштање хромозома по редоследу сортирања на основу вредности критеријумске функције. Овом методом селекције непарни хромозоми се упарују са парним хромозомима док се не дође до дна листе задржаних хромозома. Случајна селекција (метод рулета) подразумева упаривање хромозома на основу случајног избора [112]. Пондерисана случајна селекција представља комбинацију претходна два метода. На основу ранга хромозома креира се вероватноћа упаривања, на основу које се хромозоми међусобно упарују. Метод турнира подразумева избор више група јединки (хромозома) међу којима се за селекцију бирају јединке са најнижим трошком. Турнир се понавља док не буду

изабрани сви парови. Селекција у генетском алгоритму у адаптивном моделу прогнозе временских серија је пондерисана случајна селекција.

**Укрштањем** се генеришу нови хромозоми – потомци који садрже генетске варијабле са вредностима које су настале на основу вредности хромозома – родитеља. Према најчешћем облику укрштања, два хромозома – родитеља креирају два хромозома – потомка, тако да се величина популације одржава. Према најједноставнијем процесу укрштања, хромозоми – потомци наслеђују генетске варијабле или од једног или од другог родитеља. Проблем са оваквим начином укрштања је у томе што се не уводе нове информације, већ се користе исте, иницијално генерисане. Такође, није погодан за адаптивни модел прогнозе временских серија, јер би довео до брзе конвергенције далеко од глобалног оптимума. Решење је у томе да генетске варијабле у хромозомима – потомцима добијају вредности између вредности генетских варијабли хромозома – родитеља и то на такав начин да вредности буду комплементарне. То значи да, ако је у једном хромозому – потомку вредност генетске варијабле ближа једном родитељу, у другом хромозому – потомку вредност ће бити ближа другом родитељу. Овај начин укрштања се користи у адаптивном моделу прогнозе временских серија јер омогућава ефикасну конвергенцију. Међутим, конвергенција постигнута на тај начин често је усмерена ка једном од локалних оптимума. Решење је у донекле у концепту мутације.

**Мутација** подразумева промене на случајно изабраним генетским варијаблама у хромозомима – потомцима. Стопа мутације означава који проценат варијабли ће бити мутиран у једној итерацији. Такође, може се посебно одредити број хромозома и број варијабли на којима ће бити извршена мутација. Начин и стопа мутације могу да буду подложни различитим прилагођавањима. Постоји више начина одређивања броја хромозома и броја гена на којима ће се извршити мутација. Један од најуспешнијих начина је теорија шеме (енгл. *Schema Theorem*), према којој се вероватноћа мутације повећава са каснијим генерацијама популација. Ова теорема полази од претпоставке да популације у ранијим генерацијама имају вишу вредност критеријумске функције и већу разноликост, тако да је мања потреба за мутацијама. Касније генерације имају нижу вредност критеријумске функције али и мању разноликост, тако да постоји већа потреба за мутацијом [121].

**Провера конвергенције** у адаптивном моделу прогнозе временских серија извршава се тако што се проверава средња вредност резултата критеријумске функције за „бољи“, односно преостали део популације. Ако се ова средња вредност не мења у осам узастопних итерација, генетски алгоритам је означен да је конвергирао.

## 6.2. Унапређење генетског алгоритма

Проблеми који су уочени приликом извршавања генетског алгоритма у адаптивном моделу прогнозе временских серија су следећи:

1. **Неодређеност стопе мутације.** Мутација помаже изласку из локалног минимума, али на нивоу појединачних варијабли над којима се мутација извршава. Повећана стопа мутације може да значи и значајно удаљавање од достигнутог локалног оптимума. Са друге стране, смањеном стопом мутације смањује се вероватноћа да се измена деси баш на генетској варијабли која је кључна за излазак из локалног и приближавање глобалном минимуму.
2. **Потреба за већим бројем хромозома у ранијим фазама извршавања.** Већи број хромозома у популацији значи и дужи период извршавања. Међутим, већи број хромозома може да буде значајан, нарочито у раним фазама извршавања, када већа разноврсност може иницијално да одведе алгоритам према више локалних оптимума, који касније могу да конвергирају ка глобалном оптимуму.
3. **Избор генетских варијабли за мутацију.** Случајан избор генетских варијабли над којима ће мутација бити извршена може да се покаже као неадекватан у случајевима када су „родитељи“ хромозоми достигли висок ниво конвергенције ка неком од локалних оптимума.
4. **Недовољна мутација у каснијим фазама извршавања.** Постоји потреба да се повећа стопа мутације у каснијим фазама извршавања генетског алгоритма, када постоји значајна конвергенција. Решење може да буде теорија шеме, односно да се стопа мутације повећава са бројем итерација. Међутим, не постоји гаранција да се број извршавања и конвергенција повећавају сразмерно. На основу тога уочена је потреба да се стопа мутације повећава не са бројем извршавања, него са уоченом стопом конвергенције, која се може идентификовати смањеним диверзитетом популације.
5. **Немогућност мутације ка екстремним вредностима.** Када се вредност неке варијабле на нивоу целе преостале популације помери ка некој од екстремних вредности (минималној или максималној), смањује се вероватноћа да стандардна мутација може да доведе до даљег приближавања екстремној вредности. У неким случајевима, управо екстремне вредности доводе до конвергенције. Често се дешава да генетске варијабле у свим преосталим хромозомима у популацији конвергирају ка вредности која је блиска екстремној вредности. Стандардним процесима укрштања и мутације релативно мале су шансе за конвергенцију ка оптималној вредности која се налази на неком од екстрема.

Решење ових проблема је у увођењу три нова концепта у генетском алгоритму: концепт насељавања, концепт мутације засноване на диверзитету и концепт мутације ка екстремима.

### 6.2.1. Концепт насељавања

Мутација подразумева измену вредности над појединачним варијаблама. Може се десити да за излазак из локалног минимума буде кључна само једна генетска варијабла у хромозому. У случају већег броја генетских варијабли у хромозому, смањује се

вероватноћа да ће се мутација извршити баш на том, кључном гену. Повећавање броја мутираних гена, а самим тим и додела случајних вредности већем броју генетских варијабли, може да доведе до удаљавања од достигнутог локалног, без приближавања глобалном оптимуму.

Уочена је могућност да се уведе делимична мутација на нивоу целог хромозома. Ово би подразумевало ограничено померање вредности свих генетских варијабли у хромозомима који су наследници хромозома који су већ постигли одређени степен конвергенције. Ограничено кретање вредности генетских варијабли, односно „fino подешавање“, може да остави хромозом у подручју блиском већ достигнутом локалном оптимуму, али промена вредности над свим варијаблама такође повећава вероватноћу да ће промена да се деси на некој од варијабли које су кључне за излазак из локалног и приближавања глобалном оптимуму.

У раду [122] предлаже се концепт самоорганизујућег алгоритма случајних имиграната, према коме се јединке са најнижим коефицијентом у преосталој популацији замењују са јединкама које садрже случајне вредности. Међутим, иако доприносе диверзитету, овакве нове јединке истовремено и „кваре“ преостали део популације. Са друге стране, са сувише малим бројем новоуведених јединки није велика вероватноћа да ће се десити пробој у конвергенцији.

Предложено решење „насељавања“ састоји се у процесу повременог укрштања преосталог дела популације са потпуно новом популацијом која се састоји од хромозома са случајним вредностима генетских варијабли. Укрштањем нових хромозома са већ постојећим, који су достигли одређени степен конвергенције, створиће се хромозоми – наследници, који ће имати промене у свакој својој генетској варијабли. Међутим промене у варијаблама ограничене су вредностима генетских варијабли једног од хромозома – родитеља, тако да постоји и вероватноћа да ће хромозом остати у простору достигнуте конвергенције.

Слика 15 приказује преостали (бољи) део популације. Број хромозома који се задржавају након што се одбаце „лошији“ хромозоми представљен је са „ $m$ “. Ознаке  $C_1, C_2, \dots, C_m$  представљају хромозоме у преосталом делу популације. Ознаке  $V_1, V_2, \dots, V_n$  представљају генетске варијабле у хромозомима. У стандардном процесу укрштања, преостали део популације улази у процес укрштања, у коме се бирају хромозоми – родитељи на одговарајући начин, и креирају се хромозоми – потомци.

1	$C_1V_1$	$C_1V_2$	$C_1V_3$	...	$C_1V_n$
2	$C_2V_1$	$C_2V_2$	$C_2V_3$	...	$C_2V_n$
⋮					
m	$C_mV_1$	$C_mV_2$	$C_mV_3$	...	$C_mV_n$

Слика 15 – Преостали део популације у генетском алгоритму

Концепт насељавања подразумева да се у повременим итерацијама, између два стандардна процеса укрштања, уведе нова популација  $s$  која ће садржати  $m/2$  хромозома и која ће се укрстити са  $m/2$  хромозома са врха листе преостале популације. Овако укрштени хромозоми креирају нових  $m$  хромозома. Нову ранг листу хромозома креирају: 1) првобитна листа  $m$  хромозома, 2) новоуведена листа  $m/2$  хромозома са случајним вредностима, и 3)  $m$  новокреираних хромозома – потомака. За листе 2) и 3) извршава се прорачун вредности критеријумске функције и целокупна листа се сортира по вредности критеријумске функције. За следећу итерацију задржавају се  $m$  хромозома са врха листе.

Предложеним концептом насељавања не губе се хромозоми из бољег дела популације. У случају да новоуведени хромозоми и новокреирани потомци не дају одговарајуће резултате, у следећу итерацију ће ући непромењена листа од  $m$  хромозома у односу на претходну итерацију. Међутим, на овај начин се повећава вероватноћа да ће неки од новоуведених хромозома и потомака допринети конвергенцији и изласку из локалног оптимума.

У почетним фазама генетског алгоритма увођење повремених нових популација са случајним вредностима може да повећа диверзитет популације, а самим тим и да повећа могућност конвергенције у неки од локалних оптимума, и у складу са тим и могућност приближавања глобалном оптимуму.

У каснијим фазама извршавања, укрштањем популације са новоуведеном популацијом креираће се хромозоми – наследници са вредностима генетских варијабли између оних које су већ довеле до конвергенције, и случајних вредности. На овај начин се заправо креира мутација, односно „фино подешавање“ целог хромозома, а не појединачних генетских варијабли.

Колико често ће се извршавати „насељавање“ одређује се параметром који се поставља пре извршавања генетског алгоритма, и који може да има динамичку вредност, односно да се учесталост насељавања повећава са конвергенцијом генетског алгоритма.

## 6.2.2. Концепт мутације засноване на диверзитету

Концепт мутације засноване на диверзитету односи се на диверзитет генетских варијабли приликом укрштања и на диверзитет целе популације.

### 6.2.2.1. Диверзитет генетских варијабли

Приликом процеса селекције и укрштања, идентичне вредности генетских варијабли у хромозомима – родитељима могу да значе да су те варијабле већ конвергирале ка неком од локалних оптимума, и да се у процесу укрштања неће мењати вредности над овим варијаблама код хромозома – наследника. Промена вредности управо над оваквим генетским варијаблама код хромозома – наследника може да буде кључна за излазак из локалног оптимума приликом конвергенције. Према концепту мутације заснованом на диверзитету генетских варијабли, приоритет у мутацији имају оне генетске варијабле у хромозомима – наследницима, код којих су хромозоми – родитељи имали идентичне вредности тих генетских варијабли.

За реализацију овог концепта, неопходно је да се информација о идентичности варијабли код хромозома – родитеља чува код хромозома – наследника. Ова информација се чува приликом процеса укрштања.

Овим концептом тежи се препознавању ситуације када су оба хромозома – родитеља достигле висок ниво конвергенције ка истом локалном оптимуму. Мутација се извршава над оним генетским варијаблама које могу бити одговорни за ту конвергенцију.

### 6.2.2.2. Диверзитет популације

Мутација заснована на диверзитету популације подразумева да постоји потреба да се број мутираних хромозома повећава у ситуацијама када је постигнут већи диверзитет популације. Диверзитет популације се одређује на основу постигнутог просечног нивоа конвергенције за преостали, односно „бољи“ део популације. Ако се просечна вредност критеријумске функције није значајно мењала у претходним итерацијама, то може да значи да је диверзитет популације мали, и да је потребно повећати стопу мутације.

Ниво конвергенције преосталог дела популације може да се израчуна формулом за стопу конвергенције:

$$con_{rate} = \frac{\sum_{i=-1}^{-c n} avg_{cost_i}}{c n}, \quad (20)$$

где је:

$con\_rate$  – стопа конвергенције,

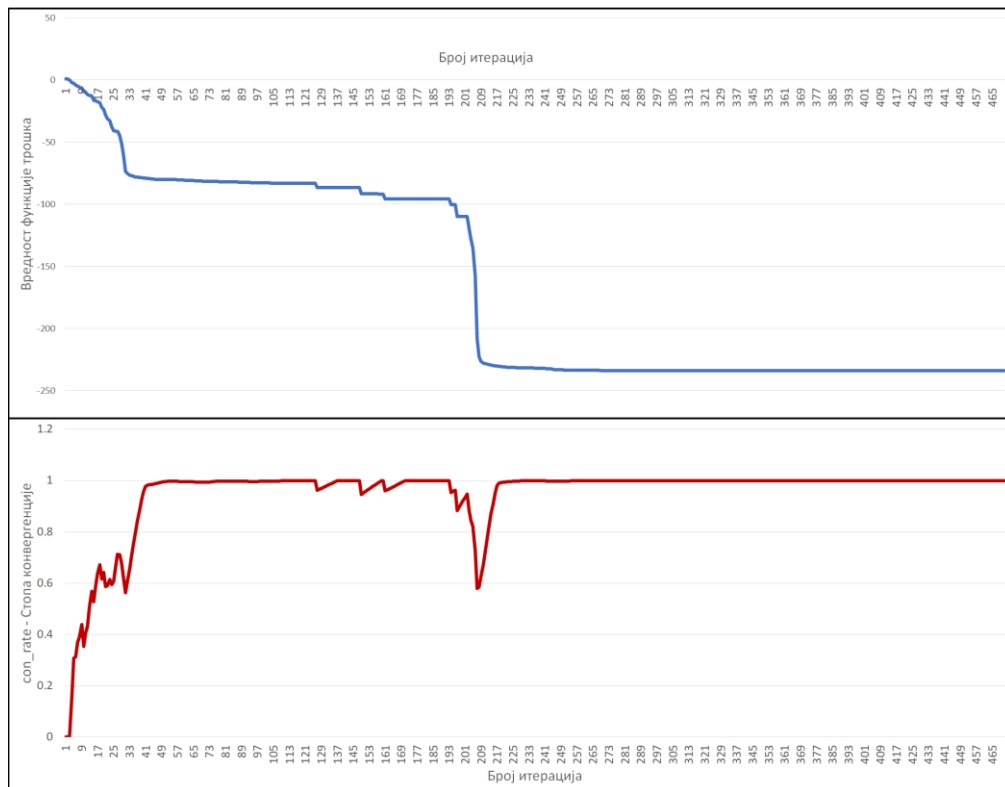
$ccn$  (број провера конвергенције, енгл. *Convergence Check Number*) – број претходних итерација на основу којих ће се извршити провера конвергенције,

$avg\_cost_i$  ( $i \in 1..ccn$ ) – просечни трошак преостале популације у претходних  $ccn$  итерација,

$avg\_cost$  – просечни трошак у тренутној преосталој популацији.

Вредност  $ccn$  може да варира у зависности од врсте оптимizacionог проблема. Ако су очекиване нагле промене у конвергенцији, ова вредност може да има малу вредност (на пример 10), што би значило да се конвергенција може уочити и на малом броју итерација. Ако се ипак очекује спорија конвергенција, овај број може имати већу вредност, што би значило да се конвергенција уочава на већем броју итерација.

Стопа конвергенције одређује максимални број мутираних хромозома. На тај начин се диверзитет популације одражава на стопу мутације. Виша стопа конвергенције значи већи број мутираних хромозома.



Слика 16 – Пример конвергенције и стопе конвергенције у генетском алгоритму

Пример конвергенције генетског алгоритма и упоредно праћење стопе конвергенције приказани су на слици 16. На горњем дијаграму прати се просечна вредност функције трошка по итерацијама. На доњем дијаграму види се да у итерацијама у којима се просечна вредност функције трошка драстично мења, стопа конвергенције опада. Мања промена у просечној вредности критеријумске функције значи да је стопа



конвергенције повећала вредност. Када нема промене у просечној вредности критеријумске функције, тада је стопа конвергенције 1. На основу ових вредности стопе конвергенције, број мутираних хромозома ће расти или ће падати.

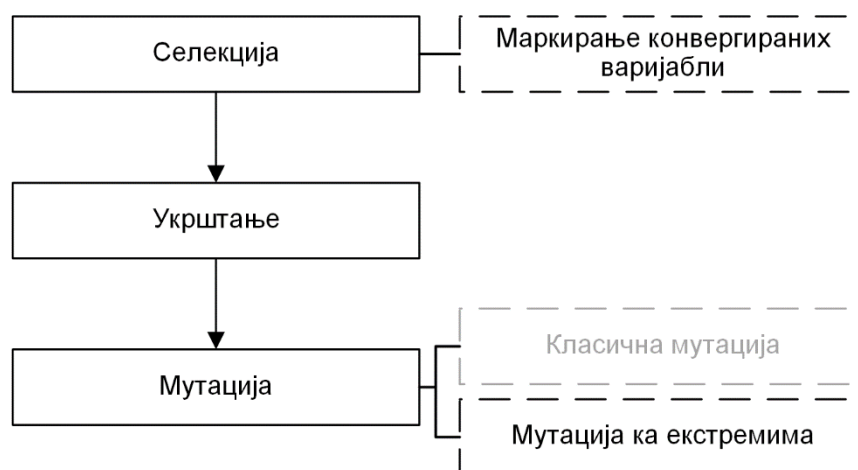
Када је достигнут дефинисани услов конвергенције, тада алгоритам мутације засноване на диверзитету популације покреће додатну мутацију. Додатна мутација подразумева мутацију над свим хромозомима – наследницима, кроз одређени број додатних итерација. У случају да се просечна вредност критеријумске функције смањи приликом додатне мутације, означава се да критеријумска функције није конвергирала ка глобалном оптимуму и наставља се процес извршавања генетског алгоритма. На овај начин се додатно проверава да ли је достигнут један од локалних оптимума, и да ли промене на вредностима генетских варијабли могу да доведу до изласка из локалног оптимума.

Концепт учесталости мутације на основу диверзитета популације примењује се и на учесталост насељавања на основу диверзитета популације.

### **6.2.3. Мутација ка екстремима**

Мутација ка екстремима идентификује генетске варијабле које су у потпуности конвергирале, односно које имају исту вредност у свим преосталим хромозомима у популацији. Процесом укрштања ове вредности се више не могу мењати. Једини начин да се те варијабле промене јесте мутација. Класична мутација поставља случајне вредности на ове варијабле. Ако је достигнута вредност блиска екстрему (минималној или максималној вредности), сразмерно мала је вероватноћа да ће мутирана вредност прећи тренутну вредност и још више се приближити ка екстрему. Мутација ка екстремима подразумева да мутација на оваквим генетским варијаблама има подједнаку вероватноћу повећања и смањења вредности генетске варијабле. На овај начин повећава се вероватноћа оптимизације генетских варијабли чија је оптимална вредност на максималној или минималној дефинисаној вредности.

Додатни кораци у генетском алгоритму којима се извршава мутација ка екстремима приказани су на слици 17. У процесу селекције маркирају се генетске варијабле које имају једнаке вредности у свим преосталим хромозомима. У процесу мутације ствара се подједнака вероватноћа за класичну мутацију и мутацију ка екстремима. Мутација ка екстремима препознаје конвергиране варијабле и у њима извршава мутацију ка екстремима, тако да постоји једнака вероватноћа померања вредности ка минималној и максималној вредности. Уколико је број конвергираних варијабли мањи од укупног броја предвиђеног за мутацију, остатак до тог броја узима се за класичну мутацију. Уколико конвергиране варијабле не постоје, извршава се само класична мутација.



Слика 17 – Мутација ка екстремима – кораци у генетском алгоритму

### 6.3. Тестирање унапређеног генетског алгоритма

Пре тестирања на адаптивном моделу прогнозе временских серија, унапређени генетски алгоритам је тестиран на низу оптимизационих проблема. Један од примера је тестирање тригонометријске функције, са пет варијабли. Овај пример је креиран тако да буде једноставан и репрезентативан у циљу представљања предности унапређеног генетског алгоритма:

$$f = \sqrt{|\cos(x)|} + \cos(y)^2 + \sin(z) + a^2 + \sqrt{b}. \quad (21)$$

У наведеном тесту потребно је оптимизовати варијабле  $x$ ,  $y$ ,  $z$ ,  $a$  и  $b$ , тако да се трага за минимумом функције. Свака варијабла ограничена је минималним и максималним вредностима, и за сваку варијаблу дефинисан је корак претраге. Граничне вредности и кораци претраге дефинисани су тако да свака од саставних делова функције има минимум у једној тачки, и да неке од функција имају више локалних минимума. На овај начин целокупна функција која се тестира има више локалних и један глобални оптимум.

Варијабле које је потребно оптимизовати имају следеће особине (табела 1):

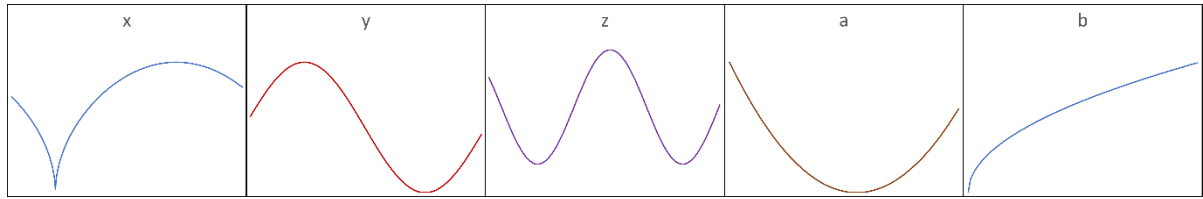
Табела 1 – Особине тестних варијабли у оптимизацији генетским алгоритмом

Варијабла	Мин. вредност	Макс. вредност	Корак
$x$	1	4	0,01
$y$	37	40	0,01
$z$	78	88	0,1
$a$	-5	4	0,1
$b$	0	100	1

Прва колона представља назив варијабле. Минимална и максимална вредност су граничне вредности за које ће се извршавати претрага. Промена вредности у претрази

оптималних вредности мора бити дељива са кораком (последња колона). Сваки ред представља једну варијаблу са својим особинама.

Са додељеним минималним и максималним вредностима и корацима, саставни делови функције који се сабирају имају облике кривих за варијабле као што приказује слика 18.



Слика 18 – Облици кривих за појединачне делове оптимизационог проблема

Као што се види, варијабле  $x$ ,  $y$  и  $z$  имају по два локална минимума. Свака од ових варијабли има по један глобални минимум. Варијабла  $b$  има минимум у пресеку криве и ординате.

Целокупна критеријумска функција која се тестира има један глобални минимум, са вредности 0,97.

За потребе теста постављена су четири различита подешавања генетског алгорита. Сва подешавања имају популацију од 64 хромозома, пондерисану случајну селекцију и укрштање са комплементарним вредностима генетских варијабли. Специфичности у подешавањима су следеће:

1. Стандардни генетски алгоритам. Мутација на 6 хромозома, и на случајно изабраном броју варијабли, у распону од 1 до 3.
2. На Подешавање 1 додат је концепт насељавања.
3. На Подешавање 2 додат је концепт мутације засноване на диверзитету.
4. На Подешавање 3 додат је концепт мутације ка екстремима.

Тест је обухватао 20 циклуса генетског алгорита за свако подешавање.

Резултати теста су приказани у табели 2. Евидентира се резултат критеријумске функције и број итерација којима се дошло до резултата.

Табела 2 – Резултати теста генетског алгоритма

Подешавање 1		Подешавање 2		Подешавање 3		Подешавање 4		
<i>f</i>	Бр. ит.	<i>f</i>	Бр. ит.	<i>f</i>	Бр. ит.	<i>f</i>	Бр. ит.	
-0,97	53	-0,96	75	-0,97	216	-0,97	146	
-0,97	63	-0,97	47	-0,97	154	-0,97	86	
-0,96	53	-0,96	77	-0,97	138	-0,97	88	
-0,94	45	-0,96	55	-0,97	160	-0,97	146	
-0,88	51	-0,95	49	-0,97	89	-0,97	183	
0,76	32	-0,97	42	-0,97	172	-0,97	111	
-0,81	85	-0,97	52	-0,97	64	-0,97	92	
-0,72	41	-0,97	36	-0,97	134	-0,97	163	
-0,82	43	-0,97	43	-0,97	197	-0,97	112	
0,03	3	0,2	58	-0,97	62	-0,97	144	
0,03	31	-0,97	44	-0,97	200	-0,97	54	
-0,96	49	0,1	49	-0,97	228	-0,97	58	
1,27	44	-0,91	67	-0,97	101	-0,97	171	
0,03	70	-0,7	74	-0,97	114	-0,97	168	
0,04	47	0,61	45	-0,97	96	-0,97	113	
-0,94	61	-0,93	62	-0,97	172	-0,97	204	
-0,97	67	-0,85	36	-0,97	172	-0,97	107	
0,05	63	-0,97	55	-0,97	131	-0,97	180	
0,12	45	-0,95	46	-0,9	185	-0,97	311	
0,44	34	0,1	53	-0,9	193	-0,97	53	
<b>просек</b>	<b>-0,3585</b>	<b>49,00</b>	<b>-0,70</b>	<b>53,25</b>	<b>-0,96</b>	<b>148,9</b>	<b>-0,97</b>	<b>134,5</b>

У подешавању 1 резултат је у неколико наврата био близак глобалном оптимуму. Међутим, очигледно је да је више пута генетски алгоритам конвергирао ка неком од локалних оптимума. Просечан број итерација је 49.

Додавање концепта насељавања у подешавању 2 довело је до тога да се у 12 од 20 тестова генетски алгоритам приближи глобалном минимуму, или га чак достигне. Просечан број итерација незнатно је повећан на 53,25.

Додавањем концепта мутације засноване на диверзитету значајно је повећан просечан број итерација. Међутим у 18 од 20 случајева достигнут је глобални минимум, док је у преостала два случаја достигнут резултат био близак глобалном минимуму.

Додавањем концепта мутације ка екстремима глобални минимум је достигнут у свих 20 тестних случајева, док је просечан број итерација благо смањен на 134,5.

Анализом конвергенције генетског алгоритма са додељеним подешавањима, и анализом самих резултата, дошло се до закључка да унапређени генетски алгоритам решава свих 5 уочених проблема.

Концептом насељавања решава се проблем 1 – *неодређеност стопе мутације*. „Фино подешавање“ се извршава над свим варијаблама у хромозому чији непосредни предак је хромозом који се „населио“ у популацију. На овај начин повећава се шанса да ће један

од оваквих хромозома да промени вредност варијабле која је кључна за приближавање глобалном оптимуму, а да се истовремено задржи достигнути ниво конвергенције.

Концептом насељавања такође се решава и проблем 2 – *потреба за већим бројем хромозома у ранијим фазама извршавања*. Увођењем нових хромозома разноврсност популације се повећава, што у раним фазама извршавања може да значи кретање хромозома ка више локалних оптимума, који касније могу да доведу до глобалног оптимума.

Концепт мутације заснован на диверзитету генетских варијабли решава проблем 3 – *избор генетских варијабли за мутацију*. Идентичне вредности одређених генетских варијабли могу да значе конвергенцију ка локалном оптимуму. Мутација таквих варијабли тежи да их преусмери према потенцијалном глобалном оптимуму варијабле.

Концепт мутације заснован на диверзитету популације решава проблем 4 – *недовољна мутација у каснијим фазама извршавања*. Конвергенција ка локалном минимуму најчешће значи смањење диверзитета популације. Повећање стопе мутације са повећањем конвергенције повећава вероватноћу изласка из простора локалног и приближавање глобалном оптимуму.

Мутација ка екстремима решава проблем 5 – *немогућност мутације ка екстремним вредностима*. Код варијабли чије су вредности у потпуности конвергирале, изједначавање вероватноће да се у мутацији вредност варијабле креће ка једној или другој граничној вредности повећава вероватноћу да се оптимизација вредности варијабле настави и преко претходно достигнуте конвергиране вредности. У наведеном примеру, варијабла  $b$  има оптималну вредност у минималној вредности. Разлика између резултата за подешавање 3 и подешавање 4 је у томе што је у подешавању 4 варијабла  $b$  у два циклуса конвергирала ка вредности која је блиска минималној, али не и ка минималној. У подешавању 4 у свим циклусима варијабла  $b$  је конвергирала ка минималној вредности.

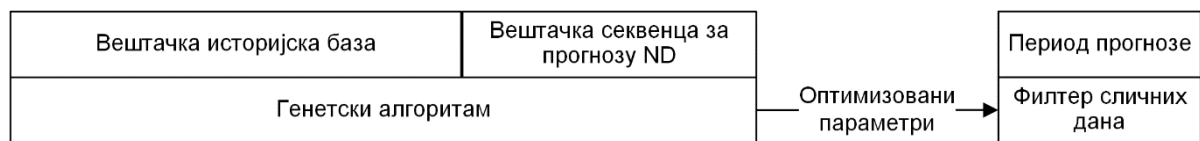
## 6.4. Примена генетског алгоритма у адаптивном моделу прогнозе

Примена метахеуристичке оптимизације ограничена је начином добављања података. Подаци о реализованој потрошњи и подаци о оствареним метеоролошким вредностима доступни су увек за претходни дан. То значи да се провера успешности оптимизације извршава само за претходне дане. Прогноза за дан унапред доступна је након једног дана, прогноза за два дана унапред доступна је два дана након прогнозе итд.

Примена метахеуристичке оптимизације генетским алгоритмом подразумева одређивање оптималних вредности параметара за одређени период из прошлости и

примену оптимизованих параметара на прогнозу за дане у будућности, у реалној прогнози.

Слика 19 приказује периоде релевантне за оптимизацију. Период за који се извршава оптимизација представља вештачку секвенцу за прогнозу и означава се са *ND*. Симулација прогнозе у периоду *ND* извршава се за сваки појединачни дан. За сваки од ових дана постоје подаци о реализованој потрошњи, тако да се може израчунати апсолутна просечна процентуална грешка (*MAPE*). Критеријумска функција у генетском алгоритму за *ND* период представља просечну *MAPE* вредност за *ND* период (*NDMAPE*).



Слика 19 – Периоди метахеуристичке оптимизације

Секвенца *ND* може да има различите дужине. Што је овај период дужи, резултати оптимизације су стабилнији – њихова вредност је релевантна у различитим периодима током године. Уколико је период *ND* краћи, оптимизовани параметри могу да буду релевантни у кратком периоду након оптимизације, али се током времена релевантност ових параметара смањује. На пример, ако је период *ND* једна година (365 дана), очекује се да ће оптимизовани параметри моћи да се користе макар током целе следеће године. Ови оптимизовани параметри су одговарајући за неке краће периоде у години уз напомену да им се релевантност не мења током времена и са променом годишњих доба. Са друге стране, ако је *ND* период 14 дана, велика је вероватноћа да ће тренд утицаја оптимизованих параметара на тачност прогнозе, коме оптимизовани параметри припадају, бити актуелан и следећих неколико дана. Релевантност оптимизованих параметара у наредним данима ће бити велика, али ће након тога падати. Са променом годишњег доба ови параметри ће потпуно изгубити релевантност.

Период прогнозе у филтеру сличних дана такође може да има различите периоде. Прогноза се извршава у временској серији, са резолуцијом од једног сата. Број дана овог периода је од 1 до 7, и зависи од расположивих независних података (подаци о временској прогнози и сл.). Различити дани у прогнози могу да имају различите оптималне параметре. Оптимизација параметара у различитим данима прогнозе се своди на три чињенице:

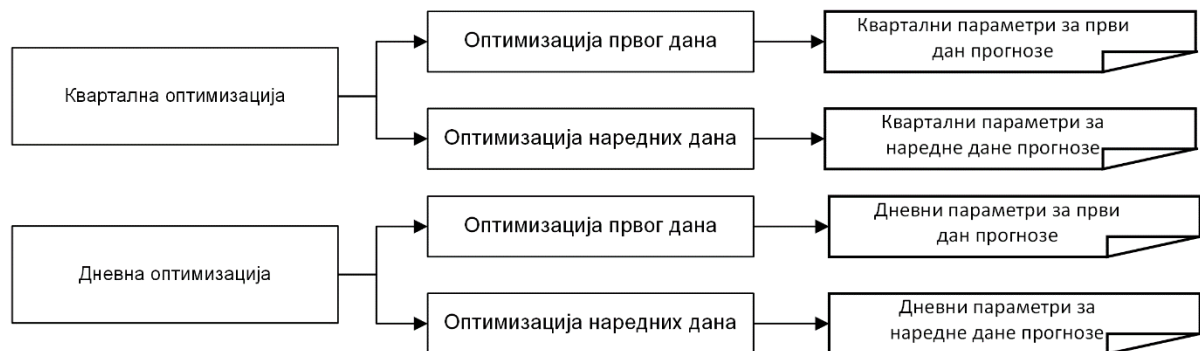
1. Инерција као фактор од утицаја на потрошњу има већи утицај на први следећи дан, док се тај утицај смањује са протоком времена.
2. Прецизност метеоролошких прогноза се смањује током времена.
3. Календарски подаци су једнако релевантни у свим периодима прогнозе.

Логично се постављају три питања:

1. Колика дужина периода  $ND$  треба да буде у адаптивном моделу прогнозе?
2. Колико често се покреће оптимизација?
3. Како се решава проблем оптимизације параметара у различитим данима прогнозе?

У адаптивном моделу прогнозе временских серија постоје два типа и два подтипа оптимизације (слика 20):

1. Квартална оптимизација
2. Дневна оптимизација
  - Оптимизација прогнозе првог дана
  - Оптимизација прогнозе наредних дана



Слика 20 – Типови оптимизације параметара у адаптивном моделу прогнозе

Оптимизација прогнозе првог дана и оптимизација прогнозе осталих дана се извршавају у свакој кварталној и свакој дневној оптимизацији.

**Квартална оптимизација** се извршава квартално и  $ND$  период у овој оптимизацији је 365 дана. Овом оптимизацијом креирају се стабилни параметри за филтер сличних дана. Параметри оптимизовани кварталном оптимизацијом зову се **квартални параметри**.

**Дневна оптимизација** се извршава дневно и  $ND$  период у овој оптимизацији је 14 дана. У овој оптимизацији креирају се параметри који су оптимални за наредних неколико дана. Пошто се ова оптимизација извршава дневно, параметри добијени овом оптимизацијом ће се користити у следећем дану. Параметри добијени дневном оптимизацијом зову се **дневни параметри**.

Дневна оптимизација се извршава кумулативно у односу на претходну дневну оптимизацију. Минималне и максималне вредности параметара ограничене су вредностима параметара добијених претходном дневном оптимизацијом. Да би се спречила прекомерна промена вредности дневних параметара, промена у односу на претходне вредности може да се креће плус/минус 10% у односу на дневне вредности добијене претходном оптимизацијом. На овај начин спречавају се утицаји на оптимизацију ванредних околности које су могле да се десе у  $ND$  периоду.

**Оптимизације прогнозе првог дана** се односи на оптимизацију за први дан прогнозе. У симулацији прогнозе у овом подтипу вештачка историјска база се за сваки дан у симулацији завршава последњим даном пре дана за који се извршава симулација прогнозе. На тај начин се у симулацији прогнозе симулира прогноза за наредни дан.

**Оптимизација прогнозе за наредне дане** се односи на оптимизацију за други и све наредне дане прогнозе. У симулацији прогнозе у овом подтипу вештачка база података се за сваки дан завршава два дана пре дана за који се извршава симулација прогнозе. На овај начин се у симулацији прогнозе симулира прогноза за други дан. Исти подтип оптимизације се користи за оптимизацију прогнозе за други дан и за оптимизацију прогнозе за све остале дане.

На основу два типа и два подтипа, у сваком моменту за сваки параметар постоје четири врсте оптимизованих вредности параметара:

1. Квартални параметри за први дан.
2. Квартални параметри за остале дане.
3. Дневни параметри за први дан.
4. Дневни параметри за остале дане.

У прогнози потрошње за предвиђени период, користе се параметри за први дан за све сате који су у дану у коме се и прогнозира потрошња. За све сате осталих дана користе се параметри за остале дане.

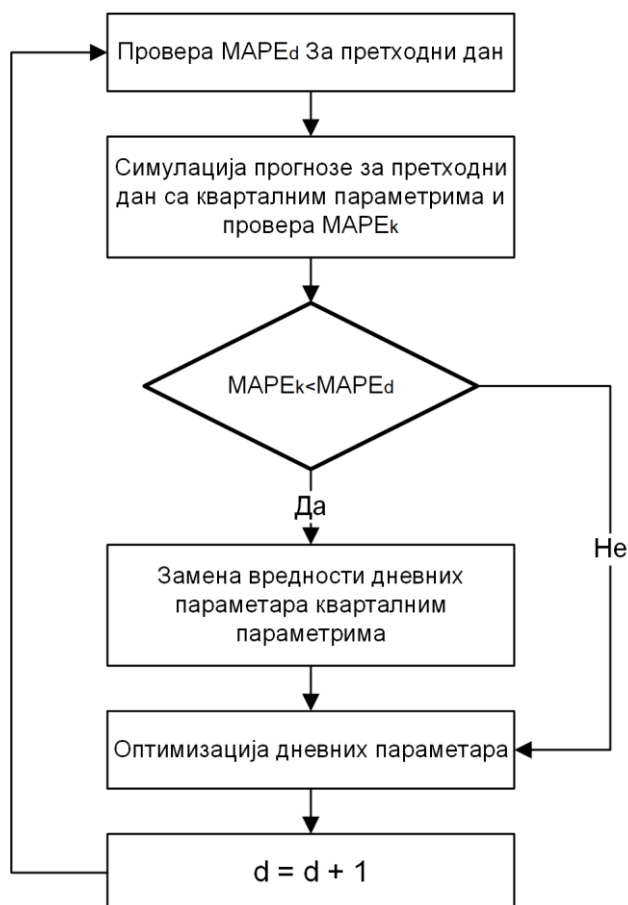
Слика 21 приказује поступак усаглашавања дневних и кварталних параметара. Ради утврђивања ефикасности дневних и кварталних параметара, израчунавају се две  $MAPE$  вредности за претходни дан: вредност добијена коришћењем дневних параметара ( $MAPE_d$ ) и вредност добијена коришћењем кварталних параметара ( $MAPE_q$ ).

У редовном поступку избора сличних дана, филтер сличних дана користи дневне параметре. Један дан након прогнозе доступни су подаци о стварној прогнози за први дан у периоду прогнозе, тако да се може израчунати  $MAPE_d$  за тај дан. Два дана након прогнозе доступни су подаци за други дан периода прогнозе, тако да се може израчунати  $MAPE_d$  и за тај дан.

Пре почетка дневне оптимизације креира се симулација прогнозе за претходна два дана, са кварталним параметрима. За сваки од ова два дана извршава се поређење  $MAPE_d$  вредности добијене коришћењем дневних параметара, и вредности  $MAPE_q$  добијене коришћењем кварталних параметара. Уколико  $MAPE_q$  има нижу вредност, вредности дневних параметара се замењују вредностима кварталних параметара. Ово важи посебно за параметре за први дан и за параметре за остале дане. Разлог овоме је претпоставка да је препозната ситуација када су вредности дневних параметара изгубиле релевантност и потребно је да њихова оптимизација крене испочетка, са иницијалним вредностима кварталних параметара.



Поступак усаглашавања дневних и кварталних параметара извршава се на исти начин и у случају оптимизације дневних параметара за први дан и у случају оптимизације дневних параметара за остале дане. У случају оптимизације за први дан,  $MAPE_d$  и  $MAPE_q$  вредности могу да се израчунају један дан након претходне прогнозе. У случају оптимизације за остале дане,  $MAPE_d$  и  $MAPE_q$  вредности могу да се израчунају два дана након претходне прогнозе



Слика 21 – Усаглашавање дневних и кварталних параметара

Описаним поступком параметри се адаптивно оптимизују, извршавањем генетског алгоритма на дневном нивоу. Уколико се покаже да оптимизација не даје очекиване резултате, вредности дневних параметара се постављају на иницијалне, односно на кварталне параметре и ове вредности представљају нову полазну тачку у дневној оптимизацији. На овај начин описани поступак има и карактеристику саморегулације и континуалног унапређења.

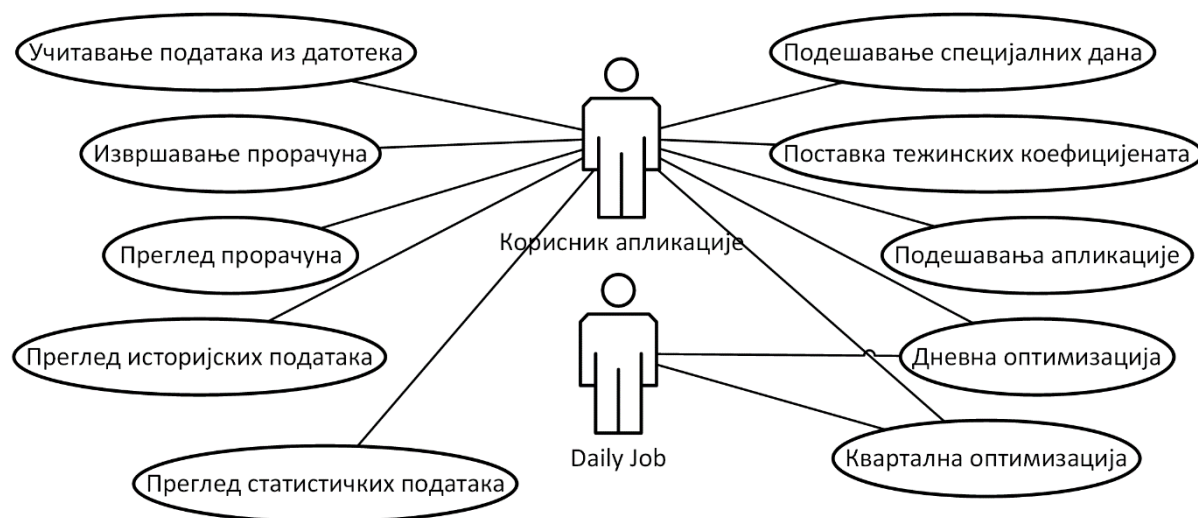
# 7. Софтверска реализација модела

Адаптивни модел прогнозе временских серија примењен је у апликацији за краткорочну прогнозу потрошње ресурса. Апликација је креирана тако да може да се односи на потрошњу било ког ресурса који испуњава услове постављене самим моделом. Апликација се активно користи у електроенергетском преносном систему „Електромрежа Србије“ за потребе краткорочне прогнозе потрошње електричне енергије, за подручје Србије, без територије Косова и Метохије.

Ради приказа примене адаптивног модела прогнозе временских серија, биће описани архитектура, дизајн и техничке карактеристике апликације.

## 7.1. Приказ функционалности

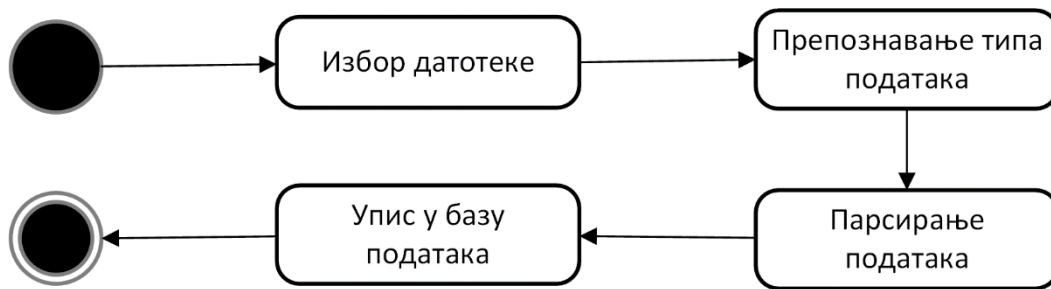
На слици 22 приказани су случајеви коришћења обухваћени у апликацији. Учесници релевантни за приказ функционалности од интереса јесу корисник апликације и *Daily Job*. Корисник апликације покреће све функционалности, док је *Daily Job* аутоматски процес који управља дневним и кварталним тренингом. У наставку ће укратко бити описане појединачне активности.



Слика 22 – Дијаграм случајева коришћења

### 7.1.1. Учитавање података из датотека

У овој активности (слика 23) учитавају се подаци о оствареној потрошњи, временској прогнози за будућност и оствареним временским параметрима за прошлост.

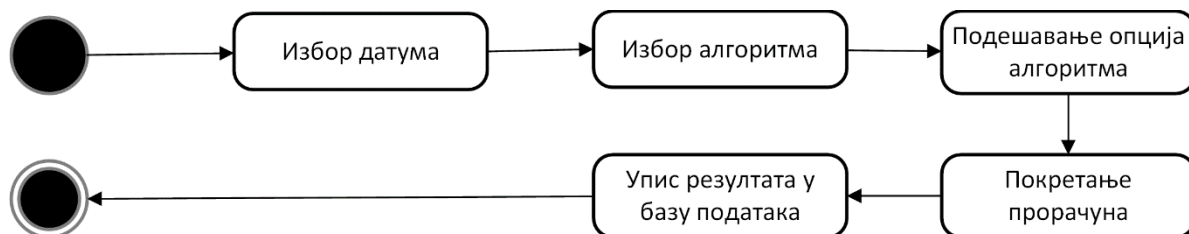


Слика 23 – Активност учитавања података из датотека

На основу назива, софтвер аутоматски препознаје тип датотеке, односно, да ли се ради о подацима временске прогнозе, временских остварења или остварене потрошње. Подаци из датотеке се парсирају и уписују у базу података.

### 7.1.2. Извршавање прорачуна

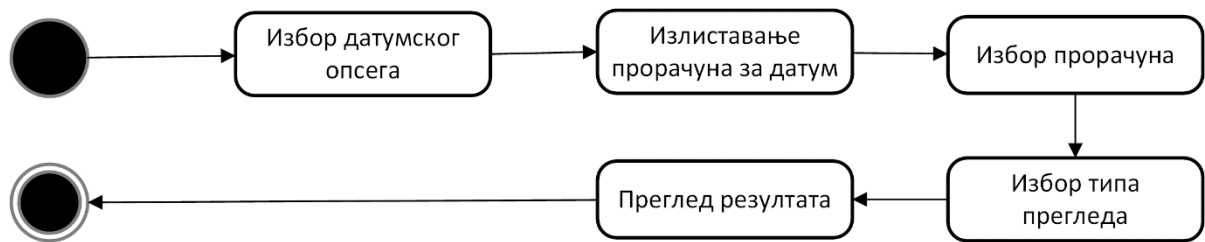
Први корак у процесу извршавања прорачуна (слика 24) јесте избор датума. Подразумевани избор датума јесте данашњи датум. Избором датума из прошлости сврха ове активности ће бити симулација прорачуна. Избором алгоритма бира се да ли се ради о прорачуну неуронском мрежом, линеарном регресијом или методом потпорних вектора. На основу изабраног алгоритма отвара се могућност подешавања одговарајућих опција алгоритма. Након покретања и извршавања прорачуна резултати прогнозе по сатима се уписују у базу података.



Слика 24 – Активност извршавања прорачуна

### 7.1.3. Преглед прорачуна

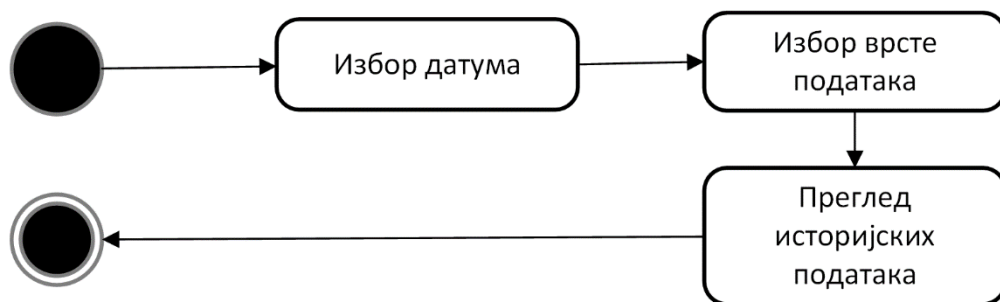
Након што су извршени, прорачуни прогнозе могу да се прегледају (слика 25). Резултати се прегледају у изабраном датумском опсегу. Тип прегледа може да буде табеларни приказ и приказ дијаграма.



Слика 25 – Активност прегледа резултата прорачуна

#### 7.1.4. Преглед историјских података

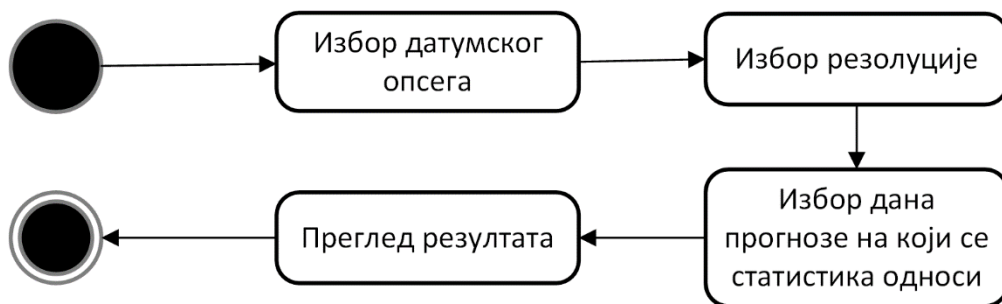
Историјски подаци (слика 26) могу да се прегледају за избрани датум из прошлости. Врста података који се прегледају може да буде временска прогноза, временска остварења и остварена потрошња.



Слика 26 – Активност прегледа историјских података

#### 7.1.5. Преглед статистичких података

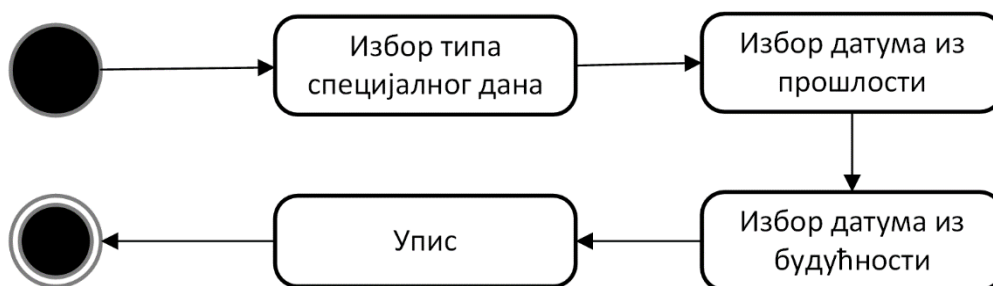
Статистички преглед (слика 27) подразумева преглед *MAPE* вредности, квадратног одступања и других статистичких података у одређеном временском периоду. Преглед статистичких података захтева избор датумског опсега. Након тога бира се резолуција: преглед може да буде на нивоу дана, недеље, месеца и године. Поред тога могуће је изабрати да ли се статистика односи на први дан прогнозе или на други дан прогнозе.



Слика 27 – Активност прегледа статистичких података

### 7.1.6. Подешавање специјалних дана

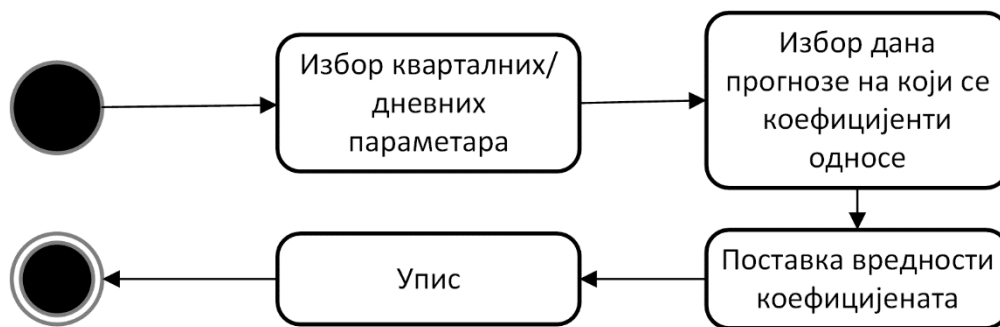
Као што је описано у поглављу 5.3. , услов за прогнозу потрошње за специјалне дане је дефинисање типова специјалних дана. Слика 28 приказује активност примене одређеног типа специјалних дана на конкретне датуме у прошлости и будућности.



Слика 28 – Подешавање специјалних дана

### 7.1.7. Поставка тежинских коефицијената

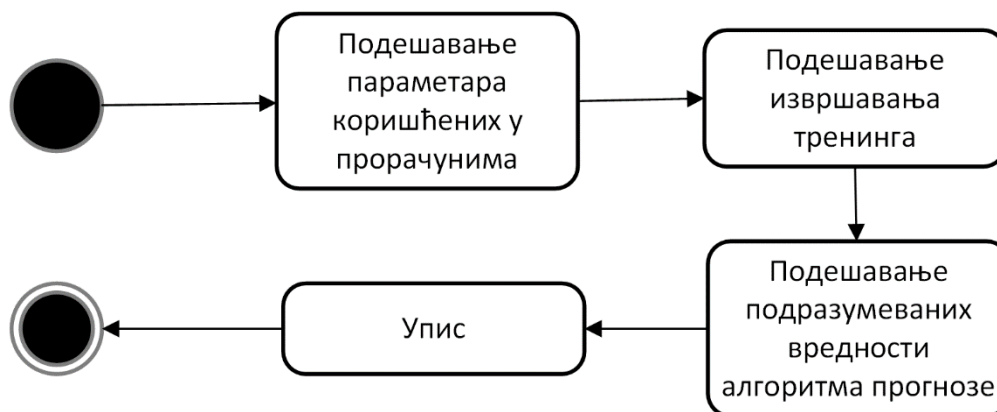
Тежински коефицијенти који се оптимизују преко генетског алгоритма могу да добију вредности и ручном поставком. Слика 29 показује да поставка вредности коефицијената може да се односи како на кварталне тако и на дневне параметре, а могуће је изабрати да ли се коефицијенти односе на први или на остале дане прогнозе.



Слика 29 – Активност поставке тежинских коефицијената

### 7.1.8. Подешавање апликације

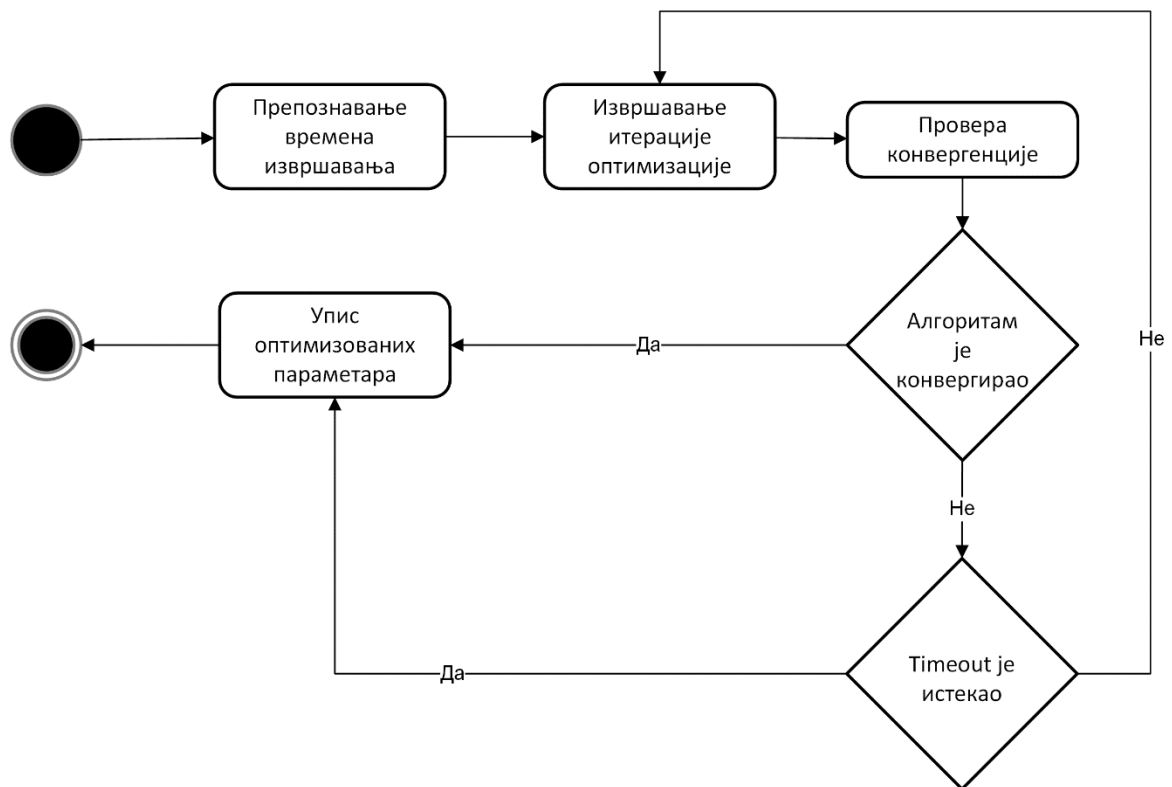
Подешавање апликације подразумева поставку вредности које се односе на целу апликацију (слика 30). То може бити подешавање параметара коришћених у прорачунима (осим тежинских коефицијената), подешавање извршавања тренинга (време покретања тренинга, *timeout*, итд.) и подешавање подразумеваних вредности за алгоритам прогнозе. Вредности алгоритма прогнозе односе се на сам тип алгоритма (неуронска мрежа, линеарна регресија или метод потпорних вектора) и на параметре који су специфични за сваки од ових алгоритама.



Слика 30 – Активност подешавања апликације

### 7.1.9. Дневна и квартална оптимизација

Активности дневне и кварталне оптимизације (слика 31) имају исти ток. Разлика је у времену позива и месту на ком се уписују оптимизовани подаци. Аутоматски *Daily Job* препознаје време у току дана када је потребно покренути оптимизацију. Након сваке извршене итерације генетског алгоритма, проверава се да ли је алгоритам конвергирао. Ако јесте, оптимизовани параметри се уписују у базу података. Ако није, проверава се да ли је истекао *timeout*. Ако је истекао, извршава се упис оптимизованих података у базу података, са вредностима које су постигнуте до момента *timeout*-а. Уколико није истекао, наставља се извршавање следеће итерације генетског алгоритма.

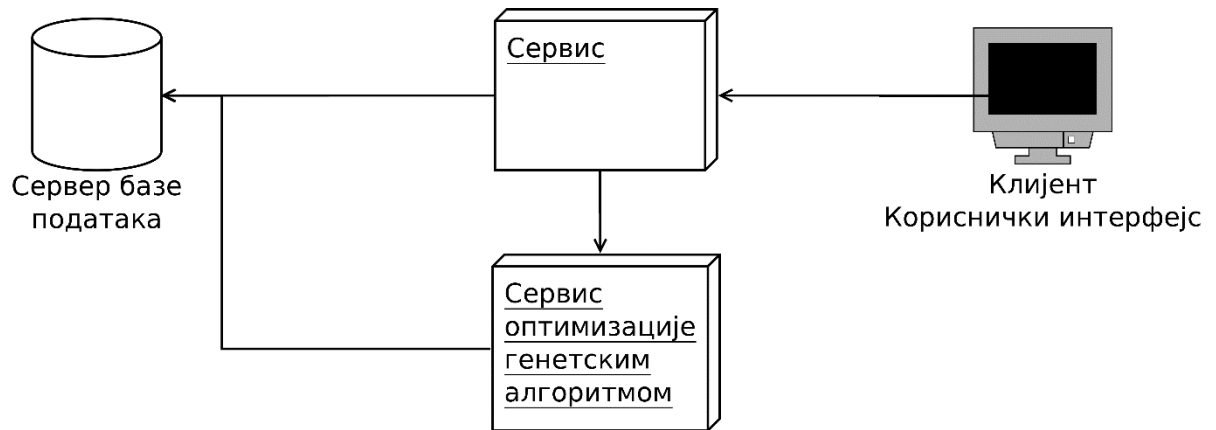


Слика 31 – Активност дневне/кварталне оптимизације

Исти поступак оптимизације извршава се за први дан прогнозе и за остале дане прогнозе.

## 7.2. Архитектура апликације

Апликативни систем прогнозе временских серија у инфраструктурним мрежама је креиран у четворослојној архитектури (слика 32). Клијентска апликација налази се на независним радним станицама и приступа сервисном делу апликације, на коме се налази пословна логика. Оптимизација генетским алгоритмом је сервис који има могућност да се налази на независном рачунару, у циљу одржавања стабилних перформанси и несметане оптимизације. Сервисни део и сервис оптимизације приступају бази података, у којој се налазе историјски подаци, подаци временске прогнозе, подаци о прогнозираној потрошњи и остали подаци апликативног система.

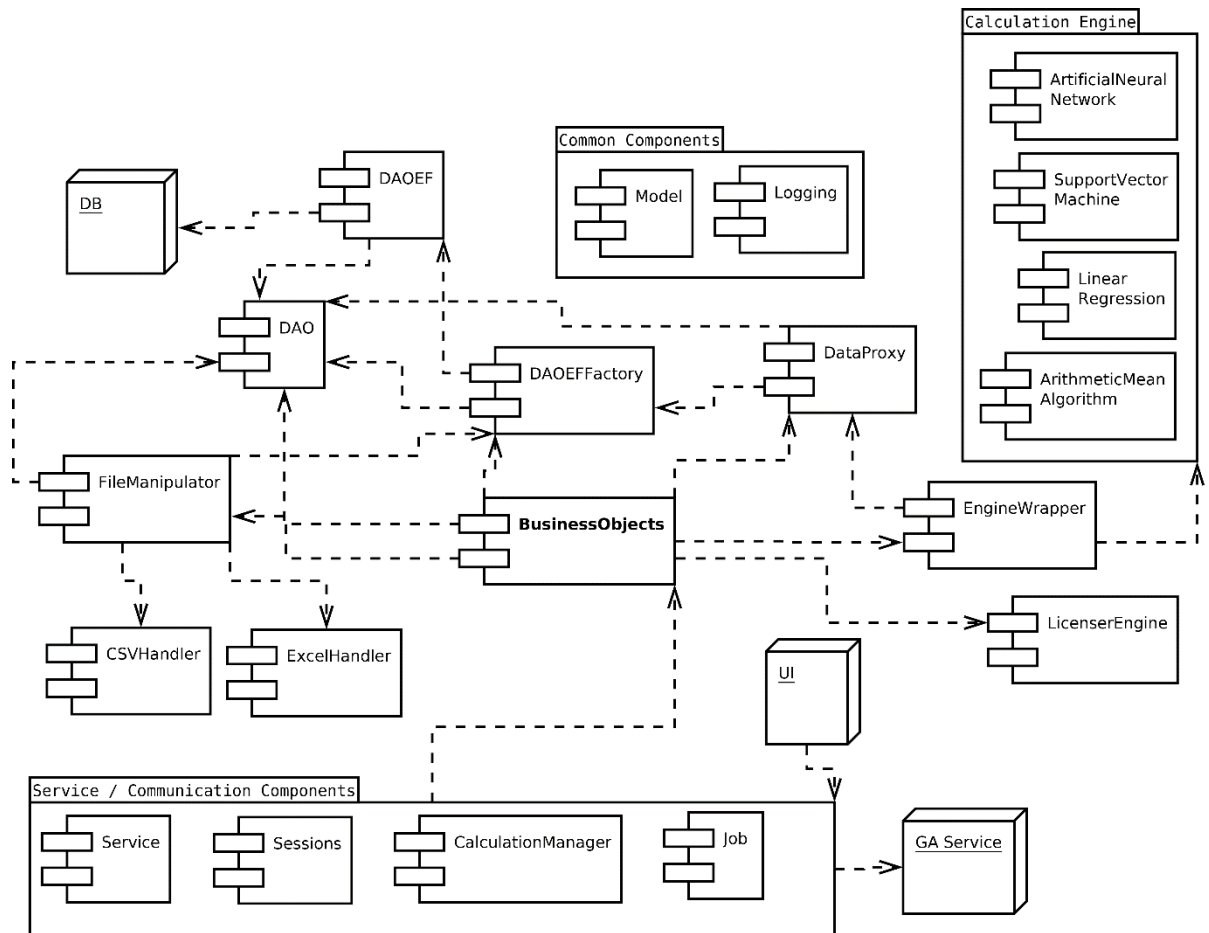


Слика 32 – Архитектура апликације

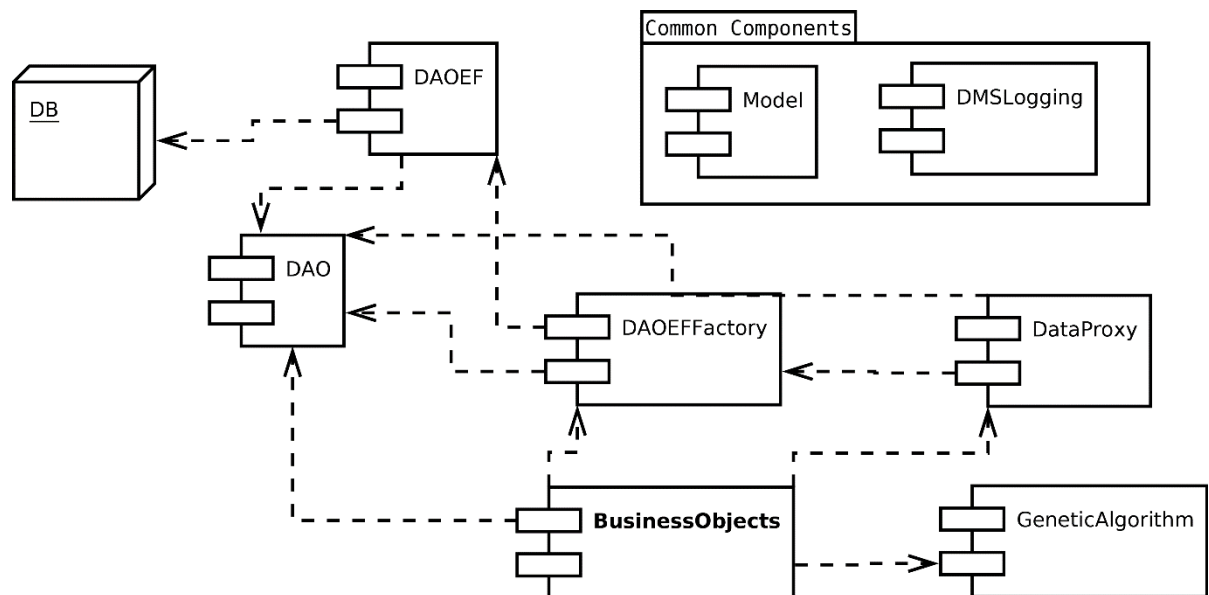
### 7.2.1. Компоненте система

Слика 33 приказује компоненте сервисног дела система. Овај подсистем комуницира са базом података, сервисом генетског алгоритма и корисничким интерфејсом. Комуникација са базом података извршава се преко компоненти приступа бази које су имплементирани коришћењем *Entity Framework* библиотеке (*DAO*, *DAOFactory* и *DAOEF*). *In-memory* база података се имплементира преко *DataProxy* компоненте. Увоз, извоз и манипулација датотекама извршавају се преко компоненти *FileManipulator*, *CSVHandler* и *ExcelHandler*. Компоненте које врше прорачун прогнозе су *ArtificialNeuralNetwork*, *SupportVectorMachine*, *LinearRegression* и *ArithmeticMeanAlgorithm*. Управљање, позивање и конфигурирање ових компоненти извршавају се кроз компоненту *EngineWrapper*. Објектни модел апликације, који се користи од стране целог апликативног система, налази се у компоненти *Model*. Компонента *BusinessObjects* извршава и координише пословну логику сервисног дела апликације. Комуникација са другим слојевима апликативног система одвија се кроз компоненте *Service*, *Sessions*, *CalculationManager* и *Job*. Компонента *Logging* је заједничка за све подсистеме и служи за логовање у апликацији. Компонента *LicenserEngine* задужена је за управљање лиценцама корисника апликације.





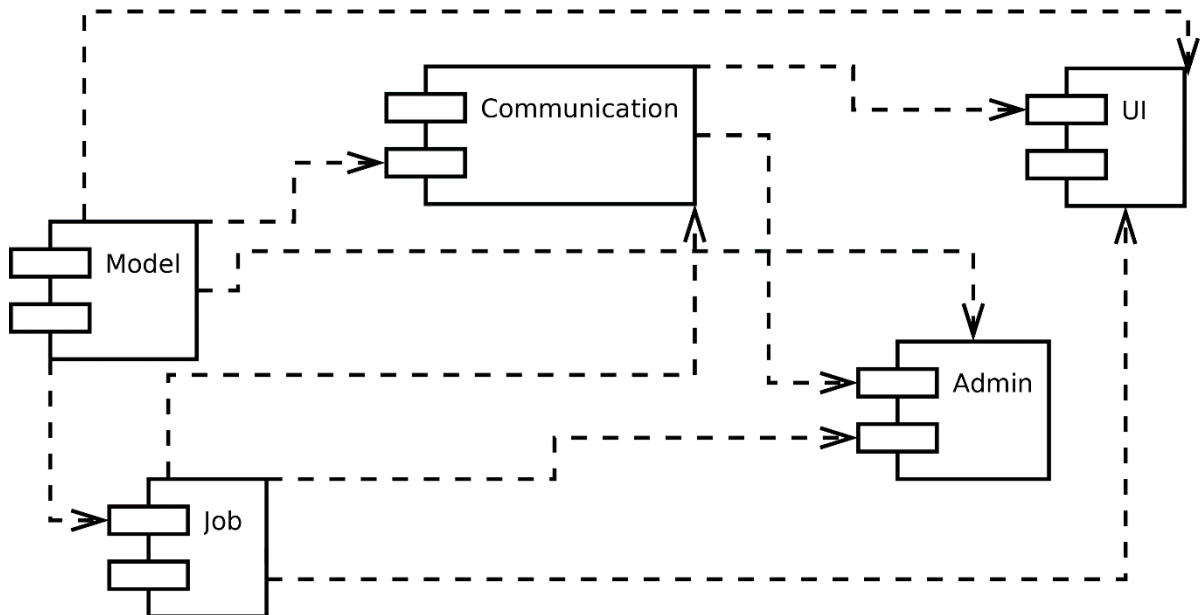
Слика 33 – Дијаграм компоненти сервисног дела апликативног система



Слика 34 – Дијаграм компоненти сервиса оптимизације генетским алгоритмом

Слика 34 приказује компоненте подсистема сервиса оптимизације генетским алгоритмом. Компоненте у овом подсистему садржане су у подсистему сервиса, али су

присутне само оне компоненте које су неопходне за оптимизацију генетским алгоритмом.



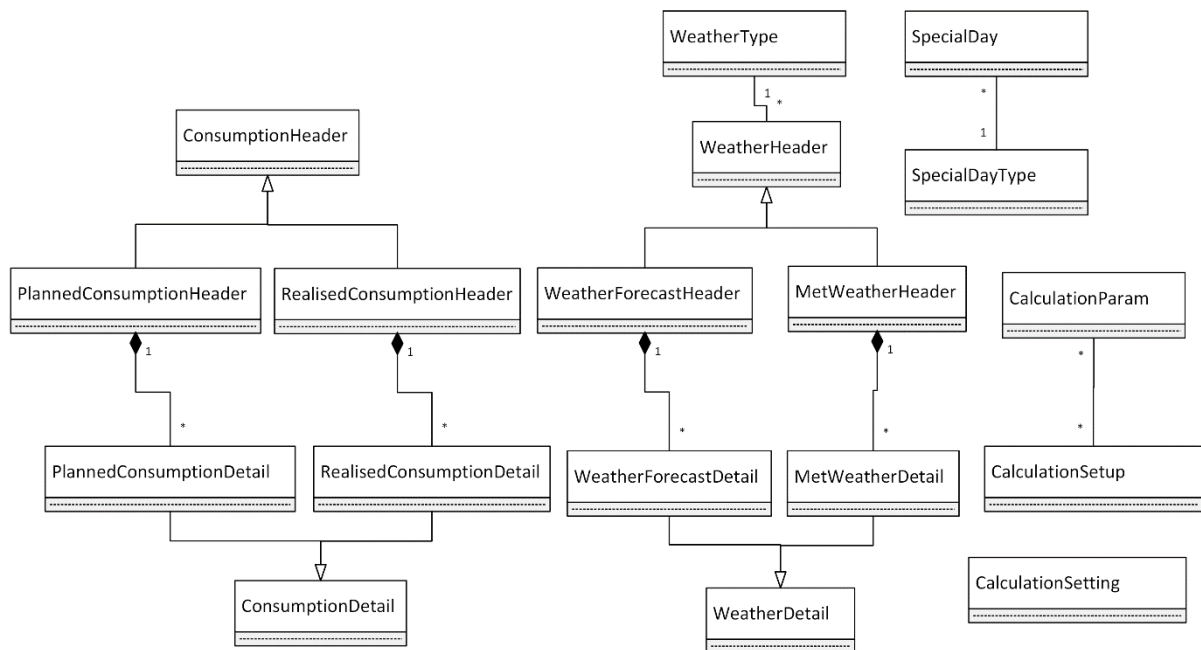
Слика 35 – Дијаграм компоненти клијентске апликације

Слика 35 приказује компоненте клијентске апликације. *UI* компонента намењена је свакодневном коришћењу апликације, док је компонента *Admin* намењена администрирању апликације. Компоненте *Communication* и *Job* служе за комуникацију са осталим подсистемима, док је компонента *Model* заједничка за цео апликативни систем.

### 7.3. Модел података

Слика 36 приказује модел података апликације, са класама које су релевантне за опис адаптивног модела прогнозе временских серија.

Сви подаци временских серија на сатном нивоу налазе се у објектима класа чији се назив завршава са „*Details*“. Тако се подаци о оствареним временским условима налазе у објектима класе *MetWeatherDetail*, а подаци о прогнозираним временским условима налази се у класи *WeatherForecastDetail*. Подаци о оствареној потрошњи ресурса налазе се у објектима класе *RealisedConsumptionDetails*, док се прорачунати подаци о потрошњи уписују у објекте класе *PlannedConsumptionDetails*. Класе које се односе на прогнозирану и остварену потрошњу у временским серијама наслеђују заједничку класу *ConsumptionDetail*, са којом деле заједничка обележја. Слично је и са класама које се односе на прогнозиране и остварене временске податке. Оне наслеђују класу *WeatherDetail*.



Слика 36 – Модел података апликације

Објекти *Header* класа представљају контејнере за *Detail* класе. *Header* класа садржи заједничке карактеристике за више *Detail* објеката, као што су датум увоза или прорачуна података, ознака да ли су подаци актуелни итд. Објекти *Header* класа временских података имају референцу на објекат класе *WeatherType* – тип временске прилике, чиме се означава да се сваки објекат односи на одређену временску прилику.

Типови специјалних дана налазе се у објектима класе *SpecialDayType*, док се конкретне појаве специјалних дана са утврђеним датумима налазе у објектима класе *SpecialDay*.

Параметри добијени оптимизацијом генетским алгоритмом описани су класама *CalculationParam* и *CalculationSetup*. *CalculationSetup* описује случај на који се оптимизација односи, као што је први дан прогнозе или други дан прогнозе.

## 7.4. Апликативни модели

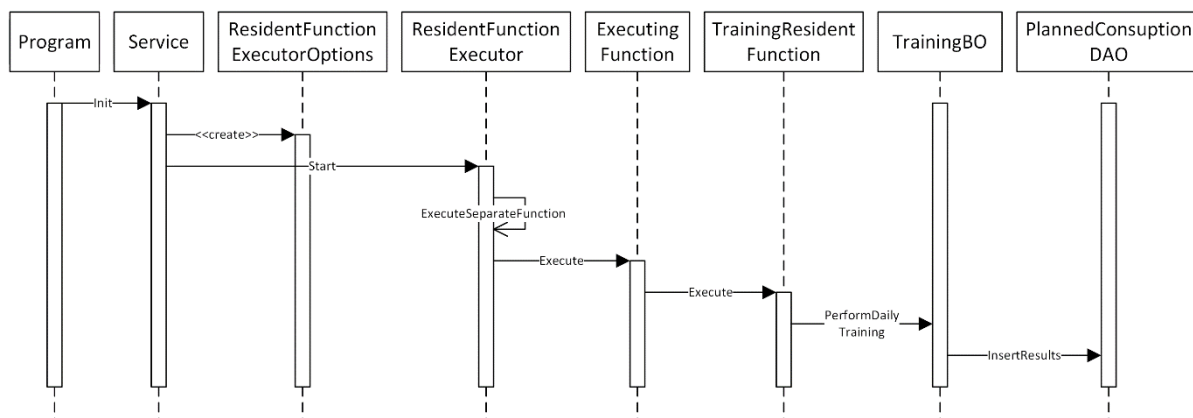
Свака приказана компонента састоји се од класа које све заједно чине велики јединствени апликативни модел. Овде ће бити представљени делови тог модела који су битни за приказ имплементације адаптивног модела прогнозе временских серија.



Слика 37 описује процес *припреме података за извршавање у алгоритму прогнозе*. Након издвајања одговарајућих сличних дана извршава се интерполација за евентуално недостајуће податке. *Substitutor* класе претварају податке дневне резолуције у податке сатне резолуције, како би сви подаци могли да буду употребљени у алгоритму прогнозе временске серије. Након тога сви подаци се валидирају. Уколико валидација није успешна, прорачун се не извршава. На крају се подаци конвертују у модел који је читљив за одговарајући алгоритам прогнозе. Објекат алгоритма прогнозе креира класа *CalculationEngineFactory*. Одговарајући *CalculationEngine*, који може да буде имплементација неуронске мреже, метода потпорних вектора, линеарне регресије или просте аритметичке средине, извршава калкулацију прогнозе над припремљеним подацима.

Пошто се различити подаци међу собом разликују, за интерполацију, валидацију и субституцију података постоје одговарајуће класе са сопственим имплементацијама ових метода.

Сликом 38 описује се механизам *резидентног извршавања оптимизације* параметара генетским алгоритмом. Самим покретањем апликације, кроз иницијализацију објекта сервиса апликације, иницијализују се и објекти задужени за покретање резидентног извршавања функција. То су објекти класа *ResidentFunctionExecutorOptions*, који служи за постављање опција резидентног извршавања функција, и *ResidentFunctionExecutor*, који служи за само резидентно покретање функција. Асинхроним методом *ExecuteSeparateFunction* издваја се објекат функције који је на реду да се изврши у одређеном моменту. Овај објекат се зове *ExecutingFunction* и служи као „покривач“ за конкретну имплементацију функције, у овом случају *TrainingResidentFunction*. Из ове класе, преко метода *Execute*, позива се метод *PerformDailyTraining* објекта класе *TrainingBO*, који садржи методе пословне логике који се односе на тренинг и оптимизацију у апликацији. Овај метод извршава оптимизацију која је описана у поглављу 6.4. и уписује резултате у базу података преко класе *PlannedConsumptionDAO*.

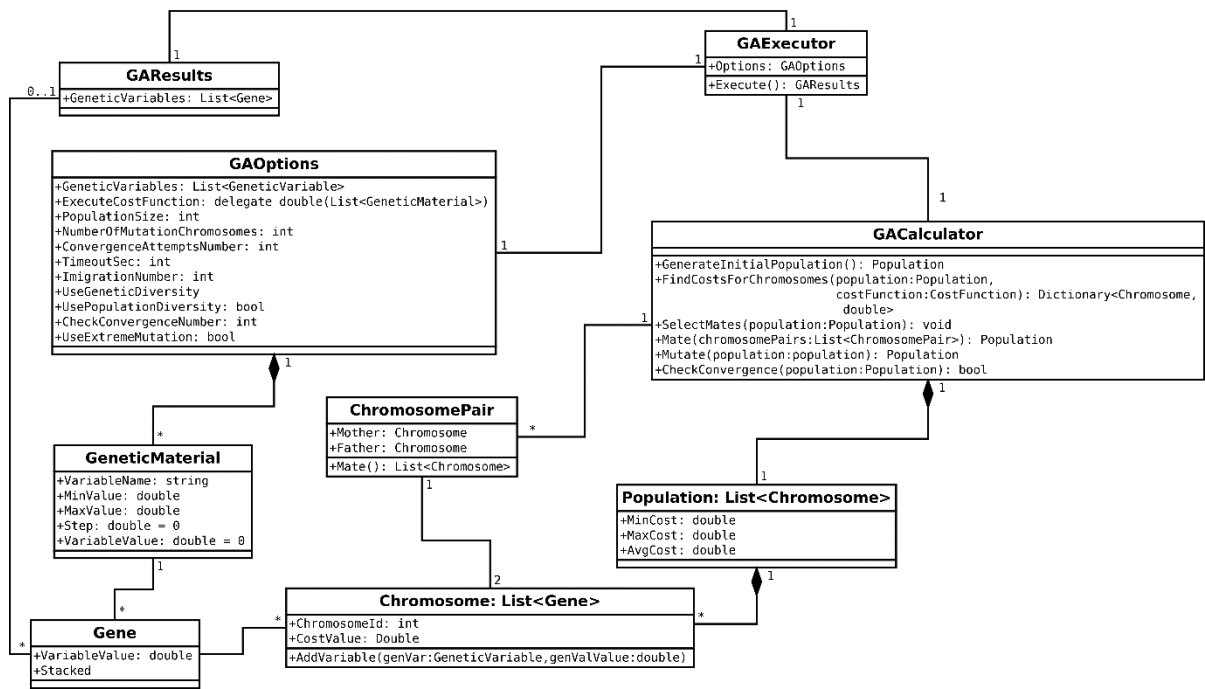


Слика 38 – Дијаграм секвенци за резидентно извршавање оптимизације

Слика 39 кроз дијаграм класа приказује **функционални модел генетског алгоритма**. Извршавање почиње од класе *GAExecutor*. У класи *GACalculator* извршавају се операције као што је креирање популације, селекција, укрштање, мутација и провера конвергенције. Ова класа садржи објекат *Population* који је заправо листа објеката типа *Chromosome*. Објекат класе *GACalculator* садржи објекте типа *ChromosomePair*, који су настали процесом селекције и користе се у процесу укрштања. Класа *Chromosome* је листа објеката класе *Gene*. Класа *Gene* садржи обележје о вредности варијабле коју представља, и динамичко обележје које означава да ли је ген „заглављен“, односно дошло је до конвергенције према неком од оптимума. Ова информација се користи за потребе мутације ка екстремима. Класа *GeneticMaterial* садржи информације о типовима гена, односно варијабли које се оптимизују. Ова класа има вредности о минималној и максималној вредности, кораку у оптимизацији и иницијалној вредности.

Објекат класе *GAOptions* се креира изван генетског алгоритма, коме се прослеђује. Овде се налазе параметри који ће утицати на начин извршавања генетског алгоритма, колекција објеката типа *GeneticMaterial*, као и критеријумска функција у форми делегата.

Резултати генетског алгоритма налазе се у објекту класе *GAResults*. Овде се налази листа гена са вредностима које су конвергирале тако да критеријумска функција враћа минималну вредност.



Слика 39 – Функционални дијаграм класа генетског алгоритма

# 8. Студије случаја

Студије случаја обухватиће примену предложеног модела прогнозе на примерима прогнозе потрошње електричне енергије и прогнозе потрошње природног гаса.

Студије случаја обухватају праћење процеса прогнозе и анализу резултата уз коришћење софтверске апликације која имплементира адаптивни модел прогнозе временских серија. Прогноза ће се извршити за две врсте ресурса: прогнозу потрошње електричне енергије у преносној мрежи за територију Србије без Косова, и за прогнозу потрошње природног гаса у дистрибутивној мрежи, за подручје једног региона у Холандији.

## 8.1. Прогноза потрошње електричне енергије у преносној мрежи

Пример прогнозе потрошње електричне енергије односи се на прогнозу у реалном систему који користи адаптивни модел прогнозе, за територију Србије без Косова и Метохије. Поступак прогнозе биће објашњен на примеру прогнозе за један типичан и један специјалан дан. Детаљан процес биће приказан за филтер сличних дана и метахеуристичку оптимизацију параметара. Као алгоритам прогнозе користиће се неуронска мрежа.

Јединица мере за количину дневне потрошње електричне енергије је MW/h.

### 8.1.1. Прогноза за типичан дан

Датум за који ће бити приказан процес прогнозе је 25. јануар 2021. године. Према експерименту, потрошња се прогнозира у поноћ између 24. и 25. јануара, и у том моменту су доступни сви подаци о претходној потрошњи, подаци о оствареним временским условима и временска прогноза за наредна 24 сата. Прогноза се извршава за само један дан, односно за временску серију од 24 сата.

#### 8.1.1.1. Оптимизација генетским алгоритмом

Критеријумска функција и варијабле које се оптимизују приказани су у поглављу 5.2. Параметри који се узимају у обзир приликом оптимизације генетским алгоритмом приказани су у табели 3. Параметри су у табели подељени у три групе: параметри временских прилика, инерциони параметри и остали параметри. За сваки параметар представљени су минимална и максимална вредност, као и корак промене. Временске прилике за које се оптимизују тежински коефицијенти су температура, ветар и



облачност. Инерциони параметри који се оптимизују су тежински коефицијенти инерције потрошње, термалне инерције и инерције последњег сата, затим дневни умањујући коефицијент, гранична вредност и број предисторијских дана. Поред тога оптимизују се тежински коефицијенти за дневну светлост и за близину дана и сумарни тежински коефицијенти за временске прилике и за инерцију.

Распон могућих вредности за све тежинске коефицијенте је од 0 до 100, са кораком 1. Дневни умањујући коефицијент може да има вредности у распону 0-1, као што је наведено у поглављу 5.1.2.2. Корак за овај коефицијент је 0,1.

Распон могућих вредности за граничну вредност је 1,1-3, при чему вредност 1,1 значи да и мало одступање од просечне стандардне девијације значи да датум није адекватан. Вредност 3 значи да само изузетно велико одступање од стандардне девијације треба да подразумева изbacивање дана из узорка сличних дана.

Табела 3 – Параметри који се оптимизују генетским алгоритмом у прогнози потрошње електричне енергије

<b>Параметри временских прилика</b>			
<b>Назив параметра</b>	<b>Минимална вредност</b>	<b>Максимална вредност</b>	<b>Корак промене</b>
Тежински коефицијент за температуру ( $w_{ct}$ )	0	100	1
Тежински коефицијент за ветар ( $w_{cwi}$ )	0	100	1
Тежински коефицијент за облачност ( $w_{csc}$ )	0	100	1
<b>Инерциони параметри</b>			
<b>Назив параметра</b>	<b>Минимална вредност</b>	<b>Максимална вредност</b>	<b>Корак промене</b>
Тежински коефицијент инерције потрошње ( $w_{cin}$ )	0	100	1
Тежински коефицијент термалне инерције ( $w_{cin}$ )	0	100	1
Тежински коефицијент инерције последњег сата ( $sc_{thin}$ )	0	100	1
Дневни умањујући коефицијент ( $ddwc$ )	0,00	1,00	0,01
Гранична вредност ( $L_{\sigma}$ )	1	3	0,1
Број предисторијских дана ( $M$ )	1	5	1
<b>Остали параметри</b>			
<b>Назив параметра</b>	<b>Минимална вредност</b>	<b>Максимална вредност</b>	<b>Корак промене</b>
Тежински коефицијент за дневну светлост ( $w_{cdd}$ )	0	100	1
Тежински коефицијент за близину дана ( $w_{c_{dp}}$ )	0	100	1
Тежински коефицијент за временске прилике ( $w_{cw}$ )	0	100	1
Тежински коефицијент за инерцију ( $w_{cin}$ )	0	100	1

Оптимизација се извршава за први дан и за остале дане прогнозе, док ће у овом примеру бити приказана оптимизација за први дан прогнозе.

Оптимизација се извршава на кварталном и на дневном нивоу. Оптимизација на кварталном нивоу извршава се последњег дана месеца јануара, априла, јула и октобра. Примена параметара оптимизованих кварталном оптимизацијом извршава се након завршетка кварталне оптимизације, која може да траје више дана. Дневна оптимизација се извршава сваку ноћ. Вредности кварталних и дневних параметара у моменту извршавања прорачуна прогнозе приказани су у табели 4.

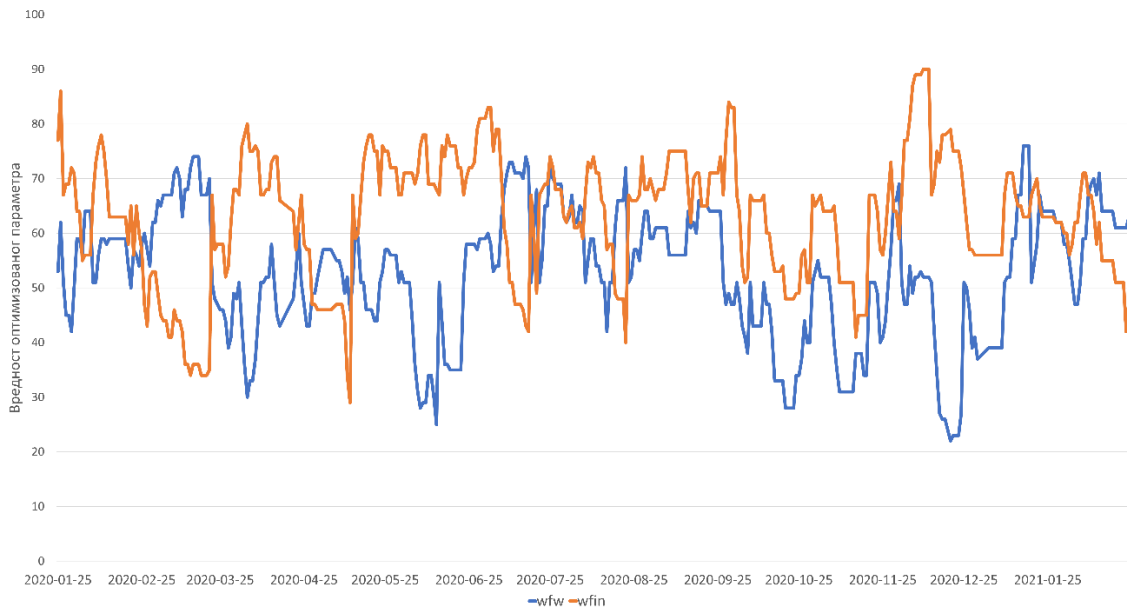
Табела 4 – Вредности параметара добијених кварталном оптимизацијом у прогнози потрошње електричне енергије

Параметри временских прилика		
Назив параметра	Вредност кварталног параметра	Вредност дневног параметра
Тежински коефицијент за температуру ( $wc_t$ )	55	80
Тежински коефицијент за ветар ( $wc_{wi}$ )	33	65
Тежински коефицијент за облачност ( $wc_{sc}$ )	49	18
Инерциони параметри		
Назив параметра	Вредност кварталног параметра	Вредност дневног параметра
Тежински коефицијент инерције потрошње ( $wc_{cin}$ )	52	72
Тежински коефицијент термалне инерције ( $wc_{tin}$ )	38	18
Тежински коефицијент инерције последњег сата ( $sc_{thin}$ )	44	32
Дневни умањујући коефицијент ( $ddwc$ )	0,15	0,15
Гранична вредност ( $L_{\sigma}$ )	1,7	1,7
Број предисторијских дана ( $M$ )	2	2
Остали параметри		
Назив параметра	Вредност кварталног параметра	Вредност дневног параметра
Тежински коефицијент за дневну светлост ( $wc_{dd}$ )	45	52
Тежински коефицијент за близину дана ( $wc_{dp}$ )	58	29
Тежински коефицијент за временске прилике ( $wc_w$ )	51	53
Тежински коефицијент за инерцију ( $wc_{in}$ )	67	77

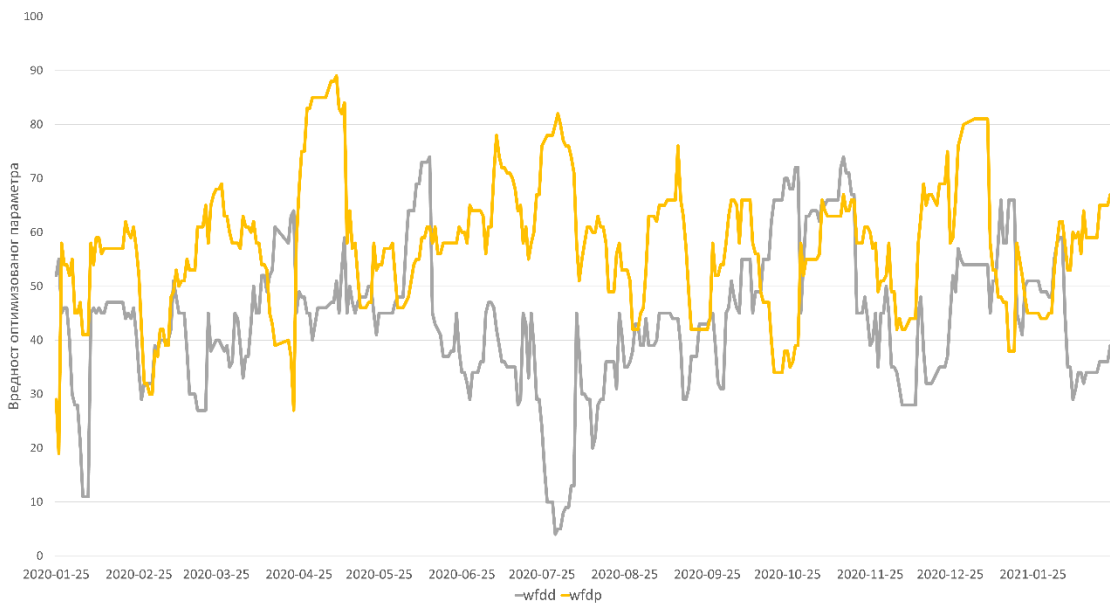
#### 8.1.1.1.1. Дневна оптимизација током времена

Кретање вредности дневних параметара током времена за период од једне године пре тренутка прогнозе приказано је на сликама 40 и 41. За репрезентативан пример изабрани су параметри Тежински коефицијент за временске прилике ( $wc_w$ ), Тежински коефицијент за инерцију ( $wc_{in}$ ), Тежински коефицијент за дневну светлост ( $wc_{dd}$ ) и Тежински коефицијент за близину дана ( $wc_{dp}$ ). Очигледно је да вредност тежинског коефицијента за дневну светлост опада у деловима године са изузетно високим, односно, изузетно ниским температурама. Разлог се може тражити у чињеници да је при екстремним временским условима примат у утицају имају временски и инерциони

фактори. Међутим, кретање вредности оптимизованих параметара, а самим тим и кретање утицаја појединачних фактора на потрошњу, може се анализирати кроз посебан рад који је изван опсега истраживања дисертације.



Слика 40 – Кретање вредности дневних параметара током времена (a)



Слика 41 – Кретање вредности дневних параметара током времена (b)

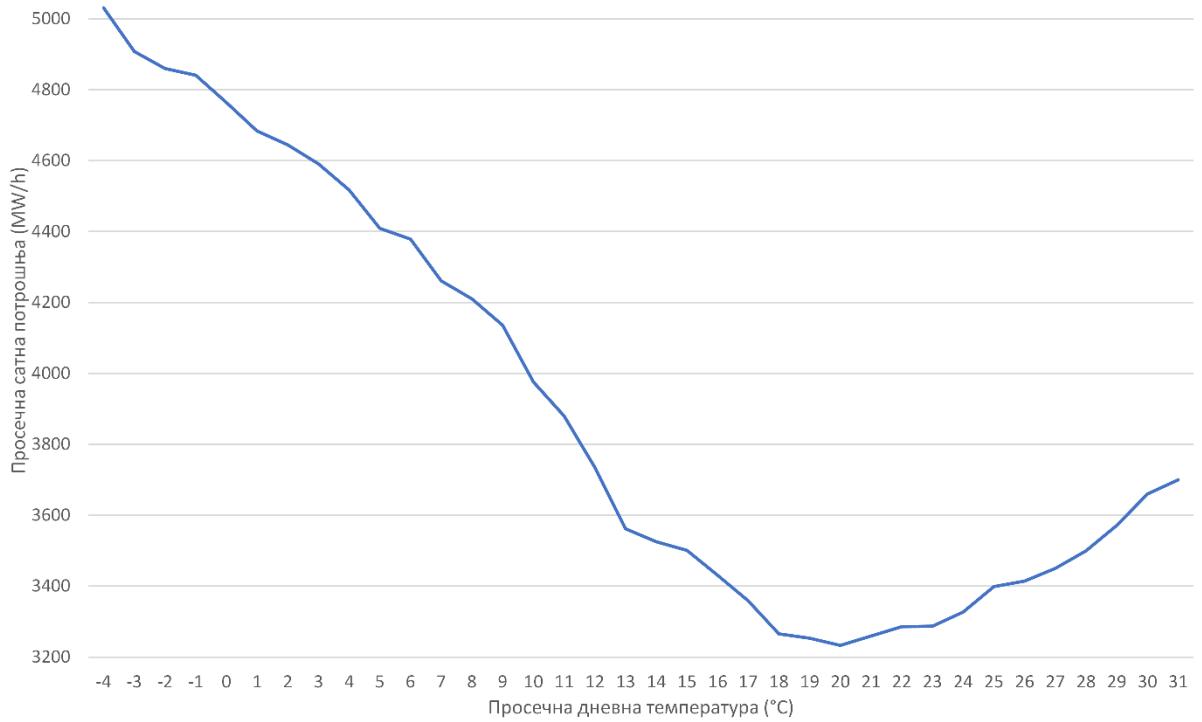
### 8.1.1.2. Селекција фактора од утицаја и партиципационих коефицијената

Историјска база података садржи метеоролошке податке и податке о потрошњи за задњих седам година. Метеоролошки подаци који су на располагању су сатне вредности температуре, облачности и брзине ветра. Поред њих, у модел су уведени фактори

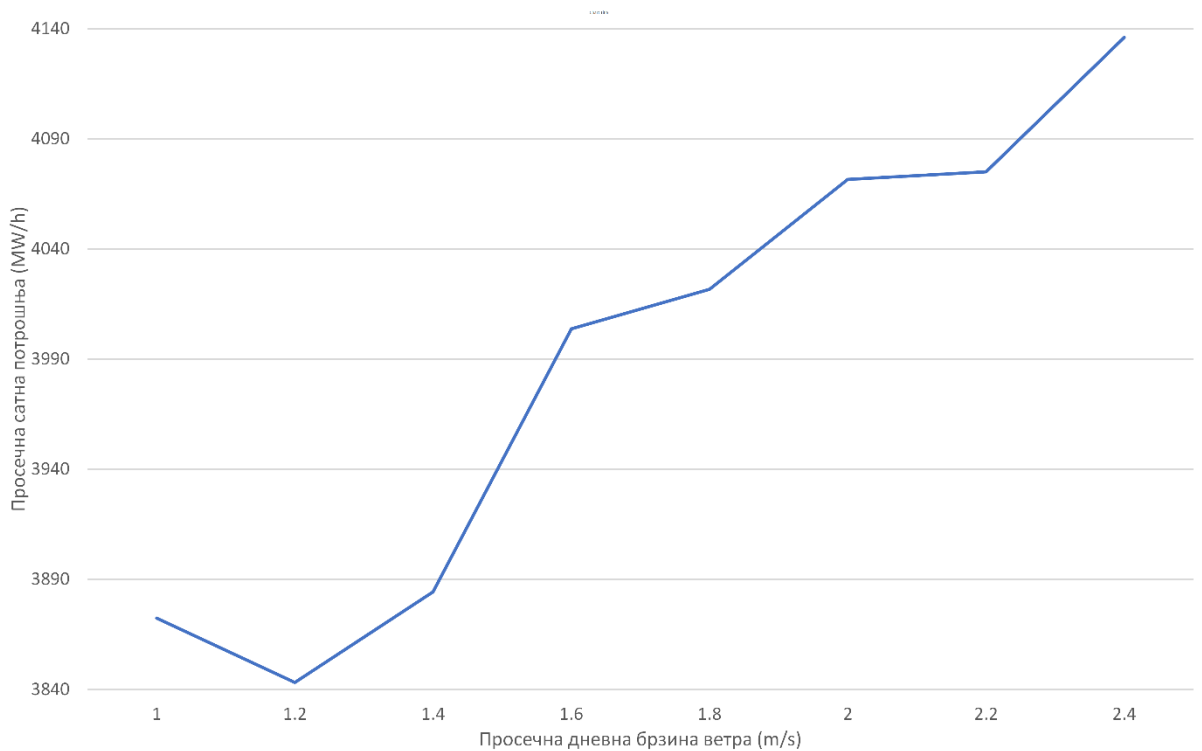
дужина дневног светла и удаљеност дана. За ове факторе потребно је утврдити да заиста имају утицај на промену потрошње, и одредити партиципационе коефицијенте за сваки појединачни фактор. Кретање просечне сатне потрошње на основу промене температуре приказано је на слици 42. Очигледно је да потрошња опада до одређене вредности температуре приближно линеарно, док након те тачке потрошња расте, али нешто слабијим интензитетом. Разлог пада потрошње налази се у смањењу потребе за загревањем домаћинства и постројења са растом температуре, док се разлог раста потрошње налази у повећању потребе за коришћењем климатизационих уређаја са растом температуре. Тачка прелома између пада и раста потрошње је 20 °С. Када се изврши регресиона линеарна анализа посебно пада и раста потрошње, тренд просечне промене потрошње за промену од једног степена целзијуса је 70 MW/h. Ову вредност означавамо са  $A1$ .

Слике 43 и 44 приказују кретање просечне сатне потрошње на основу промене просечне дневне брзине ветра. На слици 43 налазе се подаци који су измерени при температури до 20 °С, док се на слици 44 налазе подаци који су измерени при температури преко 20 °С. Граница од 20 °С узета је на основу тачке прелома у којој потрошња престаје да пада и почиње да расте (слика 42). На сликама се примећује да при температурама до 20 °С потрошња расте са порастом брзине ветра, док при температурама преко до 20 °С са порастом брзине ветра потрошња опада. Регресивном линеарном анализом добија се тренд просечне промене потрошње од 210 MW/h за промену од једног m/s. Ова вредност се означава са  $A2$ .

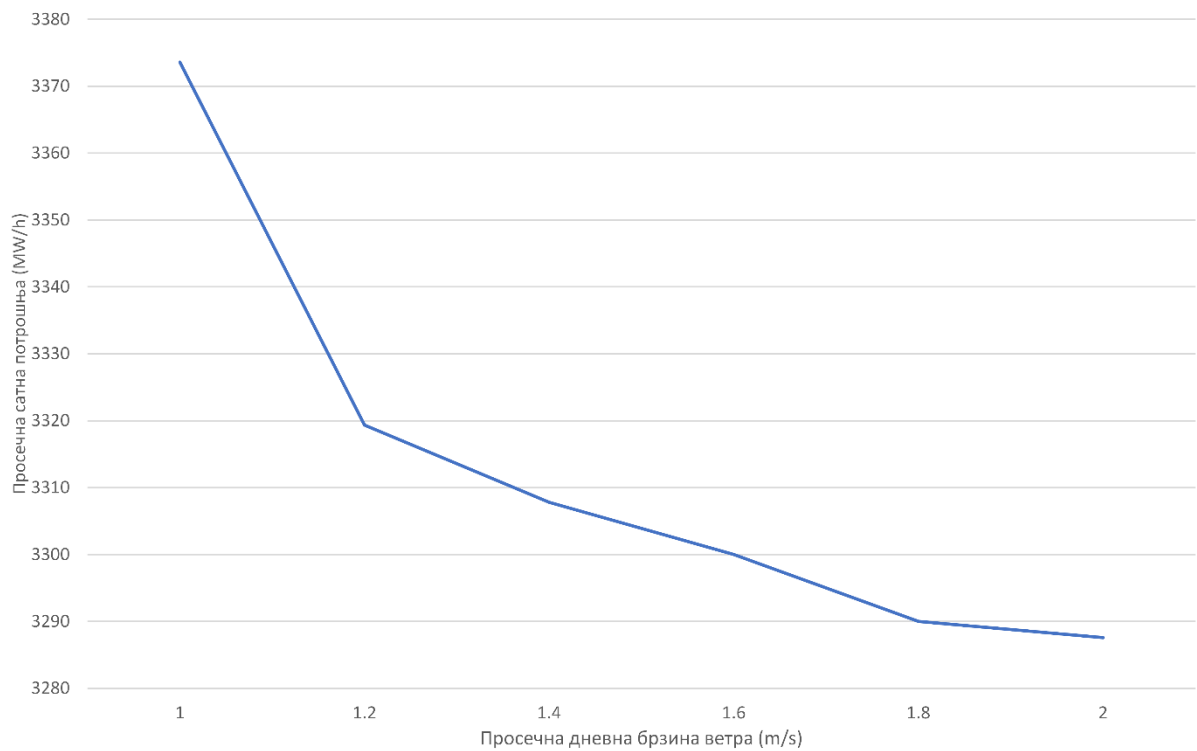
На основу објашњења из поглавља 5.1.2.1.1. партиципациони коефицијент за брзину ветра износи  $p_{c_{wi}}=A2/A1=210/70=3$ .



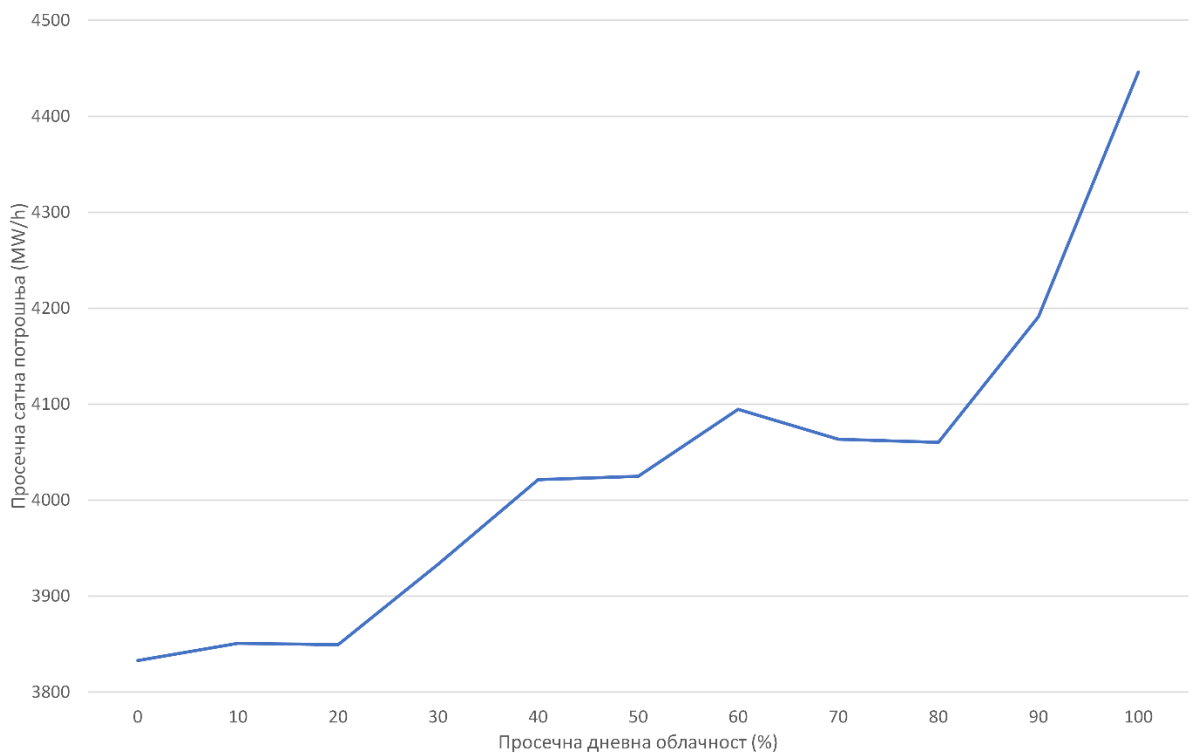
Слика 42 – Историјски тренд промене потрошње електричне енергије на основу промене температуре



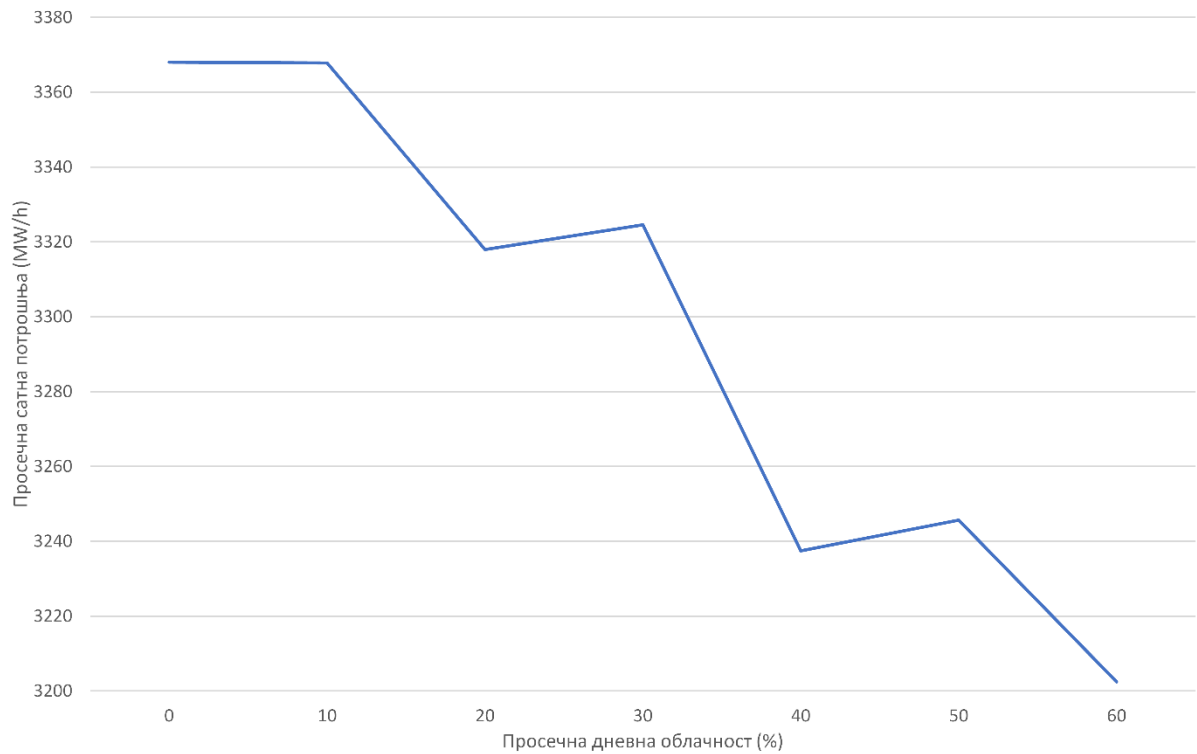
Слика 43 – Историјски тренд промене потрошње електричне енергије на основу промене брзине ветра при температури до 20 °C



*Слика 44 – Историјски тренд промене потрошње електричне енергије на основу промене брзине ветра при температури преко 20 °C*

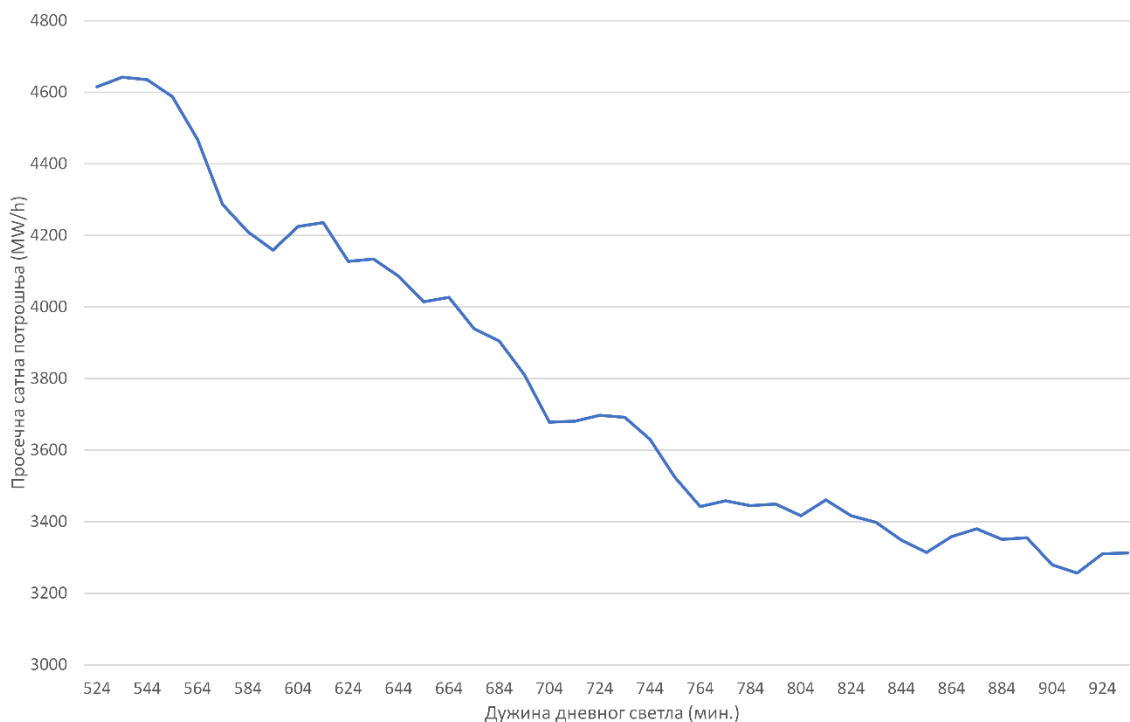


*Слика 45 – Историјски тренд промене потрошње електричне енергије на основу промене облачности при температури до 20 °C*



*Слика 46 – Историјски тренд промене потрошње електричне енергије на основу промене облачности при температури преко 20 °C*

На слици 45 види се кретање просечне сатне потрошње на основу просечне дневне облачности при просечним дневним температурама до 20 °C, док се на слици 46 налазе подаци о кретању просечне сатне потрошње на основу просечне дневне облачности при температурама преко 20 °C. Очигледно је да при температурама до 20 °C потрошња расте са повећањем облачности, док при температурама преко 20 °C са повећањем облачности потрошња опада. Регресивном линеарном анализом добија се тренд просечне промене потрошње од  $A2=7,22$  MW/h по промени од 1% просечне облачности. Партиципациони коефицијент за облачност износи  $pc_{sc}=A2/A1=7,22/70 \approx 0,10$ .

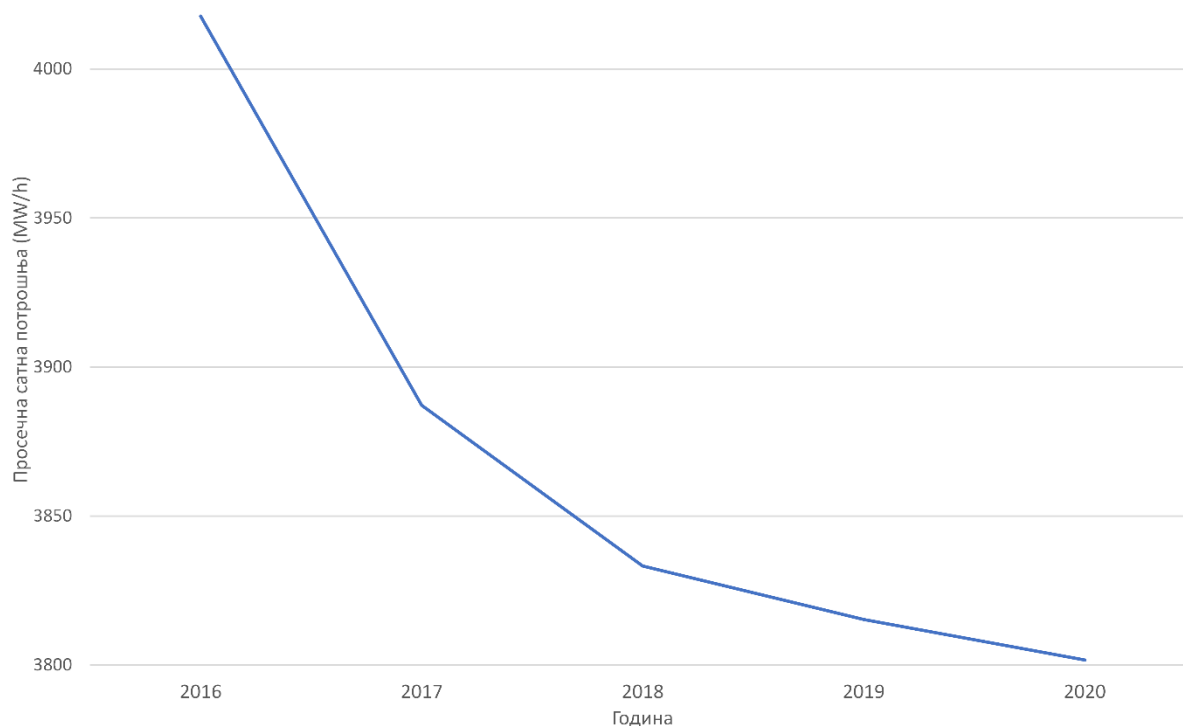


*Слика 47 – Историјски тренд промене потрошње електричне енергије на основу промене трајања дневног светла*

Слика 47 приказује кретање криве потрошње на основу промене трајања дневног светла. Очигледно је да се потрошња смањује са повећањем дужине дневног светла. Регресивном линеарном анализом утврђено је да просечно сатно смањење потрошње износи  $A2=3,45 \text{ MW/h}$  по повећању дневне светлости од једног минута. На основу објашњења из поглавља 5.1.2.3. партиципациони коефицијент за дневно светло износи  $pc_{ad}=A2/A1=3,45/70 \approx 0,05$ .

Промена просечне сатне потрошње у различитим годинама приказано је на слици 48. Очигледно је да се просечна дневна потрошња дугорочно смањује са протоком времена. Регресивном линеарном анализом утврђено је да просечно сатно смањење потрошње износи  $50 \text{ MW/h}$  годишње. Праћењем резултата прогнозе коришћењем историјских сличних дана утврђено је да на релевантност сличних дана утичу и промене у облицима дневних кривих потрошње кроз време. На основу анализе утицаја протока времена на криву потрошње, утврђен је тренд просечне промене потрошње  $A2=70 \text{ MW/h}$  по протоку једне године, односно  $0,19 \text{ MW/h}$  по протоку једног дана.





Слика 48 – Историјски тренд промене потрошње електричне енергије по годинама

На основу објашњења из поглавља 5.1.2.4. партиципациони коефицијент за близину дана износи  $pc_{dp}=A2/A1=0,19/70\approx 0,002738$ .

На основу објашњења из поглавља 5.1.2.2.1. партиципациони коефицијент за инерцију потрошње представља инверзну вредност  $A1$ :  
 $pc_{cin}=1/ A1=1/70\approx 0,014286$ .

Сви партиципациони коефицијенти за факторе који утичу на сличност приказани су у табели 5.

Табела 5 – Фактори од утицаја за прогнозу потрошње електричне енергије и њихови партиципациони коефицијенти

Фактор	Јединица мере фактора	Вредност партиципационог коефицијента
Температура	°C	1,00
Ветар	m/s	1,14
Облачност	%	0,10
Трајање дневне светлости	минут	0,05
Близина дана	дан	0,002738
Потрошња ресурса	MW/h	0. 014286

### 8.1.1.3. Релевантни параметри прогнозираног дана

Дан у недељи за који се прогнозира потрошња је понедељак. Дан припада грејној сезони. Дужина трајања дневне светлости је 572 минута (приближно 9.5 сати).

Табела 6 приказује дневне податке временске прогнозе релевантне за прогнозу потрошње.

Табела 6 – Дневни подаци временске прогнозе за прогнозирани дан

Просечна температура [°C]	Минимална Температура [°C]	Максимална Температура [°C]	Просечна брзина ветра [m/s]
3,6	1,2	11,5	3,7

Табела 7 приказује сатне податке временске прогнозе који су релевантни за прогнозу потрошње: вредности температуре, брзине ветра и облачности.

Табела 7 – Сатни подаци временске прогнозе за прогнозирани дан

Сат	Температура [°C]	Брзина ветра [m/s]	Облачност [%]
00:00	7	5	100
01:00	6	5	100
02:00	5	5	100
03:00	4	4	100
04:00	4	4	100
05:00	3	4	100
06:00	3	4	100
07:00	2	4	50
08:00	2	4	50
09:00	2	3	50
10:00	3	3	0
11:00	4	3	0
12:00	5	2	0
13:00	6	2	0
14:00	6	2	50
15:00	5	2	50
16:00	5	2	50
17:00	4	2	100
18:00	3	2	100
19:00	3	2	100
20:00	2	2	100
21:00	2	2	100
22:00	1	2	100
23:00	1	2	100

Табела 8 показује инерционе податке у односу на прогнозирани дан, односно податке о просечној температури, и просечној потрошњи за три дана који претходе прогнозираном дану. Последња колона приказује потрошњу у последњем сату претходног дана, и ова информација је доступна само за први дан који претходи прогнозираном.

Табела 8 – Инерциони подаци у односу на прогнозирани дан

Датум	Просечна температура	Просечна потрошња	Потрошња у последњем сату
2021-01-24	11,3	4083,88	4238
2021-01-23	11,6	4345,25	
2021-01-22	10,8	4607,92	

#### 8.1.1.4. Селекција података иницијалним филтером

Иницијалним филтером издвајају се подаци над којима ће бити примењен филтер сличности. Услови филтера сличности за прогнозирани дан су да селектовани дани имају исти дан у недељи, да припадају грејној сезони и да је просечна дневна температура плус/минус 5 степени у односу на прогнозирану просечну температуру дана за који се прогнозира потрошња.

Табела 9 – Дани селектовани иницијалним филтером

Датум	Мин. темп.	Макс. темп.	Прос. темп	Прос. вет.	Дуж. дн. свет.	Прос. потр. претх. дана	Прос. темп. претх. дана	Потрошња у посл. сату претх. дана
2021-01-11	-0.8	0.3	-0.3	1.5	545	4302	0.29	4694
2021-01-04	7.0	9.9	8.2	2.9	535	3888	8.00	4156
2020-12-21	3.4	8.4	5.1	3.6	527	4398	3.75	4522
2020-12-14	3.3	4.5	3.9	1.4	530	4450	5.75	4477
2020-12-07	3.0	6.9	5.3	6.8	535	4398	9.71	4352
2020-11-30	0.0	4.8	2.2	1.3	544	4637	0.88	4554
2020-11-23	1.0	8.1	4.4	1.3	557	4221	3.54	4384
2020-11-09	4.0	6.3	5.3	1.1	588	3965	8.00	4014
2020-03-23	-2.4	4.2	0.9	2.5	740	3766	13.46	4001
2020-03-16	1.0	13.7	7.1	4.6	718	4018	9.29	4458
2020-03-09	4.6	9.2	6.8	1.3	696	4162	8.63	4335
2020-01-27	2.4	7.2	3.6	1.7	575	4738	4.29	4627
2020-01-13	-1.9	-0.6	-1.4	1.0	547	4673	3.21	4943
2019-12-30	-2.8	-0.4	-1.5	1.8	531	4576	1.92	4892
2019-12-23	6.0	11.6	7.2	4.2	528	4159	14.67	4196
2019-12-02	2.2	8.6	4.6	3.8	543	4123	6.83	4400
2019-03-11	4.1	14.8	8.6	2.3	699	3763	13.54	3817
2019-02-25	-1.0	6.5	1.7	1.3	657	4598	-1.46	4815
2019-02-18	1.6	14.3	7.8	1.2	635	4269	4.46	4466
2019-02-11	3.8	12.0	7.8	3.2	615	4456	4.13	4563
2019-02-04	3.8	8.0	5.1	1.8	596	4383	11.83	4344
2019-01-28	1.4	7.4	4.3	3.8	578	4838	-3.00	4804
2019-01-21	0.5	2.6	1.2	1.7	563	4593	2.13	4662
2019-01-14	2.2	8.1	4.5	2.6	549	4878	-1.50	4899
2018-12-24	2.2	14.2	7.4	2.9	528	4676	8.00	4675
2018-12-10	3.2	9.4	6.0	2.1	533	4567	6.29	4490
2018-12-03	4.0	11.3	7.6	2.3	541	4760	-1.79	4640
2018-11-19	0.7	5.5	2.4	2.0	566	4236	4.13	4474
2018-03-19	-1.5	2.6	0.4	3.2	725	3992	12.50	4451
2018-03-05	-2.0	5.8	1.9	4.4	682	4783	0.29	4838
2018-02-19	-0.8	3.4	1.2	2.2	639	4329	5.25	4668
2018-02-12	0.4	7.7	4.0	1.9	619	4443	1.83	4639
2018-02-05	-0.8	7.1	2.6	1.2	599	4298	3.83	4480
2018-01-29	-1.1	13.0	5.6	1.2	581	4527	4.25	4755
2018-01-22	-0.6	3.2	1.2	2.1	564	4424	2.88	4704
2017-12-25	2.8	15.0	8.5	1.3	528	4800	4.38	4741
2017-12-18	-0.4	2.6	0.8	2.0	528	4450	5.08	4705
2017-12-11	4.8	11.9	8.6	5.4	531	4532	4.50	4666
2017-12-04	-1.0	3.7	1.0	1.4	539	4523	1.83	4615
2017-11-27	1.3	4.9	2.9	2.0	550	4188	9.50	4375
2017-11-20	2.5	9.1	5.1	1.9	563	4294	6.67	4299
2017-10-30	5.0	12.6	8.3	2.6	616	3906	10.33	3807
2017-03-13	3.9	10.4	6.7	2.6	708	4320	6.08	4384
2017-02-20	-0.6	4.4	1.7	1.3	642	4619	0.83	4872
2017-02-13	-5.1	4.0	-1.6	3.4	622	4809	-0.50	5057

Дани који су изабрани иницијалним филтером и над којима ће бити примењен филтер сличности приказани су у табели 9. Уз дане, у овој табели су приказане особине које су релевантне за поступак прогнозе: минимална, максимална и просечена температура, просечна брзина ветра, дужина дневне светлости, просечна сатна потрошња у претходном дану, просечна температура претходног дана и потрошња у последњем сату претходног дана.

### 8.1.1.5. Пример прорачуна коефицијента сличности

Пример прорачуна коефицијента сличности у односу на дан за који се прогнозира потрошња биће приказан за један од дана из табеле 9: 2019-12-02 (други децембар 2019. године).

#### 8.1.1.5.1. Прорачун коефицијента временских прилика

Сатни подаци о временским приликама за изабрани датум приказани су у табели 10.

Табела 10 – Сатни подаци о временским приликама за датум 2019-12-02

Сат	Температура [°C]	Брзина ветра [m/s]	Облачност [%]
00:00	3	5	50
01:00	4	5	50
02:00	4	6	50
03:00	4	5	50
04:00	4	5	50
05:00	4	6	50
06:00	4	6	50
07:00	4	6	100
08:00	4	5	100
09:00	5	5	100
10:00	6	4	100
11:00	7	5	100
12:00	8	6	100
13:00	9	3	100
14:00	8	4	100
15:00	8	1	100
16:00	5	1	100
17:00	5	2	100
18:00	4	2	100
19:00	3	3	100
20:00	3	3	100
21:00	3	2	100
22:00	2	2	100
23:00	2	2	100

На основу података из табеле 10 и из табеле 7 потребно је израчунати коефицијент временских прилика. На основу формуле (7), формула за коефицијент сличности за сатну температуру је:

$$sc_t = \sqrt{\frac{\sum_{i=1}^{24} (t_{d_i} - t_{(d-n)_i})^2}{24}}, \quad (22)$$

где је:

$sc_t$  – коефицијент сличности за температуру,

$t_{d_i}$  – прогнозирана температура у сату  $i$ , у дану за који се прогнозира потрошња (2021-01-25),

$t_{(d-n)_i}$  – остварена температура у сату  $i$ , у дану који је кандидат за изабран и сличан дан (2019-12-02).

Табела 11 – Прорачун коефицијента сличности за температуру за датум 2019-12-02

Сат ( $i$ )	$t_{d_i}$	$t_{(d-n)_i}$	$(t_{d_i} - t_{(d-n)_i})^2$
00:00	7	3	16
01:00	6	4	4
02:00	5	4	1
03:00	4	4	0
04:00	4	4	0
05:00	3	4	1
06:00	3	4	1
07:00	2	4	4
08:00	2	4	4
09:00	2	5	9
10:00	3	6	9
11:00	4	7	9
12:00	5	8	9
13:00	6	9	9
14:00	6	8	4
15:00	5	8	9
16:00	5	5	0
17:00	4	5	1
18:00	3	4	1
19:00	3	3	0
20:00	2	3	1
21:00	2	3	1
22:00	1	2	1
23:00	1	2	1
$\frac{\sum_{i=1}^{24} (t_{d_i} - t_{(d-n)_i})^2}{24}$			3,96
		$sc_t$	1,99

Табела 11 приказује упоредне сатне вредности температуре за два дана која се пореде, квадратну разлику тих вредности у сваком сату, просек квадратне разлике и коначну вредност коефицијента сличности за температуру  $sc_t = 1,99$ .

На сличан начин се израчунавају коефицијенти сличности за брзину ветра и облачност. Табела 12 приказује податке за доступне факторе временских прилика: израчунате

квадратне девијације за температуру, брзину ветра и облачност, затим партиципационе коефицијенте, израчунате коефицијенте сличности, као и тежинске коефицијенте за сваки од фактора.

Табела 12 – Доступни подаци за факторе временских прилика за датум 2019-12-02

Фактор ( <i>i</i> )	$\sqrt{\frac{\sum_{i=1}^{24} (w_{d_i} - w_{(d-n)_i})^2}{24}}$	Партиципациони коефицијент за фактор ( <i>pc<sub>i</sub></i> )	Вредност коефицијента сличности за фактор ( <i>sc<sub>i</sub></i> )	Тежински коефицијент за фактор ( <i>wc<sub>i</sub></i> )
Температура ( <i>t</i> )	1,99	1,00	<i>sc<sub>t</sub></i> = 1,99	<i>wc<sub>t</sub></i> = 80
Ветар ( <i>w<sub>i</sub></i> )	1,44	1,14	<i>sc<sub>w<sub>i</sub></sub></i> = 1,64	<i>wc<sub>w<sub>i</sub></sub></i> = 65
Облачност ( <i>sc</i> )	54,96	0,10	<i>sc<sub>sc</sub></i> = 5,50	<i>wc<sub>sc</sub></i> = 18

На основу формуле (7), коефицијент временских прилика се израчунава на следећи начин:

$$sc_w = \frac{1,99 \times 80 + 1,64 \times 65 + 5,50 \times 18}{80 + 65 + 18} = 2,24.$$

#### 8.1.1.5.2. Прорачун инерционог коефицијента

Табела 13 приказује инерционе податке за датум 2019-12-02: просечну температуру и просечну потрошњу за претходна три дана, као потрошњу у последњем сату првог претходног дана.

Табела 13 – Инерциони подаци у односу на датум 2019-12-02

Датум	Просечна температура	Просечна потрошња	Потрошња у последњем сату
2019-12-01	2,80	4133,00	4400
2019-11-30	6,40	4123,17	
2019-11-29	11,50	4137,96	

На основу формуле (8) и на основу података из табеле 8 и табеле 13, коефицијент инерције потрошње се израчунава на следећи начин:

$$sc_{cin} = \frac{|4083,88 - 4133,00| \times 0,15^0 + |4345,25 - 4123,17| \times 0,15^1 + |4607,92 - 4137,96| \times 0,15^2}{0,15^0 + 0,15^1 + 0,15^2} \times 0,14286 = 1,13.$$

Примена формуле (9) на податке о просечним температурама из табела 8 и 13 изводи се на следећи начин:

$$sc_{tin} = \frac{|11,30 - 2,80| \times 0,15^0 + |11,60 - 6,40| \times 0,15^1 + |11,80 - 11,50| \times 0,15^2}{0,15^0 + 0,15^1 + 0,15^2} = 7,92.$$

Инерција последњег сата израчунава се применом формуле (10) на податке о потрошњи у последњем сату из табела 8 и 13:

$$sc_{hin} = |4238 - 4400| \times 0,14286 = 2,31.$$

Табела 14 – Доступни подаци за инерционе факторе

Фактор	Вредност коефицијента сличности за фактор	Тежински коефицијент за фактор ( $w_{c_i}$ )
Инерција потрошње ( $cin$ )	$sc_{cin} = 1,13$	$w_{c_t} = 72$
Термална инерција ( $tin$ )	$sc_{tin} = 7,92$	$w_{c_{wi}} = 18$
Инерција последњег сата ( $lhin$ )	$sc_{lhin} = 2,31$	$w_{c_{sc}} = 32$

Табела 14 приказује доступне податке за инерционе факторе: вредности коефицијента сличности и тежинске коефицијенте за инерцију потрошње, термалну инерцију и инерцију последњег сата.

Коначни инерциони коефицијент израчунава се на основу формуле (11):

$$sc_{in} = \frac{72 \times 1,13 + 18 \times 7,92 + 32 \times 2,31}{72 + 18 + 32} = 2,44.$$

#### 8.1.1.5.3. Прорачун коефицијента трајања дневне светлости

Дужина трајања дневног светла за датум 2019-12-02 је 538 минута. На основу формуле (12) израчунава се коефицијент трајања дневне светлости:

$$sc_{ad} = |572 - 538| \times 0,05 = 1,7.$$

#### 8.1.1.5.4. Прорачун коефицијента близине дана

Дан 2019-12-02 удаљен је од дана за који се прогнозира потрошња (2021-01-25) 419 дана. Применом формуле (13) израчунава се коефицијент близине дана:

$$419 \times 0,002738 = 1,15.$$

#### 8.1.1.5.5. Прорачун резултујућег коефицијента сличности

Табела 15 садржи израчунате подкоефицијенте, заједно са тежинским коефицијентима сваког подкоефицијента.



Табела 15 – Вредности подкоефицијената и њихови тежински коефицијенти

Подкоефицијент сличности	Вредност	Тежински коефицијент
Коефицијент временских прилика ( $sc_w$ )	2,24	53
Инерциони коефицијент ( $sc_{in}$ )	2,44	77
Коефицијент трајања дневне светлости ( $sc_{dd}$ )	1,70	52
Коефицијент близине дана ( $sc_{dp}$ )	1,15	29

На основу ових података и формуле (14) израчунава се резултујући коефицијент сличности за дан 2019-12-02, у односу на дан за који се прогнозира потрошња (2021-01-25):

$$s_c = \frac{53 \times 2,24 + 77 \times 2,44 + 52 \times 1,70 + 29 \times 1,15}{53 + 77 + 52 + 29} = 2,03.$$

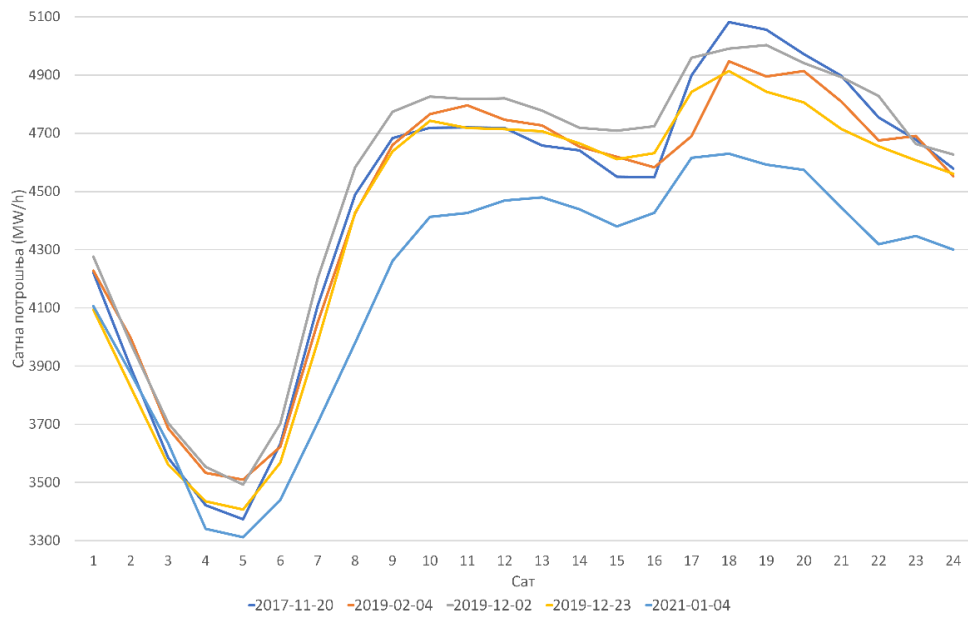
Вредност резултујућег коефицијента сличности за датум 2019-12-02 рангираће тај дан у односу на остале дане. На основу ранга одређује се да ли ће датум бити изабран за филтрирање у филтеру девијације потрошње.

У овом примеру број дана који се бирају је  $z=5$ . Овај број је изабран због прегледности примера и због тога што је итеративним тестирањем утврђено да пет дана даје довољно добре резултате у овом конкретном случају.

#### 8.1.1.6. Филтер девијације потрошње

Табела 16 садржи израчунате коефицијенте сличности за све датуме изабране иницијалним филтером. Поред резултујућег коефицијента сличности ( $sc$ ), табела садржи и подкоефицијенте сличности за временске прилике ( $sc_w$ ), инерцију ( $sc_{in}$ ), трајање дневне светлости ( $sc_{dd}$ ) и близину дана ( $sc_{dp}$ ). Дани су сортирани по резултујућем коефицијенту сличности. За филтер девијације потрошње изабрано је пет дана са највећом вредности резултујућег коефицијента сличности.

Криве сатних вредности потрошње за иницијално изабране дане приказане су на слици 49. Очигледно је да 4 од 5 дана има сличне криве потрошње, док крива једног дана (2021-01-04) изразито одступа од осталих. Разлог овоме може бити то да тај дан има одређене специфичности које нису могле бити препознате постојећим алгоритмом. У овом конкретном случају, у питању је дан који је номинално радни дан, међутим у великој мери запослени су користили слободне дане, пошто су се тај и следећи дан налазили између два нерадна празника. Због тога је крива потрошње тог дана имала особине нерадног дана. Задатак филтера девијације потрошње је да препозна овакав дан, да га одстрани из узорка и на његово место стави следећи дан са највећом вредности коефицијента сличности.



Слика 49 – Криве сатних вредности потрошње електричне енергије за иницијални скуп изабраних дана

Табела 16 – Израчунати коефицијенти за изабране датуме

Датум	$sc_w$	$sc_{in}$	$sc_{dd}$	$sc_{dp}$	$sc$
2019-02-04	4,23	1,16	0,80	1,97	1,95
2019-12-02	2,24	2,44	1,70	1,15	2,03
2019-12-23	3,35	2,27	1,47	1,09	2,18
2017-11-20	3,97	2,47	0,30	3,18	2,41
2021-01-04	4,04	3,36	1,23	0,06	2,55
2020-11-23	4,08	4,13	0,5	0,17	2,68
2017-11-27	3,96	3,10	0,73	3,16	2,74
2020-12-07	6,49	2,17	1,23	0,13	2,74
2020-12-21	2,44	6,56	1,50	0,10	3,39
2020-11-09	5,19	5,32	0,53	0,21	3,40
2020-03-09	4,97	2,94	4,13	0,88	3,46
2018-12-10	3,75	5,29	1,30	2,13	3,49
2018-11-19	3,62	6,25	0,20	2,18	3,54
2019-02-11	3,15	5,98	1,43	1,95	3,59
2019-02-18	5,65	4,19	2,10	1,94	3,73
2020-12-14	4,74	6,10	1,40	0,11	3,78
2020-01-27	3,89	7,58	0,10	1,00	3,91
2017-03-13	3,77	3,74	4,53	3,87	3,96
2018-02-05	4,57	6,02	0,90	2,97	3,97
2020-03-16	4,28	4,57	4,87	0,86	4,06
2017-10-30	3,65	6,57	1,47	3,24	4,12
2018-12-24	3,54	7,93	1,47	2,09	4,43
2018-03-19	4,76	4,49	5,10	2,86	4,48
2020-11-30	5,02	8,76	0,93	0,15	4,71
2018-12-03	3,65	9,39	1,03	2,15	4,89
2020-03-23	4,44	6,32	5,60	0,84	4,92
2018-02-12	3,44	9,02	1,57	2,95	4,95
2019-01-21	4,99	9,33	0,30	2,01	5,01
2018-01-22	4,61	9,52	0,27	3,01	5,11
2018-02-19	5,44	7,79	2,23	2,93	5,16
2017-12-25	5,38	8,69	1,47	3,09	5,31
2021-01-11	5,96	10,05	0,90	0,04	5,39
2017-12-18	4,69	9,48	1,47	3,10	5,43
2017-12-11	5,02	9,37	1,37	3,12	5,45
2019-01-28	2,33	12,52	0,20	1,99	5,48
2017-12-04	5,38	9,44	1,10	3,14	5,50
2018-01-29	5,12	10,64	0,30	2,99	5,65
2019-03-11	5,11	9,17	4,23	1,88	5,93
2019-01-14	3,38	13,52	0,77	2,03	6,25
2019-12-30	6,65	12,60	1,37	1,07	6,75
2019-02-25	5,03	12,67	2,83	1,92	6,85
2020-01-13	7,44	13,33	0,83	1,03	7,08
2017-02-20	5,65	12,98	2,33	3,93	7,27
2018-03-05	4,42	13,37	3,67	2,89	7,29
2017-02-13	4,76	16,35	1,67	3,95	8,12

Стандардна девијација потрошње за сваки од изабраних дана рачуна се коришћењем формуле (15). Ове вредности приказане су у табели 17. У истој табели приказана је и сума стандардних девијација.

Табела 17 – Стандардна девијација потрошње електричне енергије за иницијални скуп изабраних дана

Датум	Стандардна девијација ( $\sigma_i$ )
2019-02-04	84,20
2019-12-02	126,26
2019-12-23	42,72
2017-11-20	84,20
2021-01-04	236,48
$\sum_{i=1}^z \sigma_i$	550,93

Највећу вредност стандардне девијације има датум 2021-01-04, и за њега се израчунава допринос девијацији потрошње  $C_\sigma$ , према формули (15):

$$C_\sigma = \frac{5 \times 236,48}{550,93} = 2,15.$$

Проверава се да ли је вредност већа од граничне вредности  $L_\sigma$ , која је одређена оптимизацијом генетским алгоритмом, и има вредност 1,7. Пошто  $C_\sigma$  има већу вредност од  $L_\sigma$ , датум 2021-01-04 се замењује са следећим датумом са најнижом вредности коефицијента сличности ( $sc$ ). То је датум 2020-11-23, чија вредност  $sc$  је 2,68. Ова вредност не прелази 200% најмање вредности  $sc$  изабраног скупа дана, тако да је дан одговарајући.

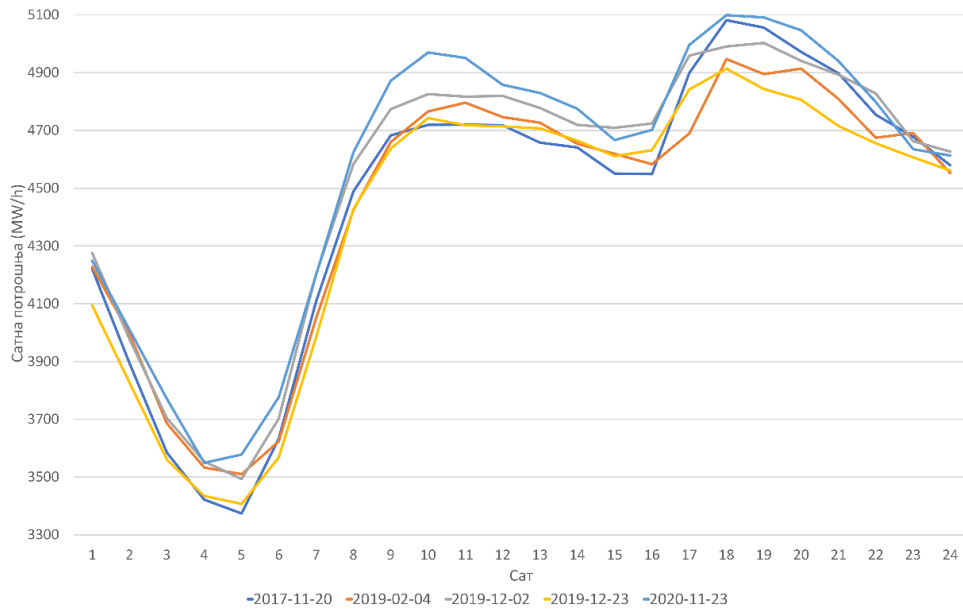
Табела 18 – Стандардна девијација за нови скуп изабраних дана

Датум	Стандардна девијација ( $\sigma_i$ )
2019-02-04	58,95
2019-12-02	59,26
2019-12-23	52,40
2017-11-20	85,15
2020-11-23	97,04
$\sum_{i=1}^z \sigma_i$	352,80

Нови скуп дана, са замењеним датумом чија је потрошња одступала, приказан је у табели 18. Криве сатних вредности потрошње за изабране дане приказане су на слици 50. У овом скупу дана, највећу вредност стандардне девијације има датум 2020-11-23, тако да се за њега рачуна допринос девијацији потрошње  $C_\sigma$ :

$$C_\sigma = \frac{5 \times 97,04}{352,80} = 1,38.$$

Вредност израчунатог доприноса девијацији потрошње мањи је од граничне вредности  $L_\sigma$ , тако да се датум задржава, а скуп дана који је изабран представља и коначни скуп дана који ће се користити у алгоритму прогнозе.



Слика 50 – Криве сатних вредности потрошње електричне енергије за нови скуп изабраних дана

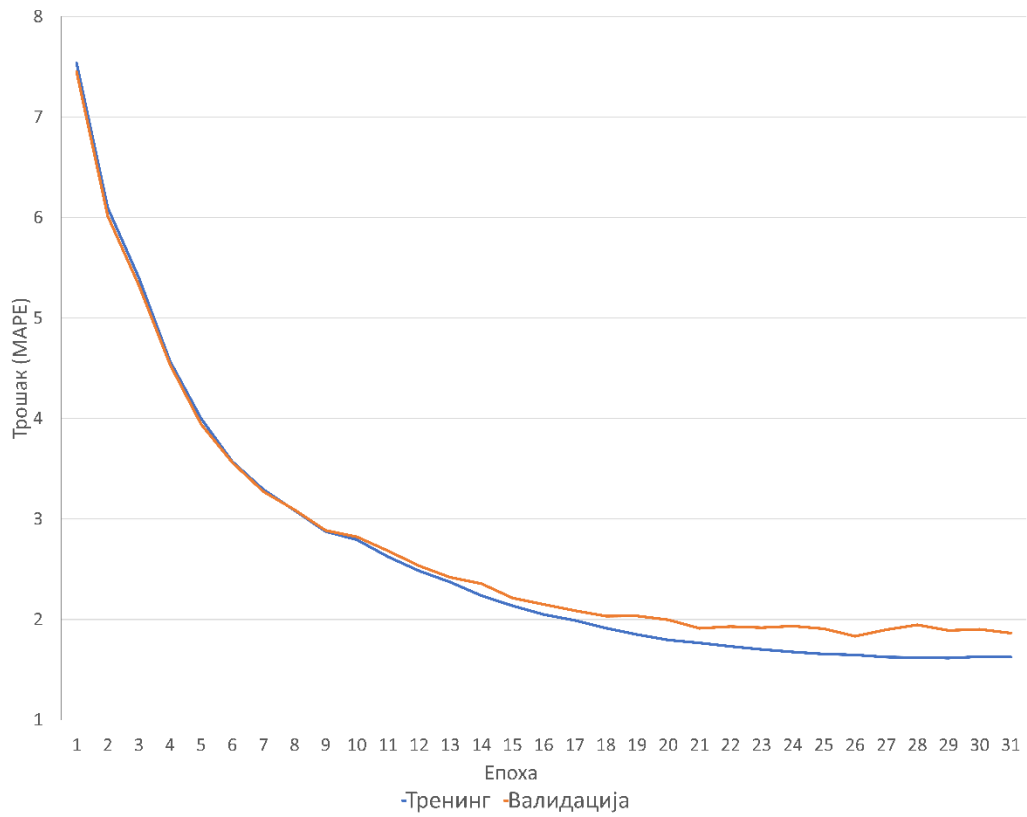
### 8.1.1.7. Прорачун прогнозе потрошње електричне енергије неуронском мрежом

Селектовани дани се користе као улаз у алгоритам прогнозе. Прогноза се извршава за сваки сат изабраног дана. Изабрани алгоритам прогнозе је *feed-forward* неуронска мрежа. Број узорака као улаз у тренинг неуронске мреже, односно број сличних дана, може да се подешава ручно за сваки прогнозирани дан.

За имплементацију неуронске мреже коришћен је *Python* програмски језик и *Keras* библиотека [123]. Као независне променљиве коришћени су подаци о температури, инерцији оптерећења и трајању дневног светла.

Вредности хипер параметара неуронске мреже оптимизоване су континуалним праћењем кривих учења у узорцима [124]. Сваки случај који се прати обухвата тренинг и валидацију са различитим вредностима хипер параметара. Хипер параметри који се оптимизују су: број скривених слојева, број неурона у скривеним слојевима, број епоха, број пакета, оптимизатор, активациона функција, као и улазне колоне. Из филтера сличних дана селектован је подједнак број дана за тренинг и за валидацију. Криве учења пореде се са кривама валидације и уочавају се криве које теже минималном трошку и минималном одступању криве учења од криве валидације.

Крива учења изабране неуронске мреже за један сат из сета података за тренинг приказана је на слици 51. Изабрана функција трошка је просечна апсолутна процентуална грешка.



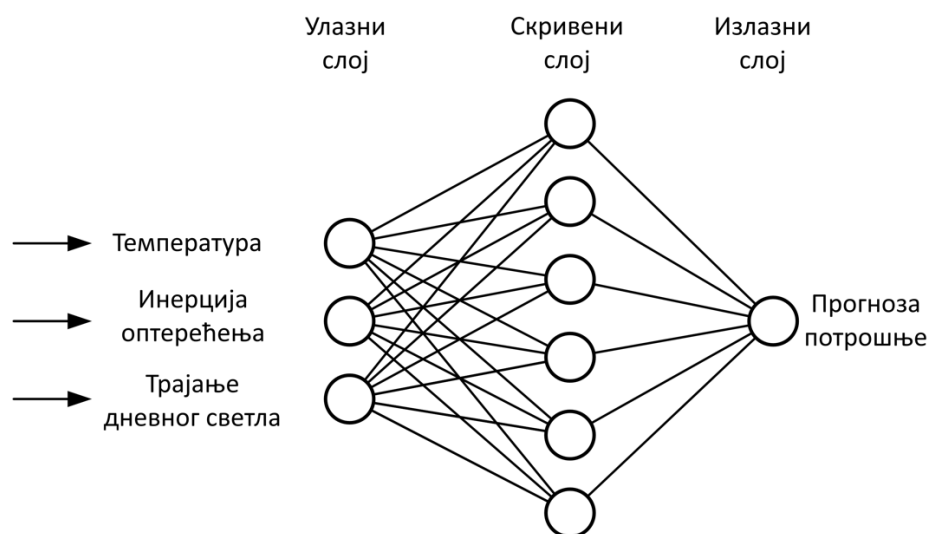
Слика 51 – Криве учења и валидације за изабрану структуру неуронске мреже

Вредности хипер параметара који се оптимизују приказане су у табели 19. За сваки параметар наведен је опсег вредности, корак (уколико постоји) и оптимизована вредност.

Табела 19 – Оптимизоване вредности хипер параметара неуронске мреже за прогнозу потрошње електричне енергије у адаптивном моделу прогнозе

Назив хипер параметра	Опсег вредности	Корак	Оптимизована вредност
Број скривених слојева	1-4	1	1
Број неурона у скривеном слоју	1-20	1	6
Број пакета	1-5	1	1
Број епоха	10-500	1	31
Стопа учења	0,01-0,05	0,01	
Оптимизатор	Gradient Descent, RMSprop, Adam, Adadelta, Adagrad, Adamax, NAdam, FTRL	-	Adam
Активациона функција	ReLU, Sigmoid, Softmax, Softplus, Softsign, Hyperbolic tangent, SELU, ELU, Exponential	-	ReLU
Улазне колоне	-	-	Температура Инерција оптерећења Трајање дневног светла

Архитектура *feed-forward* неуронске мреже која се користи као алгоритам прогнозе потрошње електричне енергије у адаптивном моделу прогнозе приказана је на слици 52.



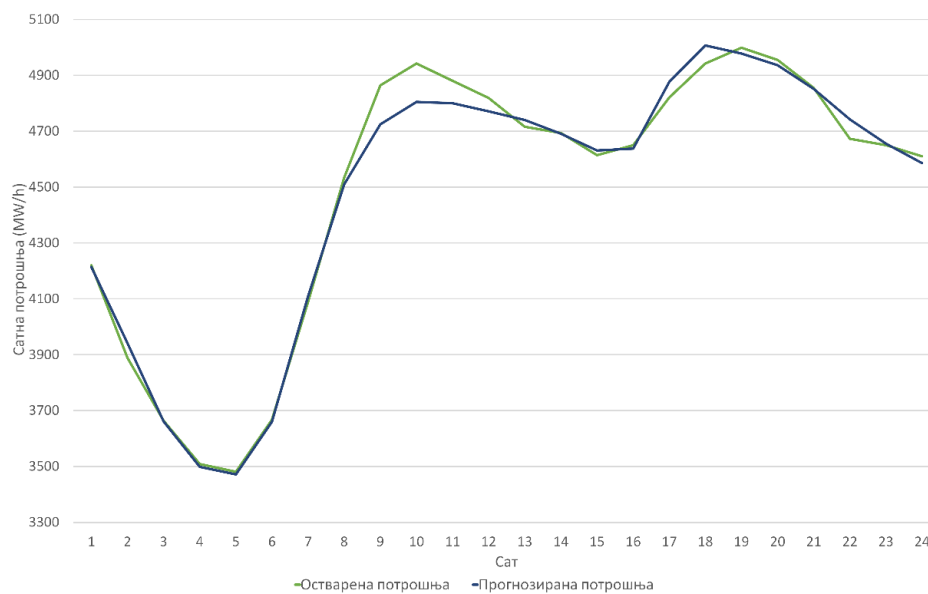
Слика 52 – Архитектура неуронске мреже за прогнозу у адаптивном моделу

### 8.1.1.8. Резултати прогнозе за изабрани дан

Табела 20 приказује податке о прогнозираној и оствареној потрошњи за сваки сат прогнозираног дана и апсолутну и процентуалну разлику у потрошњи по сату. Графички приказ прогнозе и остварења потрошње за прогнозирани дан приказан је на слици 53.

Коначна *MAPE* вредност прогнозираног дана од 0,76% показује малу разлику између прогнозираног и оствареног профила потрошње, што имплицира да је резултат на

примеру прогнозираног дана прихватљив од стране електроенергетског преносног система који користи апликативну примену адаптивног модела прогнозе.



Слика 53 – Криве прогнозираних и остварених сатних вредности потрошње за дан 2021-01-25



Табела 20 – Прогнозирана и остварена потрошња електричне енергије по сатима за прогнозирани дан

Сат	Прогнозирана потрошња ( $C_f$ ) [MW]	Остварена потрошња ( $C_r$ ) [MW]	$ C_f - C_r $ [MW]	Сатна MAPE вредност [%]
1	4214	4220	6	0,14
2	3941	3889	52	1,34
3	3661	3664	3	0,08
4	3499	3509	10	0,28
5	3472	3481	9	0,26
6	3660	3667	7	0,19
7	4109	4089	20	0,49
8	4510	4534	24	0,53
9	4725	4864	139	2,86
10	4805	4942	137	2,77
11	4800	4880	80	1,64
12	4771	4819	48	1,00
13	4740	4716	24	0,51
14	4691	4694	3	0,06
15	4631	4614	17	0,37
16	4638	4650	12	0,26
17	4877	4821	56	1,16
18	5007	4942	65	1,32
19	4978	4999	21	0,42
20	4936	4955	19	0,38
21	4851	4855	4	0,08
22	4742	4673	69	1,48
23	4655	4650	5	0,11
24	4586	4610	24	0,52
<b>Просечна MAPE вредност</b>				<b>0,76</b>

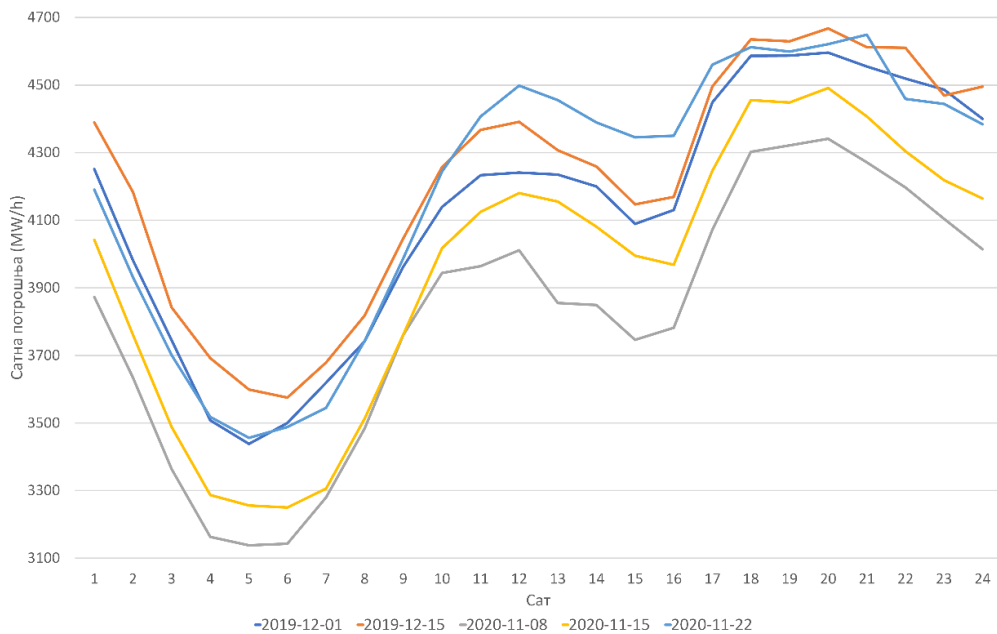
### 8.1.2. Прогноза за специјални дан

Пример прогнозе специјалног дана направљен је за датум 1. јануар 2021. године. Овај датум се празнује нерадно на целој територији за који се прогнозира потрошња. Осим чињенице да је нерадан, овај дан има и бројне друге специфичности које утичу на понашање дневног профила потрошње. Процес прогнозе у експерименту се извршава у поноћ између 31. децембра 2020. године и 1. јануара 2021. године. У моменту прогнозе доступни су сви подаци о претходној потрошњи, подаци о оствареним временским условима и временска прогноза за наредних 24 сата.

Када је покренут поступак прогнозе за датум 2021-01-01, филтер сличних дана је изабрао 5 дана који су одређени као најсличнији прогнозираном дану. Поступак избора сличних дана описан је у претходном поглављу. Изабрани дани приказани су у табели 21, а профили потрошње за изабране дане приказани су на слици 54.

Табела 21 – Слични дани изабрани филтером сличних дана у прогнози за специјални дан 2021-01-01

2019-12-01
2019-12-15
2020-11-08
2020-11-15
2020-11-22



Слика 54 – Криве сатних вредности потрошње електричне енергије за изабране сличне дане у односу на специјални дан

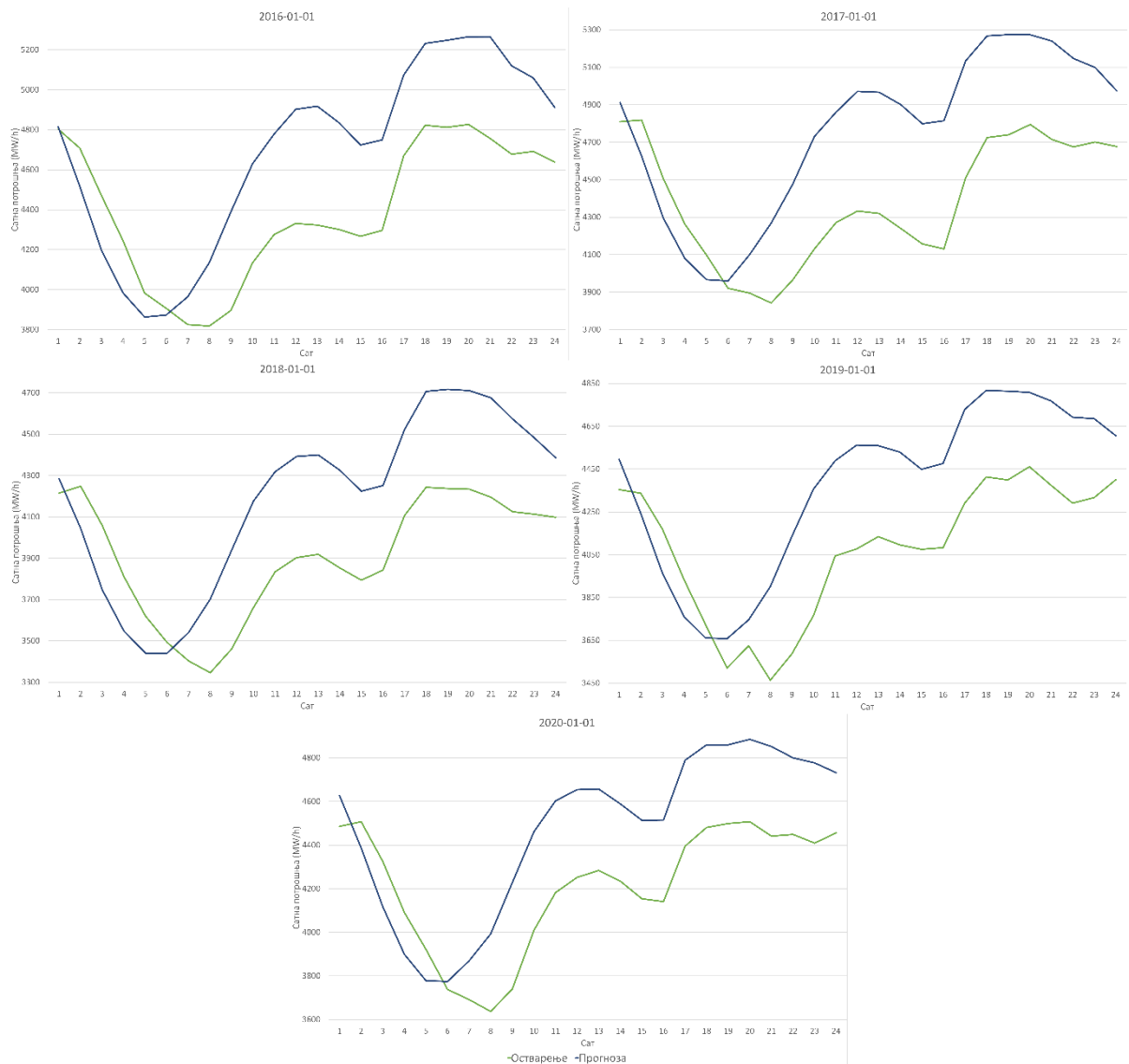
Пошто је прогнозирани специјални дан нерадан, сви селектовани слични дани имају дан у седмици недеља.

Када се селектовани слични дани искористе у алгоритму неуронске мреже ради прогнозе потрошње, без даље употребе филтера специјалних дана, добију се резултати као што се види у табели 22.

Табела 22 – Међупрогноза потрошње електричне енергије за специјални дан

Сат	Прогнозирана потрошња ( $C_f$ ) [MW]	Остварена потрошња ( $C_r$ ) [MW]	$ C_f - C_r $ [MW]	Сатна <i>MAPE</i> вредност [%]
1	4149	3994	155	3,88
2	3898	4069	171	4,20
3	3628	3873	245	6,33
4	3434	3551	117	3,29
5	3377	3368	9	0,27
6	3391	3258	133	4,08
7	3486	3226	260	8,06
8	3660	3248	412	12,68
9	3904	3318	586	17,66
10	4120	3640	480	13,19
11	4219	3762	457	12,15
12	4264	3862	402	10,41
13	4201	3908	293	7,50
14	4156	3782	374	9,89
15	4064	3671	393	10,71
16	4080	3739	341	9,12
17	4365	4043	322	7,96
18	4518	4119	399	9,69
19	4517	4066	451	11,09
20	4543	4076	467	11,46
21	4499	4047	452	11,17
22	4418	3966	452	11,40
23	4344	4038	306	7,58
24	4291	4068	223	5,48
<b>Просечна <i>MAPE</i> вредност</b>				<b>8,72</b>

Очигледно је да је просечна *MAPE* вредност од 8,72% веома висока и неприхватљива од стране електроенергетског преносног система који користи апликативни систем. Разлог за овако велику грешку је специфично понашање профила потрошње у наведеном специјалном дану. Задатак филтера специјалних дана је да креира прогнозу за овај дан која би била прихватљива од стране преносног система. Према поступку прогнозе за специјалне дане (слика 12), сатне *MAPE* вредности из табеле 22 представљају међупрогнозу потрошње. Корак који следи је идентификовање специјалних дана одговарајућег типа из прошлости. То су дани који су унапред означени да припадају истом типу специјалног дана као и прогнозирани дан. У овом случају то ће бити 1. јануар за претходних 5 година. За ових 5 дана ће се извршити симулација прогнозе.



Слика 55 – Резултати симулираних прогноза за историјске специјалне дане

Слика 55 приказује резултате симулираних прогноза за сваки од специјалних дана из историјске базе података. Може се уочити да постоји шаблон грешке за сваки од дана. Подаци о прогнозираним вредностима и остварена потрошња користе се као улаз за модел линеарне регресије. Овај модел се креира за сваки сат појединачно. У овом примеру једина независна променљива је међупрогноза потрошње за специјални дан ( $C_{tspd}$ ). То значи да број преосталих фактора ( $P$ ) има вредност 0. Разлог овоме је искуствени закључак да постоји довољна линеарност између међурезултата и остварења.

Резултати линеарне регресије приказани су у табели 23. Излаз из регресионе анализе за сваки сат су нагиб ( $\beta_{tspd}$ ) и одсечак ( $\beta_0$ ). Ове вредности се, према формули (19), примењују на међурезултат прогнозираног специјалног дана. Резултат је прогнозирана вредност потрошње ( $C_{spd}$ ), као коначна прогноза потрошње за сваки сат прогнозираног специјалног дана.

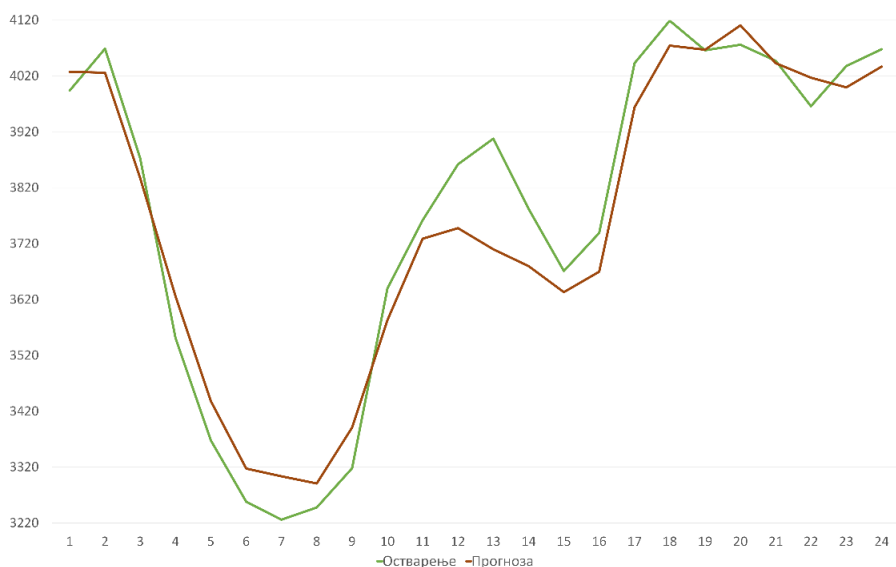
Табела 23 – Резултати линеарне регресије за специјални дан

Сат	Нагиб ( $\beta_{tspd}$ )	Одсечак ( $\beta_0$ )	Међурезултат ( $C_{tspd}$ ) за 2021-01-01 [MW/h]	Прогнозирана вредност потрошње ( $c_{spd}$ ) [MW/h]
1	0,98	-35	4149	4027,60
2	1,04	-10	3898	4025,94
3	1,06	5	3628	3837,45
4	1,05	10	3434	3625,70
5	1,03	-50	3377	3438,22
6	0,99	-50	3391	3317,53
7	0,96	-35	3486	3303,80
8	0,90	15	3660	3290,99
9	0,86	16	3904	3390,98
10	0,86	20	4120	3583,33
11	0,88	10	4219	3728,86
12	0,88	11	4264	3747,88
13	0,88	13	4201	3709,87
14	0,88	16	4156	3679,64
15	0,89	19	4064	3633,26
16	0,89	50	4080	3669,76
17	0,90	51	4365	3964,09
18	0,90	-5	4518	4074,57
19	0,90	-6	4517	4067,12
20	0,91	-10	4543	4110,55
21	0,90	8	4499	4043,03
22	0,90	20	4418	4016,84
23	0,92	22	4344	3999,76
24	0,94	5	4291	4036,97

У табели 24 упоредно су приказане прогнозиране и вредности потрошње за специјални дан 2021-01-01. Графички приказ односа прогнозираних и остварених вредности приказан је на слици 56. Просечна сатна *MAPE* вредност је 1,58%, што је прихватљиво за електроенергетски преносни систем који користи софтверску имплементацију адаптивног модела прогнозе.

Табела 24 – Прогнозирана и остварена потрошња по сатима за прогнозирани специјални дан

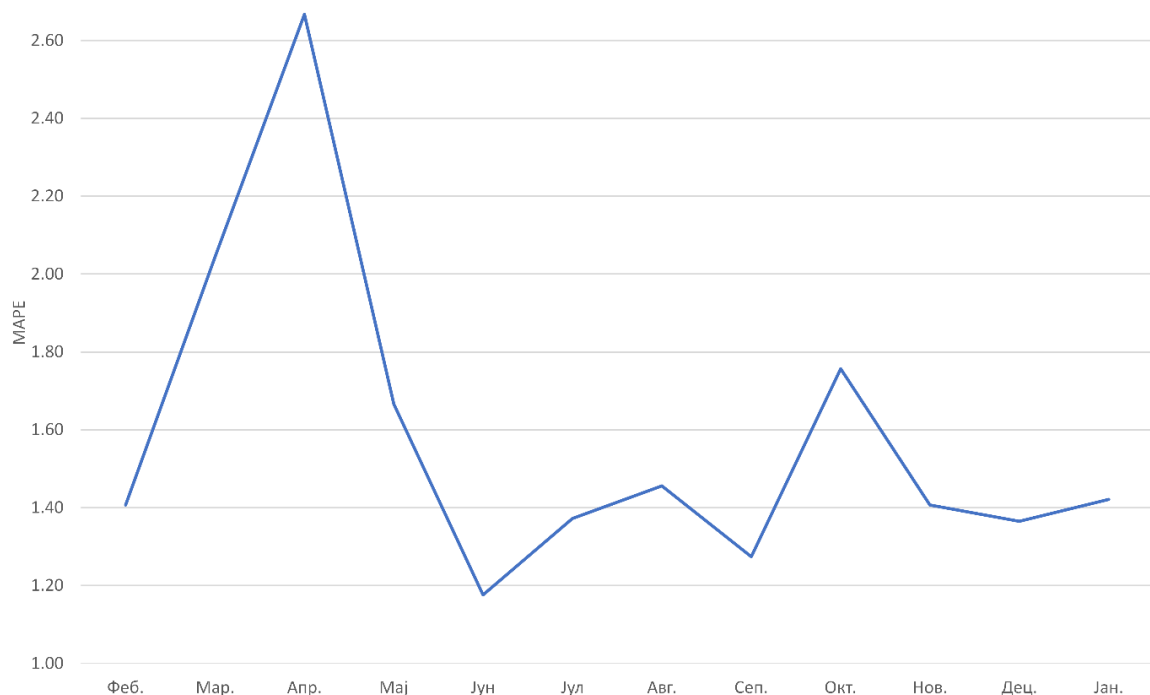
Сат	Прогнозирана вредност потрошње ( $c_{spd}$ ) [MW/h]	Остварена потрошња ( $C_r$ ) [MW]	$ c_{spd} - C_r $ [MW]	Сатна МАРЕ вредност [%]
1	4027,60	3994	33,60	0,84
2	4025,94	4069	43,06	1,06
3	3837,45	3873	35,55	0,92
4	3625,70	3551	74,70	2,10
5	3438,22	3368	70,22	2,08
6	3317,53	3258	59,53	1,83
7	3303,80	3226	77,80	2,41
8	3290,99	3248	42,99	1,32
9	3390,98	3318	72,98	2,20
10	3583,33	3640	56,67	1,56
11	3728,86	3762	33,14	0,88
12	3747,88	3862	114,12	2,95
13	3709,87	3908	198,13	5,07
14	3679,64	3782	102,36	2,71
15	3633,26	3671	37,74	1,03
16	3669,76	3739	69,24	1,85
17	3964,09	4043	78,91	1,95
18	4074,57	4119	44,43	1,08
19	4067,12	4066	1,12	0,03
20	4110,55	4076	34,55	0,85
21	4043,03	4047	3,97	0,10
22	4016,84	3966	50,84	1,28
23	3999,76	4038	38,24	0,95
24	4036,97	4068	31,03	0,76
<b>Просечна МАРЕ вредност</b>				<b>1,58</b>



Слика 56 – Криве прогнозираних и остварених сатних вредности потрошње за специјални дан 2021-01-01

### 8.1.3. Анализа резултата прогнозе

Слика 57 приказује просечне сатне резултате прогнозе потрошње на месечном нивоу, за период од фебруара 2020. године до јануара 2021. године. Очигледно је да је *MAPE* вредност почела да расте у марту и највишу вредност је достигла у априлу. Осим ова два месеца, највишу вредност је достигла у октобру. У осталим месецима *MAPE* вредност је испод 1,5%. Један од разлога за раст у марту и априлу вероватно је то што је у питању период када је почела епидемија COVID-19. Навике предузећа, запослених и осталих потрошача тада су се нагло промениле, и самим тим су профили потрошње почели да се мењају. Када се овакво понашање усталило, адаптивни модел прогнозе се прилагодио новонасталој ситуацији, кроз прилагођавање параметара генетским алгоритмом. Други разлог раста грешке у априлу је тај што је у априлу промена грејне сезоне. Пошто је један од критеријума селекције дана у иницијалном избору управо припадање истој грејној сезони, филтер сличних дана губи могућност избора дана који су временски блиски прогнозираном дану. Из истог разлога је повећана грешка у октобру. Међутим, како је у октобру то једини проблем, грешка је далеко мања него у априлу.



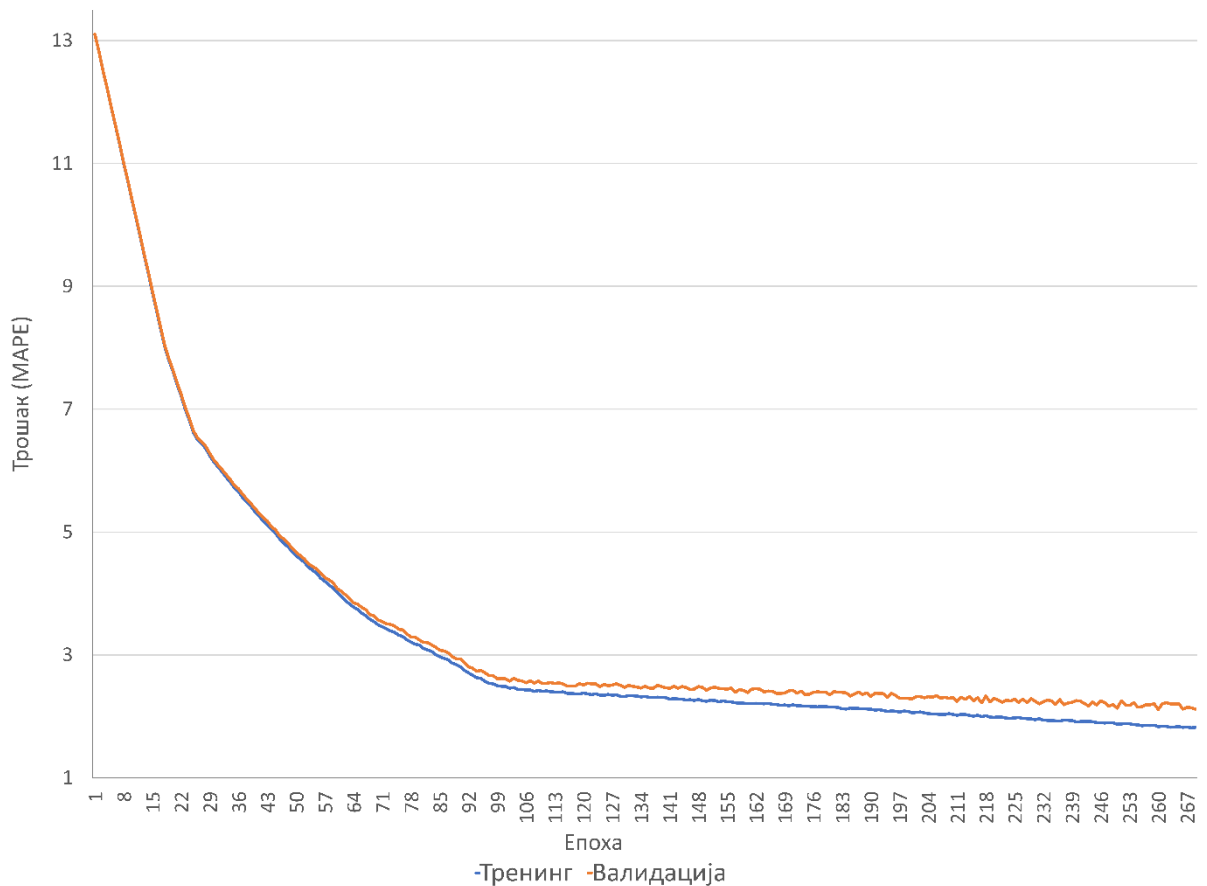
Слика 57 – Просечна сатна *MAPE* вредност по месецима

### 8.1.4. Поређење са резултатима прогнозе коришћењем саме неуронске мреже

Ради утврђивања доприноса адаптивног модела прогнозе, креиран је прорачун за исти период коришћењем *feed-forward* неуронске мреже, без адаптивног модела прогнозе.

Узорак за тренинг неуронске мреже је целокупна историја података. За разлику од адаптивног модела прогнозе, тренинг и прогноза се извршавају за сва 24 сата у једном пролазу.

У овом примеру, криве учења за оптимизацију хипер параметара неуронске мреже праћене су на сету од шест година из историјске базе података, док су валидационе криве праћене на узорку последње године. Крива учења и крива валидације изабране структуре неуронске мреже приказане су на слици 58.



Слика 58 – Криве учења и валидације на примеру саме неуронске мреже

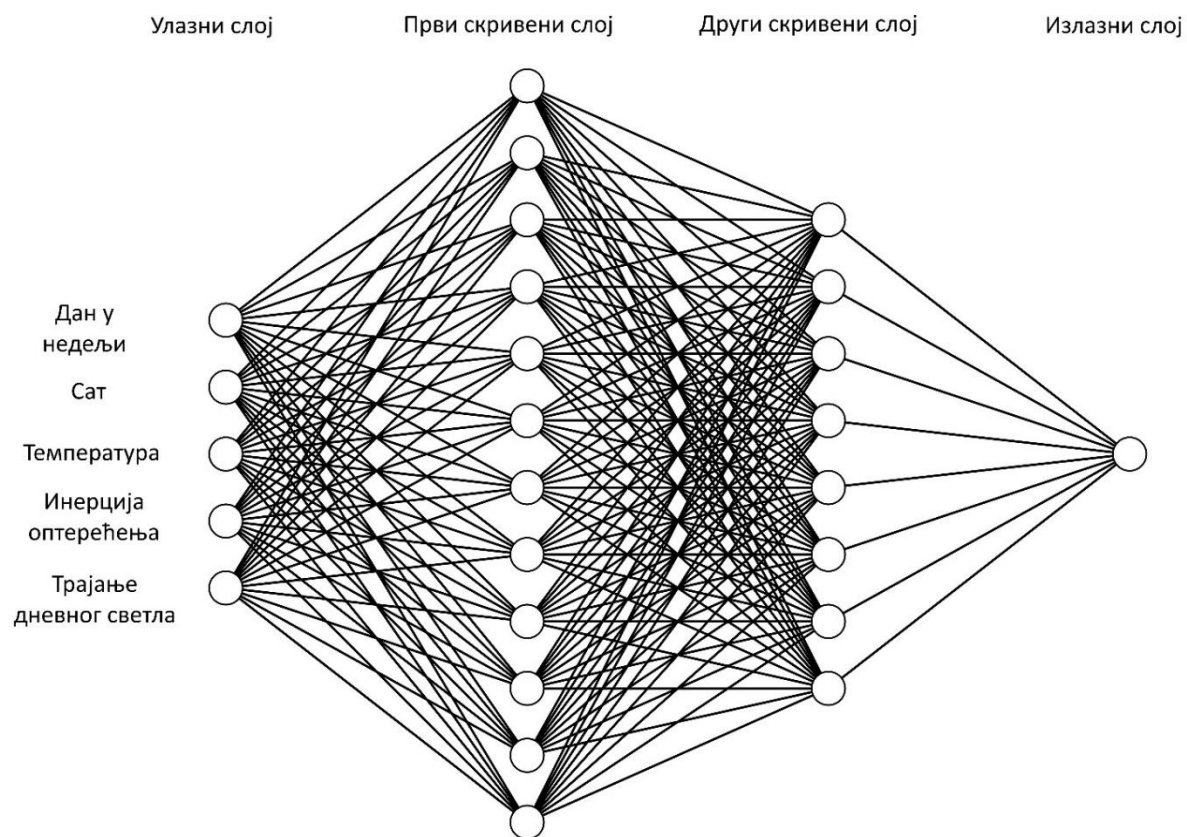
Вредности оптимизованих хипер параметара приказани су у табели 25.



Табела 25 – Оптимизоване вредности хипер параметара на примеру саме неуронске мреже

Назив хипер параметра	Опсег вредности	Корак	Оптимизована вредност
Број скривених слојева	1-4	1	2
Број неурона у првом скривеном слоју	1-20	1	12
Број неурона у другом скривеном слоју	1-20	1	8
Број пакета	1-5	1	5
Број епоха	10-500	1	269
Стопа учења	0,01-0,05	0,01	
Оптимизатор	Gradient Descent, RMSprop, Adam, Adadelta, Adagrad, Adamax, NAdam, FTRL		Adam
Активациона функција	ReLU, Sigmoid, Softmax, Softplus, Softsign, Hyperbolic tangent, SELU, ELU, Exponential		ReLU
Улазне колоне	-	-	Дан у недељи Сат Температура Инерција оптерећења Трајање дневног светла

Архитектура *feed-forward* неуронске мреже на примеру прогнозе потрошње електричне енергије коришћењем саме неуронске мреже приказана је на слици 59.



Слика 59 – Архитектура на примеру саме неуронске мреже

У табели 26 компаративно су приказане просечне сатне *MAPE* вредности добијене коришћењем саме неуронске мреже и адаптивног модела прогнозе.

Табела 26 – Резултати прогнозе неуронском мрежом и адаптивним моделом прогнозе

Месец	<i>MAPE</i> (неуронска мрежа)	<i>MAPE</i> (адаптивни модел прогнозе)
Фебруар	1,90	1,41
Март	2,50	2,04
Април	3,57	2,67
Мај	2,12	1,67
Јун	1,59	1,18
Јул	1,91	1,37
Август	1,93	1,46
Септембар	1,78	1,27
Октобар	2,33	1,76
Новембар	1,82	1,41
Децембар	1,72	1,37
Јануар	1,90	1,42
<b>Просечна <i>MAPE</i> вредност</b>	<b>2,09</b>	<b>1,58</b>

Из табеле се види да су резултати добијени коришћењем адаптивног модела прогнозе далеко прецизнији. Просечна годишња вредност *MAPE* је 2,09 за неуронску мрежу, док вредност за адаптивни модел прогнозе износи 1,58.

Табела 27 – Просечно време извршавања прогнозе неуронском мрежом и адаптивним моделом прогнозе

Тип прогнозе	Време				
	дневне метахеуристичке оптимизације	селекције сличних дана	извршавања прогнозе (тренинг + прогноза)	укупно (са метахеуристичком оптимизацијом)	укупно (без метахеуристичке оптимизације)
<b>Неуронска мрежа</b>	-	-	~240 минута	~240 минута	~240 минута
<b>Адаптивни модел прогнозе</b>	~128 минута	~5 секунди	~30 минута	~158 минута	~30 минута

Табела 27 приказује упоредно просечно време извршавања процеса прогнозе, за саму неуронску мрежу и за адаптивни модел прогнозе. У оба случаја прогноза је вршена на окружењу са 8-језгарним процесором и радном меморијом од 64GB. У односу на саму неуронску мрежу, у прогнозу адаптивним моделом прогнозе додатно су укључени кораци дневне метахеуристичке оптимизације и селекције сличних дана. Дневна метахеуристичка оптимизација се извршава независно од корисничког поступка прогнозе, тако да је укупно време извршавања приказано са и без обухваћеног овог времена.

Може се видети да је просечно извршавање прогнозе неуронском мрежом за 24 сата прогнозе 240 минута (око 10 минута по једном сату). Ово време обухвата тренинг неуронске мреже и прорачун на прогнозираном дану. Време метахеуристичке

оптимизације је око 128 минута, док извршавање прогнозе траје око 39 минута (око 1,25 минута по сату). Разлог зашто прогноза неуронском мрежом траје оволико краће је у томе што је у случају адаптивног модела прогнозе изабран скуп од 5 сличних дана, у односу на прогнозу неуронском мрежом где је узорак целокупна историја из базе података. Селекција сличних дана је занемарљива у односу на остала времена. Укупно време извршавања прогнозе адаптивним моделом прогнозе износи око 158 минута, што је за 82 минута мање у односу на прогнозу само неуронском мрежом. Ако се у укупно време не рачуна метахеуристичка оптимизација, време извршавања адаптивним моделом прогнозе је око 30 минута. Дакле, перформансе адаптивног модела прогнозе надмашују перформансе прогнозе самом неуронском мрежом.

### **8.1.5. Поређење са другим техникама**

Ради провере ефикасности предложеног решења, извршено је поређење са сродним техникама које се баве прогнозом потрошње електричне енергије. Прорачун прогнозе је извршен коришћењем једне упоредне технике над истим сетом података и резултати два решења су упоређени. Поред тога, резултати предложеног решења упоређени су и са резултатима других техника над независним сетовима података.

#### **8.1.5.1. Упоредне прогнозе над истим сетом података**

Ради упоредног приказа резултата адаптивног модела прогнозе у односу на друге технике, потрошња електричне енергије прогнозирана је за исти период коришћењем Хибридног *PSO-SVM* метода за краткорочну прогнозу потрошње електричне енергије [15]. Овај хибридни модел базиран је на оптимизацији ројева честица (*PSO*) и на методи потпорних вектора (*SVM*). Овим решењем се селекују слични дани из историјских података, користећи се временском близином дана, метеоролошким подацима (пре свега температуром) и значајним метеоролошким променама. *SVM* израчунава корелацију између метеоролошких података и потрошње електричне енергије, док *PSO* модул оптимизује *SVM* параметре.

Резултати на месечном нивоу добијени коришћењем *PSO-SVM* метода приказани су и упоређени адаптивним моделом прогнозе у табели 28.

Табела 28 – Упоредни резултати прогнозе добијени Хибридном PSO-SVM методом и адаптивним моделом прогнозе

Месец	MAPE (PSO-SVM)	MAPE (адаптивни модел прогнозе)
Фебруар	2,07	1,41
Март	2,19	2,04
Април	2,96	2,67
Мај	2,29	1,67
Јун	1,39	1,18
Јул	1,57	1,37
Август	1,59	1,46
Септембар	1,91	1,27
Октобар	2,12	1,76
Новембар	1,99	1,41
Децембар	1,57	1,37
Јануар	1,95	1,42
<b>Просечна MAPE вредност</b>	<b>1,97</b>	<b>1,58</b>

Евидентно је да адаптивни модел прогнозе производи мању грешку у односу на PSO-SVM метод. При анализи ових резултата треба имати у виду и да је PSO-SVM показује супериорност првенствено на подручјима са значајним температурном варијацијама.

### 8.1.5.2. Поређење са резултатима других техника краткорочне прогнозе потрошње електричне енергије

У доступној литератури, резултати различитих техника најчешће су представљени над специфичним сетовима података. Иако поређење резултата адаптивног модела прогнозе са таквим резултатима не може да буде у потпуности репрезентативно, може да се стекне увид у позиционирање предложеног решења у односу на друга доступна решења. Резултати свих наведених техника односе се на прогнозу једног дана унапред, и за све резултате посматра се средња апсолутна процентуална грешка.

Прогноза таласастом неуронском мрежом базираном на сличним данима [73] тестирана је за регион Нове Енглеске у Сједињеним Америчким Државама, где је прогнозирано оптерећење упоређено са остварењима у 2006. години. Процентуална грешка по месецима се креће од 1,24% у октобру, до 2,22% у јуну. Ови резултати упоређени су са резултатима добијеним адаптивним моделом прогнозе (табела 29).

Табела 29 – поређење са резултатима прогнозе коришћењем таласасте неуронске мреже базиране на сличним данима

Месец	MAPE (таласаста неуронска мрежа базирана на сличним данима)	MAPE (адаптивни модел прогнозе)
Јануар	1,59	1,42
Фебруар	1,30	1,41
Март	1,37	2,04
Април	1,27	2,67
Мај	1,33	1,67
Јун	2,22	1,18
Јул	1,57	1,37
Август	2,07	1,46
Септембар	1,98	1,27
Октобар	2,12	1,76
Новембар	1,24	1,41
Децембар	1,68	1,37
<b>Просечна MAPE вредност</b>	<b>1,65</b>	<b>1,58</b>

Просечна апсолутна процентуална грешка добијена прогнозом таласастом неуронском мрежом базираном на сличним данима је 1,65%, што је грешка већа за 4% релативно (0,07% апсолутно) односу на просечне годишње резултате приказане у дисертацији. Међутим 2020. година, у којој су презентовани резултати адаптивног модела прогнозе садржи период почетка пандемије COVID-19. Једина два месеца у којима MAPE премашује 2% у адаптивном моделу прогнозе су управо два месеца почетка пандемије: март и април 2020. године. Када се из разматрања искључе ова два месеца, адаптивни модел прогнозе даје резултате прогнозе 15% боље од резултата таласастом неуронском мрежом (0,21% у апсолутном износу). Поред овога, број итерација тренинга у таласастој неуронској мрежи је 700, док је број итерација у адаптивном моделу прогнозе 31. Рад [73] не садржи податке о времену извршавања тренинга и прогнозе, али се на основу броја итерација може закључити да је у овом погледу адаптивни модел прогнозе супериоран.

Техника са називом „Сличност секвенци шаблона – Неуронске мреже“ [82] је верификована на примеру краткорочне прогнозе оптерећења за аустралијску државу Нови Јужни Велс, за три године: 2009, 2010 и 2011. Просечна апсолутна грешка добијена коришћењем ове технике над задатим сетом података је 3,16%, при чему је највећа грешка у летњим месецима (децембар – фебруар), док је грешка у зимским месецима несразмерно мања. У раду се наводи да је могуће објашњење за већу грешку у летњим месецима употреба расхладних уређаја у данима који су на тој територији веома топли, као и специјални дани (државни празници и сл.) и летњи распуст. Овај рад представља пример другачијих календарских, метеоролошких и друштвених околности у односу на студију случаја из дисертације.

У студији случаја Хибридног ВА-ELM модела [92] анализирани су прогноза и остварења за град Јангџуен у Кини, за три дана у јуну месецу 2013. године. МАРЕ за ова три дана се креће између 0,5% и 0,7%. Приказани резултат превазилази већину других резултата који се могу пронаћи у доступној литератури. Међутим, тешко је извршити адекватно поређење са адаптивним моделом прогнозе, с обзиром да су наведени резултати за само три дана. Поред тога, ради се о прогнози за један један град који је релативно мали по површини, броју становника и оптерећењу електричне енергије у односу на подручје анализирано у студији случаја прогнозе потрошње електричне енергије у дисертацији. У односу на пример из овог рада, адаптивни модел прогнозе има далеко шири и конкретнију верификацију.

Модел предложен у раду [125] користи LSTM (*Long Short-Term Memory*) архитектуру рекурентне неуронске мреже за краткорочну прогнозу потрошње. Модел је верификован на Белгијском преносном систему *Elia*, за територију целе Белгије. Потрошња је прогнозирана на сатном нивоу, за период од једне године. Ова студија случаја је упоредива са студијом случаја у дисертацији, с обзиром да се ради о подручјима која су упоредива по површини, броју становника и потрошњи. Просечна апсолутна процентуална грешка у овом експерименту за стандардне дане је 2,13%.

Наведени експерименти краткорочне прогнозе електричне енергије се одвијају у различитим околностима у односу на студију случаја у дисертацији. И поред тога, очигледно је да адаптивни модел прогнозе даје резултате прогнозе који се могу сматрати супериорним у односу на наведене примере. У односу на постојеће технике, адаптивни модел прогнозе унапређује резултате пре свега тако што омогућава флексибилну оптимизацију сличних дана коришћењем генетског алгорита. Иако су перформансе прогнозе у мањој мери су приказане у доступној литератури, адаптивни модел прогнозе се показује као супериоран и у овом погледу.

## **8.2. Прогноза потрошње природног гаса у дистрибутивној мрежи**

Пример прогнозе потрошње природног гаса у дистрибутивној мрежи јесте симулација и биће приказан за територију једног региона у Холандији. Поступак прогнозе кроз адаптивни модел прогнозе временских серија биће приказан за један типичан дан у току зимског периода на почетку 2018. године.

У симулацији прогнозе поступак започиње у поноћ на почетку дана за који се прогнозира потрошња. У том моменту доступни су подаци о претходној потрошњи, подаци о оствареним временским условима и временска прогноза за наредна 24 сата. Прогноза се извршава за један дан, односно за временску серију од 24 сата.

Поред прогнозе од 24 сата, биће приказани и резултати прогнозе за 32 дана у току зимског периода на почетку 2018. године.

Јединица мере за потрошњу је један кубни метар ( $m^3$ ).

Тачна локација региона и време за који се прогнозира потрошња нису приказани из разлога поверљивости података.

### **8.2.1. Расположиви подаци и коришћени параметри у поступку прогнозе**

Историјски подаци о потрошњи у експерименту обухватају период од 1. јануара 2014. године до 15. фебруара 2018. године. Независни подаци који су расположиви јесу временска остварења за подручје једног региона и календарски подаци. Од временских остварења расположиви су температура, ветар и проценат облачности.

### **8.2.2. Оптимизација генетским алгоритмом**

Параметри који су узети у обзир приликом оптимизације генетским алгоритмом приказани су у табели 30. Оптимизација се извршава за температуру и за ветар. Оптимизација за облачност се не извршава, пошто је итеративним тестирањем утврђено да облачност нема значајан утицај на сатну потрошњу природног гаса.

Распон могућих вредности за тежинске коефицијенте исти је као у случају прогнозе потрошње електричне енергије.

У примеру ће бити искоришћена оптимизација само на кварталном нивоу, која се извршила један дан пре почетка прогнозе која се извршава у експерименту. Прорачуни прогнозе за све дане у примеру користиће вредности параметара добијене кварталном оптимизацијом. Последња колона у табели 30 приказује вредности параметара оптимизованих кварталном оптимизацијом генетским алгоритмом.



Табела 30 – Параметри који се оптимизују генетским алгоритмом у прогнози потрошње природног гаса

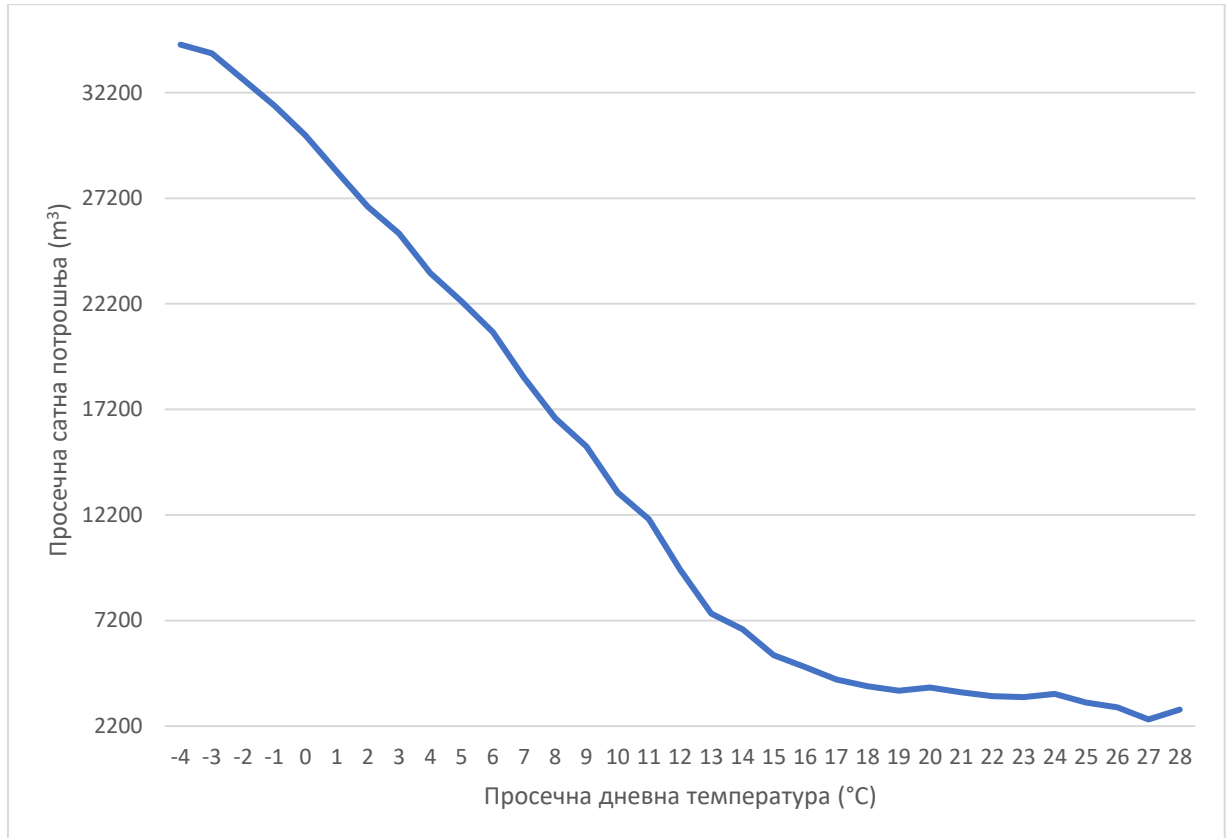
Параметри временских прилика				
Назив параметра	Минимална вредност	Максимална вредност	Корак промене	Вредност кварталног параметра
Тежински коефицијент за температуру ( $w_{Ct}$ )	0	100	1	65
Тежински коефицијент за ветар ( $w_{Cwi}$ )	0	100	1	34
Инерциони параметри				
Назив параметра	Минимална вредност	Максимална вредност	Корак промене	
Тежински коефицијент инерције потрошње ( $w_{Ccin}$ )	0	100	1	70
Тежински коефицијент термалне инерција ( $w_{Ctin}$ )	0	100	1	55
Тежински коефицијент инерције последњег сата ( $sc_{lhin}$ )	0	100	1	57
Дневни умањујући коефицијент ( $ddwc$ )	0,00	1,00	0,01	0,13
Гранична вредност ( $L_{\sigma}$ )	1	3	0,1	2,5
Број предисторијских дана ( $M$ )	1	5	1	4
Остали параметри				
Назив параметра	Минимална вредност	Максимална вредност	Корак промене	
Тежински коефицијент за близину дана ( $w_{Cdp}$ )	0	100	1	51
Тежински коефицијент за временске прилике ( $w_{Cw}$ )	0	100	1	70
Тежински коефицијент за инерцију ( $w_{Cin}$ )	0	100	1	41

### 8.2.3. Селекција фактора од утицаја и партиципационих коефицијената

Подаци расположиви у тренутку прорачуна прогнозе јесу метеоролошки подаци и подаци о потрошњи за последње четири године. Метеоролошки подаци који су на располагању су сатне вредности температуре, облачности и брзине ветра. Међутим, тежински коефицијент за облачност добијен кварталном оптимизацијом је изразито низак у односу на остале коефицијенте, те ће се стога као метеоролошки фактори користити само температура и брзина ветра. Такође је закључено да у овом случају дужина дневног светла не утиче у значајној мери на потрошњу природног гаса.

Кретање просечне потрошње природног гаса на основу промене температуре приказано је на слици 60. Очигледно је да повећање температуре доводи до смањења потрошње

природног гаса. Овај тренд је приближно линеаран за распон просечне температуре од  $-4^{\circ}\text{C}$  до  $15^{\circ}\text{C}$ . За температуре преко  $15^{\circ}\text{C}$  овај тренд се смањује, да би при већим температурама тај тренд нестао. На основу расположивих података не може се закључити да у овом конкретном случају високе просечне температуре утичу на повећање потрошње природног гаса.

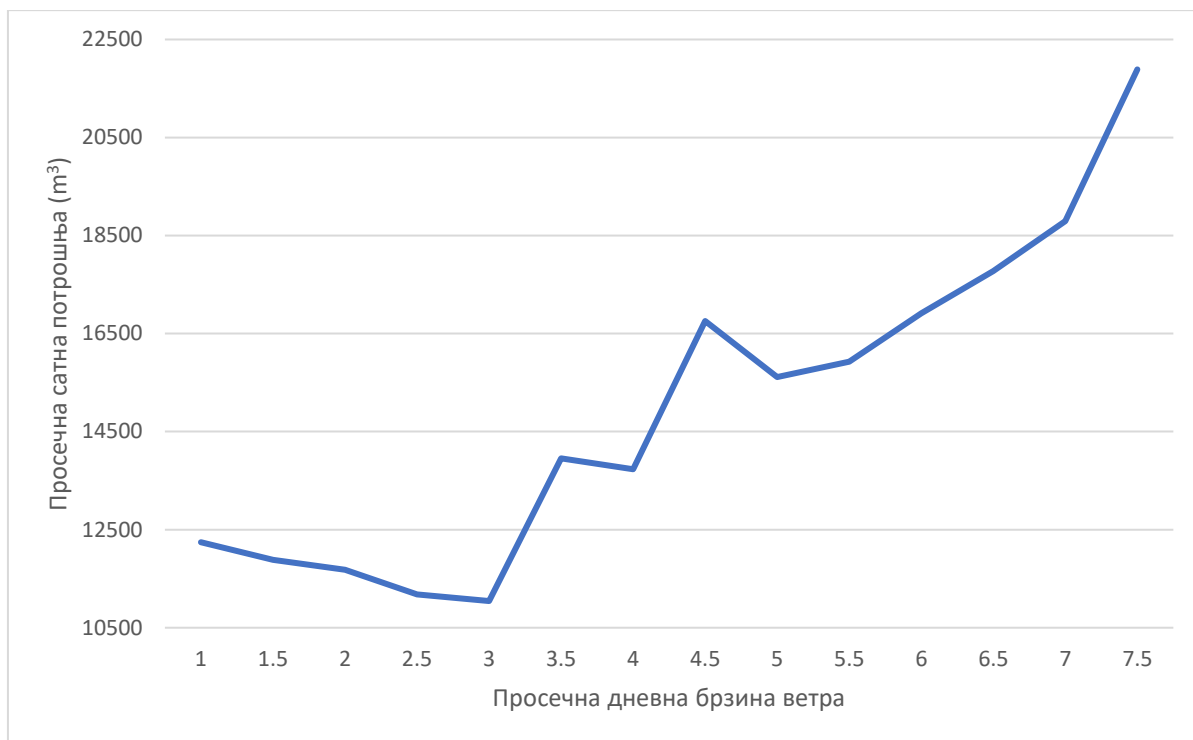


Слика 60 – Историјски тренд промене потрошње природног гаса на основу промене температуре

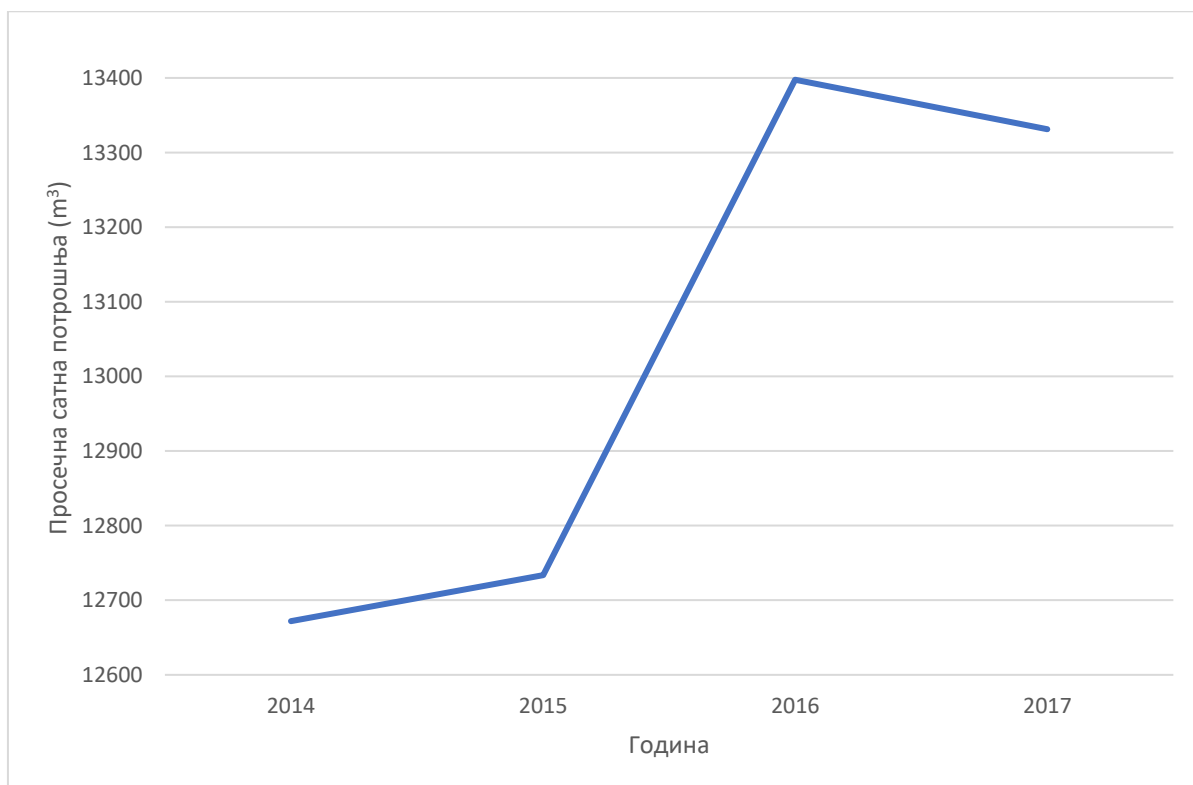
Слика 61 приказује кретање просечне потрошње природног гаса на основу промене просечне дневне брзине ветра. На основу расположивих података може се закључити да је повећање просечне дневне брзине ветра у корелацији са просечном сатном потрошњом природног гаса.

Тренд промене просечне сатне потрошње у различитим годинама приказан је на слици 62. Генерални тренд је повећање потрошње природног гаса протоком година. Просечно сатно повећање потрошње износи  $220 \text{ m}^3$  годишње. Ова вредност се поклапа и са трендом утицаја протока времена на просечну криву потрошње. Просечни тренд промене по протоку једног дана је  $0,60 \text{ m}^3$ .

На основу описаних поступака у поглављу 5.1.2.1.1. , вредност  $A1$  је  $1000 \text{ m}^3$ , а партиципациони коефицијенти за факторе који утичу на сличност приказани су у табели 31.



Слика 61 – Историјски тренд промене потрошње природног гаса на основу промене брзине ветра



Слика 62 – Историјски тренд промене потрошње природног гаса по годинама

Табела 31 – Фактори од утицаја и њихови партиципациони коефицијенти

Фактор	Јединица мере фактора	Вредност партиципационог коефицијента
Температура	°C	1,00
Ветар	m/s	0,74
Близина дана	дан	0,6
Инерција потрошње ресурса	m <sup>3</sup>	0.001

#### 8.2.4. Релевантни параметри прогнозираног дана

Дан у недељи за који се прогнозира потрошња је понедељак. Грејне сезоне у овом примеру нису дефинисане.

Табела 32 приказује дневне податке временске прогнозе релевантне за прогнозу потрошње природног гаса.

Табела 32 – Дневни подаци временске прогнозе за прогнозирани дан

Просечна температура [°C]	Минимална Температура [°C]	Максимална Температура [°C]	Просечна брзина ветра [m/s]
4,35	0,5	6,6	6,41

Табела 33 приказује сатне податке временске прогнозе релевантне за прогнозу потрошње природног гаса, а то су температура и брзина ветра.

Табела 33 – Сатни подаци временске прогнозе за прогнозирани дан

Сат	Температура [°C]	Брзина ветра [m/s]
00:00	0.80	4
01:00	0.50	4
02:00	0.60	4
03:00	0.50	4
04:00	1.70	4
05:00	2.50	5
06:00	2.60	5
07:00	3.70	6
08:00	4.20	5
09:00	4.80	6
10:00	5.20	7
11:00	5.50	7
12:00	6.40	8
13:00	6.20	8
14:00	5.50	8
15:00	5.10	8
16:00	5.20	8
17:00	5.60	8
18:00	5.90	8
19:00	6.50	8
20:00	6.40	8
21:00	5.90	8
22:00	6.60	7
23:00	6.40	6

Табела 34 садржи инерционе податке у односу на прогнозирани дан: податке о просечној температури и просечној потрошњи за три дана који претходе прогнозираном дану. Последња колона приказује потрошњу у последњем сату претходног дана. Табела има 4 реда, на основу вредности  $M$  (број предисторијских дана) одређене генетским алгоритмом.

Табела 34 – Инерциони подаци у односу на прогнозирани дан

Датум	Просечна температура	Просечна потрошња	Потрошња у последњем сату
2021-xx-[-1]	1.6	25425	16022
2021-xx-[-2]	4.2	22545	
2021-xx-[-3]	3.4	24375	
2021-xx-[-4]	4.4	22109	

### 8.2.5. Селекција података иницијалним филтером

У овом примеру, дан у недељи као услов селекције користи се тако што су посебни типови дана радни дани и викенд. Итеративним понављањем тестних прорачуна, овакав начин селекције се показао као најефикаснији. Други услов у иницијалном филтеру је да је просечна дневна температура плус/минус 5 степени у односу на прогнозирану просечну температуру дана за који се прогнозира потрошња.

Иницијалним филтером издвојено је 236 дана. Ради прегледности, у табели 35 приказује се првих 50 дана, са параметрима и израчунатим коефицијентима сличности.

Оптимизација генетским алгоритмом поставила је релативно високу граничну вредност:  $L_{\sigma} = 2,5$ . Ово може да значи да избацивање дана са девијантним кривама из узорка доприноси резултатима само у случају изузетно велике девијације. У овом примеру филтер девијације потрошње није пронашао такве дане.

У овом примеру, број дана  $z$  који се издвајају као улаз у неуронску мрежу налази се у опсегу од 10 до 40, уз узимање у обзир и максималног коефицијента сличности. Максимални коефицијент сличности је 3.15. У табели 35 последњи дан који се издваја као улаз у неуронску мрежу је 2014-01-13, са коефицијентом сличности 3,14. Укупно 25 дана је селектовано као улаз у алгоритам прогнозе.

Табела 35 – Првих 50 дана селектованих иницијалним филтером

Датум	Мин. темп.	Макс. темп.	Прос. темп	Прос. вет.	Прос. потр. претх. дана	Прос. темп. претх. дана	Потрошња у посл. сату претх. дана	Коеф. сличности (sc)
2017-12-29	0.10	5.60	2.37	5.75	18562	7.37	14528	1.70
2017-01-13	0.20	3.80	1.75	5.54	27609	3.34	15142	1.79
2017-01-03	-0.20	7.00	3.55	6.13	28580	-0.41	15907	1.96
2017-12-18	0.70	2.80	1.79	1.88	27207	2.90	15240	2.37
2016-12-21	-0.80	6.00	3.34	5.04	21737	4.88	15797	2.44
2017-02-01	1.00	5.90	3.40	3.63	24717	5.60	15139	2.47
2017-01-12	1.90	6.30	4.65	5.71	27390	1.29	13741	2.47
2017-01-09	1.00	5.40	3.34	3.88	30600	-2.83	15906	2.47
2015-01-05	0.10	3.70	1.74	3.54	27006	2.59	15812	2.47
2017-12-13	0.50	5.70	3.30	6.46	28068	1.06	14785	2.53
2016-11-14	-3.80	5.10	1.46	3.75	22879	3.46	15289	2.61
2017-01-30	2.00	6.40	4.20	4.08	34408	-3.13	14821	2.66
2014-12-15	2.70	6.00	4.26	5.00	26507	4.56	14927	2.69
2018-01-08	-2.10	2.70	0.50	5.29	20987	7.51	18052	2.70
2017-02-07	0.00	1.90	0.78	4.25	19394	8.27	14547	2.84
2014-12-11	3.60	5.90	4.56	6.79	23955	3.54	13998	2.95
2017-12-04	4.30	7.40	5.86	3.83	24326	2.05	13209	2.98
2014-12-01	-0.70	1.20	0.38	4.33	22436	5.45	14710	3.01
2017-11-27	3.30	7.10	5.53	6.21	15628	10.85	12422	3.01
2016-01-14	0.40	3.70	2.20	5.00	19402	6.72	13398	3.03
2017-12-11	-3.10	0.90	-0.43	3.83	23995	5.44	16323	3.07
2015-02-23	1.60	7.40	4.18	6.08	27589	2.94	15399	3.13
2017-12-19	1.90	6.50	4.76	2.29	25783	2.59	13670	3.14
2015-01-28	2.40	5.70	4.15	7.33	30957	-0.57	14186	3.14
2014-01-13	0.70	8.50	5.46	3.63	20678	9.06	16939	3.14
2016-12-30	-2.90	3.90	-0.41	2.75	17987	7.92	16806	3.16
2018-01-09	-0.60	5.70	2.32	3.25	21288	6.59	16126	3.19
2017-01-31	1.20	3.20	2.35	1.71	31953	1.16	14472	3.20
2014-12-09	0.50	4.50	2.54	3.96	27401	2.85	14056	3.20
2017-12-28	1.20	4.30	2.82	4.92	17653	8.56	12969	3.24
2015-01-07	2.10	6.10	3.77	4.54	26279	3.12	14126	3.24
2017-12-08	0.60	5.10	2.02	5.38	22798	5.86	12120	3.25
2014-01-27	2.20	5.30	3.83	4.96	28270	2.19	14715	3.28
2017-11-30	0.50	3.90	2.05	1.92	23132	3.18	12487	3.29
2014-12-08	-0.70	4.70	2.05	4.04	32375	-0.37	14245	3.29
2017-01-10	3.30	4.80	4.14	3.54	36710	-4.60	13934	3.35
2015-02-16	-1.80	7.00	1.71	2.25	28293	1.08	15947	3.36
2015-01-26	2.90	5.90	4.07	4.46	34984	-0.92	15074	3.41
2015-01-29	0.20	2.80	1.44	5.54	28328	1.74	15314	3.43
2016-01-05	-1.00	4.20	1.58	3.04	19851	5.18	14886	3.45
2015-11-27	0.30	6.80	4.07	5.08	23577	2.21	12433	3.47
2017-12-15	1.00	3.90	2.59	2.25	30940	-0.43	13575	3.49
2015-03-04	0.60	6.20	3.62	5.71	21839	4.16	14439	3.51
2017-12-14	1.80	4.20	2.90	5.04	28074	1.57	13414	3.53
2017-12-12	0.10	2.50	0.72	4.46	26923	2.02	15317	3.54
2017-12-01	-1.70	3.60	0.93	1.75	23617	5.53	13564	3.54
2016-03-01	-5.20	6.40	0.70	4.67	26766	0.87	14952	3.57
2016-02-15	-3.20	2.90	0.36	4.33	23387	3.05	14800	3.57
2015-02-13	-1.10	8.60	3.80	3.54	25519	5.63	15559	3.57
2015-02-02	-0.50	3.80	1.08	3.50	31096	1.44	15983	3.57

## 8.2.6. Прорачун прогнозе потрошње природног гаса неуронском мрежом

Као и у претходном примеру, изабрана библиотека неуронске мреже је *Keras*. Подаци селектованих дана који се користе као улаз у неуронску мрежу су температура и ветар. Прогноза се извршава за сваки сат појединачно.

Хипер параметри који су одређени праћењем кривих учења приказани су у табели 36.

Као улаз у тренинг неуронске мреже користе се вредности температуре, ветра, као и остварена потрошња у претходном сату. Историјски независни подаци садрже податке о потрошњи за претходни сат. У првом сату прогнозе, независни подаци прогнозе такође садрже овај податак, пошто је напоменуто да су у моменту покретања прогнозе на располагању сви историјски подаци. У сваком следећем сату прогнозе, уместо остварене потрошње користи се прогнозирана потрошња претходног сата.

Табела 36 – Оптимизоване вредности хипер параметара неуронске мреже за прогнозу потрошње природног гаса

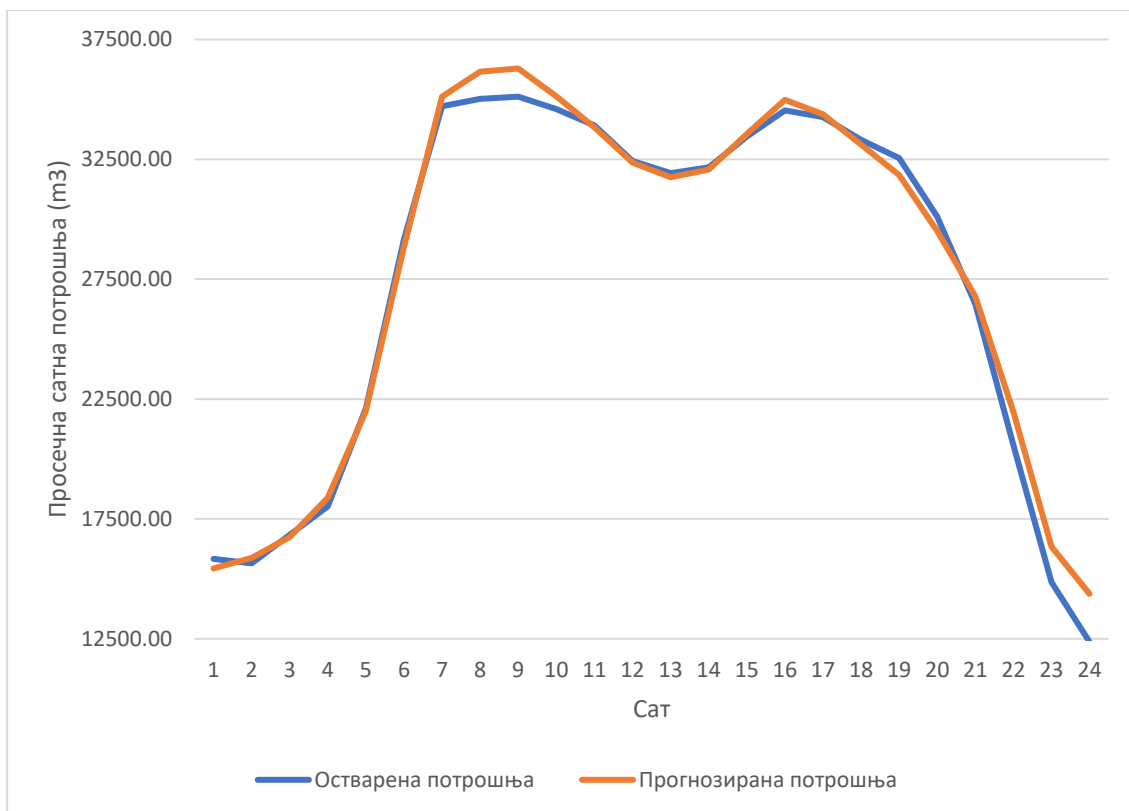
Назив хипер параметра	Опсег вредности	Корак	Оптимизована вредност
Број пакета	1-5	1	1
Број епоха	10-500	1	55
Стопа учења	0,01-1	0,01	0,07
Оптимизатор	Gradient Descent, RMSprop, Adam, Adadelta, Adagrad, Adamax, NAdam, FTRL		Adam
Активациона функција	ReLU, Sigmoid, Softmax, Softplus, Softsign, Hyperbolic tangent, SELU, ELU, Exponential		ReLU
Број скривених слојева	1-4	1	1
Број неурона у скривеном слоју	1-20	1	8
Улазне колоне			Температура Ветар Остварена потрошња у претходном сату

## 8.2.7. Резултати прогнозе за изабрани дан

Табела 37 приказује податке о прогнозираној и оствареној потрошњи природног гаса за сваки појединачни сат прогнозираног дана и апсолутну процентуалну грешку у потрошњи по сату. Коначна *MAPE* вредност прогнозираног дана је 2,32%.

Ови подаци у графичком облику приказани су на слици 63.





Слика 63 – Криве прогнозираних и остварених вредности потрошње природног гаса

Табела 37 – Поређење прогнозиране и остварене потрошње природног гаса

Сат	Прогнозирана потрошња ( $C_f$ ) [m <sup>3</sup> ]	Остварена потрошња ( $C_r$ ) [m <sup>3</sup> ]	$ C_f - C_r $ [m <sup>3</sup> ]	Сатна <i>MAPE</i> вредност [%]
1	15838.09	15437.43	400.66	2.60
2	15650.67	15872.81	222.14	1.40
3	16843.05	16740.18	102.87	0.61
4	18024.28	18389.13	364.84	1.98
5	22115.66	22010.23	105.43	0.48
6	29160.95	28836.07	324.88	1.13
7	34711.96	35116.70	404.74	1.15
8	35027.50	36157.38	1129.88	3.12
9	35114.19	36290.48	1176.29	3.24
10	34601.43	35123.23	521.79	1.49
11	33909.97	33831.44	78.54	0.23
12	32426.06	32361.51	64.55	0.20
13	31931.78	31757.64	174.13	0.55
14	32160.98	32072.30	88.69	0.28
15	33444.02	33550.58	106.56	0.32
16	34547.40	34987.09	439.69	1.26
17	34262.57	34381.68	119.11	0.35
18	33308.64	33110.97	197.67	0.60
19	32547.13	31841.28	705.85	2.22
20	30111.80	29509.53	602.27	2.04
21	26464.72	26777.89	313.18	1.17
22	20572.61	21959.93	1387.32	6.32
23	14860.93	16343.22	1482.29	9.07
24	12374.56	14374.97	2000.41	13.92
<b>Просечна <i>MAPE</i> вредност</b>				<b>2,32</b>

### 8.2.8. Резултати прогнозе за један месец

Табела 38 садржи просечне *MAPE* вредности за 29 дана у зиму на почетку 2018. године. Просечна *MAPE* вредност за цео период износи 3,81%, док се просечне *MAPE* вредности за појединачне дане крећу у распону од 1,88% до 6,86%.

Табела 38 – Резултати прогнозе потрошње природног гаса по данима, коришћењем адаптивног модела прогнозе и неуралне мреже

Дан	Просечна дневна <i>MARE</i> вредност [%]
2018-xx-[01]	3.16
2018-xx-[02]	2.80
2018-xx-[03]	6.86
2018-xx-[04]	2.32
2018-xx-[05]	2.41
2018-xx-[06]	3.48
2018-xx-[07]	3.70
2018-xx-[08]	4.21
2018-xx-[09]	3.60
2018-xx-[10]	3.18
2018-xx-[11]	6.27
2018-xx-[12]	2.26
2018-xx-[13]	3.44
2018-xx-[14]	2.16
2018-xx-[15]	2.11
2018-xx-[16]	4.84
2018-xx-[17]	6.80
2018-xx-[18]	2.60
2018-xx-[19]	3.34
2018-xx-[20]	4.68
2018-xx-[21]	1.88
2018-xx-[22]	4.11
2018-xx-[23]	4.06
2018-xx-[24]	4.40
2018-xx-[25]	3.93
2018-xx-[26]	2.76
2018-xx-[27]	6.38
2018-xx-[28]	4.88
2018-xx-[29]	3.96
<b>Просечна <i>MARE</i> вредност за период [%]</b>	<b>3.81</b>

### 8.2.9. Поређење са резултатима прогнозе потрошње природног гаса коришћењем саме неуронске мреже

За исти период као и у случају прогнозе коришћењем адаптивног модела креирана је и прогноза потрошње природног гаса коришћењем неуронске мреже, без адаптивног модела прогнозе. Узорак за неуронску мрежу је целокупна историја података.

Праћењем кривих учења и валидације дошло се до хипер параметара као што је приказано у табели 39.

Табела 39 – Оптимизоване вредности хипер параметара на примеру саме неуронске мреже за прогнозу потрошње природног гаса

Назив хипер параметра	Опсег вредности	Корак	Оптимизована вредност
Број скривених слојева	1-4	1	2
Број неурона у првом скривеном слоју	1-20	1	20
Број неурона у другом скривеном слоју	1-20	1	6
Број пакета	1-5	1	3
Број епоха	10-500	1	220
Стопа учења	0,01-0,05	0,01	
Оптимизатор	Gradient Descent, RMSprop, Adam, Adadelta, Adagrad, Adamax, NAdam, FTRL		Adam
Активациона функција	ReLU, Sigmoid, Softmax, Softplus, Softsign, Hyperbolic tangent, SELU, ELU, Exponential		ReLU
Улазне колоне	-	-	Дан у недељи Сат Температура Ветар Остварена потрошња у претходном сату

Табела 40 садржи просечне *MAPE* вредности добијене након прорачуна самом неуронском мрежом. Просечна *MAPE* вредност за цео период износи 6,12%.

Из ових резултата види се да адаптивни модел прогнозе значајно доприноси тачности резултата прогнозе потрошње природног гаса у конкретном примеру прогнозе за дистрибутивну мрежу.

Табела 40 – Резултати прогнозе потрошње природног гаса по данима, коришћењем саме неуралне мреже

Датум	Просечна дневна MAPE вредност [%]
2018-xx-[01]	8.41
2018-xx-[02]	3.35
2018-xx-[03]	4.80
2018-xx-[04]	3.63
2018-xx-[05]	5.16
2018-xx-[06]	5.29
2018-xx-[07]	5.35
2018-xx-[08]	8.33
2018-xx-[09]	8.80
2018-xx-[10]	3.75
2018-xx-[11]	6.35
2018-xx-[12]	5.53
2018-xx-[13]	6.07
2018-xx-[14]	8.51
2018-xx-[15]	8.65
2018-xx-[16]	6.24
2018-xx-[17]	3.49
2018-xx-[18]	3.49
2018-xx-[19]	6.26
2018-xx-[20]	3.61
2018-xx-[21]	6.78
2018-xx-[22]	8.69
2018-xx-[23]	8.50
2018-xx-[24]	7.33
2018-xx-[25]	6.14
2018-xx-[26]	5.91
2018-xx-[27]	8.43
2018-xx-[28]	4.15
2018-xx-[29]	6.34
<b>Просечна MAPE вредност за период [%]</b>	<b>6.12</b>

Табела 41 приказује просечно време извршавања процеса прогнозе неуронском мрежом подешене на исти начин, посебно са и без прикљученог адаптивног модела прогнозе. Окружење на коме су извршене прогнозе исте су као у случају претходног експеримента.

Табела 41 – Просечно време извршавања прогнозе потрошње природног гаса  
неуронском мрежом и адаптивним моделом прогнозе

Тип прогнозе	Време				
	дневне метахеуристичке оптимизације	селекције сличних дана	извршавања прогнозе (тренинг + прогноза)	укупно (са метахеуристичком оптимизацијом)	укупно (без метахеуристичке оптимизације)
Неуронска мрежа	-	-	~310 минута	~310 минута	~310 минута
Адаптивни модел прогнозе	~116 минута	~5 секунди	~40 минута	~156 минута	~40 минута

Из табеле се види да употреба адаптивног модела прогнозе позитивно утиче на перформансе. Прорачун који укључује метахеуристичку оптимизацију унапређује брзину извршавања за 154 минута, док у случају одвојеног извршавања метахеуристичке оптимизације унапређење износи приближно 270 минута.

### 8.2.10. Поређење са резултатима других техника краткорочне прогнозе потрошње природног гаса

Као и у случају резултата других техника краткорочне прогнозе потрошње електричне енергије, у прогонзи природног гаса тешко је извршити адекватно поређење две студије случаја које се одвијају на различитим сетовима података. У случају прогнозе потрошње природног гаса овај проблем је још више изражен пошто постоји већа разлика међу географским подручјима у погледу развијености гасне инфраструктуре. Овде се наводе примери резултата краткорочне прогнозе потрошње природног гаса ради стицања опште слике о кретањима вредности резултата адаптивног модела прогнозе и других техника.

Техника предложена у раду [8] користи неуронску мрежу дубоког учења као механизам прогнозе потрошње природног гаса. Студија случаја односи се на прогнозу потрошње у 62 оперативна подручја на територији Сједињених Америчких Држава, за период од једне године. Као показатељ резултата користи се пондерисана просечна апсолутна процентуална грешка (енгл. *WMAPE – Weighted Mean Absolute Percent Error*). У односу на *MAPE*, *WMAPE* искључује специјалне дане из резултата. Преложена техника производи грешку од 5,58%.

Техника представљена у раду [14] је комбинација неуронске мреже и расплинуте логике за краткорочну прогнозу потрошње природног гаса. Студија случаја у овом раду представља прогнозу за подручје Ирана, у периоду 21.06.2008. до 30.06.2008. године. Коришћењем предложене технике за прогнозу, просечна апсолутна процентуална грешка за овај период износи 3%.

## 9. Закључак и правци даљег истраживања

Адаптивни модел прогнозе у инфраструктурним мрежама, који је представљен у дисертацији, има за циљ унапређење постојећих техника прогнозе потрошње ресурса у паметним инфраструктурним мрежама, које се огледа у повећању тачности крајњих резултата прогнозе.

Већина постојећих решења за прогнозу потрошње ресурса односи се на само једну врсту ресурса. Такође, постојећа решења се у највећој мери баве оптимизацијом алгоритама прогнозе, док оптимизација узорка није била заступљена у довољној мери. Идеја која је представљена, а истовремено и кључна предност у односу на постојећа решења, јесте оптимизација узорка за алгоритме машинског учења. Друга предност је могућност примене модела на различите врсте ресурса.

Један од основних изазова оптимизације узорка јесте то што узорци могу да се односе на различите врсте фактора који утичу на квалитет узорка и на саму прогнозу. Са повећањем броја ресурса за које се прогнозира потрошња, проблем разноликости могућих узорака се повећава.

Други изазов представља чињеница да се временом мењају околности у којима се бирају узорци. Оптимално подешен начин избора узорка у једном тренутку не значи да ће тај начин бити одговарајући у промењеним календарским, метеоролошким и друштвеним околностима.

Одговор на представљене изазове представља адаптивност предложеног модела. Адаптивност у овом контексту има следећа значења:

- Адаптивност архитектуре, тако да се могу додавати нови фактори, уклањати постојећи фактори и мењати утицаји појединачних фактора на процес прогнозе.
- Адаптивност прогнозе у односу на промењене околности које могу да утичу на прогнозу. У складу са овим је и адаптивност избора узорка, тако да се у сваком моменту бира оптималан узорак сличних дана из историјске базе података, ради што ефикаснијег извршавања алгорита прогнозе.
- Адаптивност у односу на специфичне, односно специјалне дане, тако да за све дане током године буде извршена оптимална прогноза.

Приказани модел је креиран тако да може да се користи за различите врсте ресурса у инфраструктурним мрежама, и у складу са тим су обрађене теоријске основе краткорочне прогнозе електричне енергије, гаса и воде.

Презентована архитектура омогућава коришћење више различитих фактора који утичу на избор узорка а самим тим и на процес прогнозе. Фактори који су предложени су фактори временских прилика, инерциони фактори, фактор дужине трајања дневног светла и фактор близине дана. Поред ових фактора, архитектура је отворена за увођење нових и укидање постојећих фактора.

Оптимизација избора узорка у приказаном моделу извршава се континуално, коришћењем унапређеног генетског алгоритма као изабране метахеуристичке методе. Унапређени генетски алгоритам изабран је теоријском и практичном анализом. Теоријска анализа подразумевала је разматрање потреба приказаног модела и могућности постојећих решења. Практична анализа подразумевала је експеримент провере ефикасности оптимизације ројева честица, стандардног генетског алгоритма и унапређеног генетског алгоритма. Унапређени генетски алгоритам је показао најбоље резултате и стога је изабран.

Метахеуристичком оптимизацијом генетским алгоритмом препознаје се утицај сваког појединачног фактора. Утицај фактора представљен је увођењем тежинских коефицијената за сваки појединачан фактор.

Адаптивни модел прогнозе омогућава прилагођавање фактора промењеним календарским и временским околностима. Извршавање оптимизације обавља се на дневном и кварталном нивоу. Дневна оптимизација извршава се на периоду од последње две недеље, док се квартална оптимизација извршава на периоду последње године. Механизам оптимизације поседује могућност поређења резултата добијених кварталном и дневном оптимизацијом. Уколико се успостави да годишња оптимизација даје боље резултате, параметри добијени дневном оптимизацијом се изједначавају са параметрима добијеним кварталном оптимизацијом. На тај начин се спречава да дневна оптимизација на мањем узорку одведе прорачун у погрешном правцу.

Генетски алгоритам је унапређен увођењем иновативних техника које омогућавају ефикаснију оптимизацију. Унапређење се налази у корацима селекције, укрштања и мутације. Концепти који су уведени названи су *концепт насељавања*, *концепт мутације засноване на диверзитету* и *концепт мутације ка екстремима*.

У модел је уграђен механизам за препознавање специјалних дана и за оптимизацију прогнозе за овакве дане. Овај механизам се ослања на симулацију прогнозе у историјским специјалним данима, утврђивање тренда грешака у таквим симулацијама и примена тих трендова на реални прогнозирани специјални дан.



На основу адаптивног модела прогнозе временских серија креирана је апликација која се успешно користи у електроенергетском преносном систему „Електро mreжа Србије“. У овом раду приказани су дизајн и архитектура апликације.

Приказани модел је верификован са две студије случаја. Прва студија случаја односи се на поступак прогнозе електричне енергије у преносној мрежи, док се друга студија случаја односи на прогнозу потрошње природног гаса у дистрибутивној мрежи.

У првој студији случаја испраћен је поступак прогнозе електричне енергије у преносној мрежи за један обичан дан и један специјални дан, у реалном систему који користи предложени модел. И у случају прогнозе за обичан дан и у случају прогнозе за специјални дан модел је дао одговарајуће резултате у односу на захтеве преносног система који користи апликацију која имплементира модел. Приказани су и резултати за једну годину, и упоређени су са резултатима прогнозе коришћењем саме неуронске мреже. Адаптивни модел прогнозе показује боље резултате у односу на саму неуронску мрежу, и у погледу резултата прогнозе и у погледу перформанси.

У другој студији случаја креирана је симулација прогнозе потрошње природног гаса у дистрибутивној мрежи за један дан. Процес прогнозе у овој симулацији даје резултате очигледно боље у односу на прогнозу коришћењем саме неуронске мреже, без адаптивног модела прогнозе. Ови резултати су бољи и у погледу резултата и у погледу перформанси.

## **9.1. Правци даљег истраживања**

Један од праваца даљег истраживања може да обухвати анализу практичне примене модела над другим врстама ресурса, као што је вода.

Како модел поседује могућност коришћења различитих алгоритама прогнозе, студија случаја може да обухвати анализу резултата разврстаних по различитим алгоритмима, који поред неуронске мреже обухватају и линеарну регресију и метод потпорних вектора.

У даљем раду биће размотрено одређивање оптималног броја дана у узорку за алгоритам прогнозе. Овај број може да се разликује у зависности од тога који алгоритам прогнозе је изабран. Такође, постоји простор за разматрање минималног квалитета сличног дана, да би ушао у узорак. Овде се под квалитетом сличног дана подразумева вредност коефицијента сличности.

Логично се поставља и питање како се може управљати прогнозом у случају да не постоји довољан број сличних дана са очекиваним квалитетом. Оваква ситуација се између осталог може десити у случајевима прогнозе потрошње у дану са екстремним временским условима у неочекиваном календарском добу године (на пример екстремно високе температуре у мају или септембру, или екстремно ниска температура у октобру или априлу). У таквим случајевима постоји могућност да историјска база података не

садржи довољан број дана чија сличност са прогнозираним даном би била довољна за успешну прогнозу. Даље истраживање ће ићи у правцу решавања оваквих ситуација.

У студији случаја приказано је кретање вредности дневних параметара током времена. У даљим правцима рада ови подаци се могу искористити у анализи утицаја појединачних фактора кроз време.

# Литература

- [1] B. Yildiz, J. Bilbao and A. Sproul, "A review and analysis of regression and machine learning models on commercial building electricity load forecasting," *Renewable and Sustainable Energy Reviews*, vol. 73, pp. 1104-1122, 2017.
- [2] A. Sinha, R. Tayal, A. Vyas, P. Pandey and O. P. Vyas, "Forecasting Electricity Load With Hybrid Scalable Model Based on Stacked Non Linear Residual Approach," *Frontiers in Energy Research*, vol. 9, 2021.
- [3] T. Doan and J. Kalita, "Selecting Machine Learning Algorithms Using Regression Models," in *IEEE International Conference on Data Mining Workshop (ICDMW)*, 2015.
- [4] S. Rong and Z. Bao-wen, "The research of regression model in machine learning field," in *MATEC Web of Conferences, 6th International Forum on Industrial Design*, 2018.
- [5] K. Y. Lee, Y. T. Cha and J. H. Park, "Short-term load forecasting using an artificial neural network," *IEEE Transactions on Power Systems*, vol. 7, no. 1, pp. 124-132, 1992.
- [6] D. Srinivasan, A. Liew and C. Chang, "A neural network short-term load forecaster," *Electric Power Systems Research*, vol. 28, no. 3, pp. 227-234, 1994.
- [7] D.-x. Niu, S. Hui-feng and W. Desheng Dash, "Short-term load forecasting using bayesian neural networks learned by Hybrid Monte Carlo algorithm," *Applied Soft Computing*, vol. 12, no. 6, pp. 1822-1827, 2012.
- [8] G. D. Merkel, R. J. Povinelli and R. H. Brown, "Short-Term Load Forecasting of Natural Gas with Deep Neural Network Regression," *Energies*, vol. 11, no. 8, pp. 1-12, 2018.
- [9] D. Ivezić, "Short-term natural gas consumption forecast," *FME Transactions*, vol. 34, pp. 165-169, 2006.
- [10] A. Jain, A. Kumar Varshney and U. Chandra Joshi, "Short-Term Water Demand Forecast Modelling at IIT Kanpur Using Artificial Neural Networks," *Water Resources Management*, vol. 15, no. 5, pp. 299-321, 2001.

- [11] J. Blancas and J. Noel, "Short-Term Load Forecasting Using Fuzzy Logic," in *2018 IEEE PES Transmission & Distribution Conference and Exhibition - Latin America (T&D-LA)*, Lima, 2018.
- [12] Rustum Mamlook, Omar Badran and Emad Abdulhadi, "A Fuzzy Inference Model Forshort-Term Load Forecasting," *Energy Policy, Volume 37, Issue 4*, pp. 1239-1248, 2009.
- [13] D. Ranaweera, N. Hubele and G. Karady, "Fuzzy logic for short term load forecasting," *International Journal of Electrical Power & Energy Systems*, vol. 18, no. 4, pp. 215-222, 1996.
- [14] A. Azadeh, A. Ghanbari and S. M. Asadzadeh, "An adaptive network-based fuzzy inference system for short-term natural gas demand estimation: Uncertain and complex environments," *Energy Policy*, vol. 38, no. 3, pp. 1529-1536, 2010.
- [15] A. Selakov, D. Cvijetinović, L. Milović, S. Mellon and D. Bekut, "Hybrid PSO–SVM Method For Short-Term Load Forecasting During Periods With Significant Temperature Variations In City Of Burbank," *Applied Soft Computing, Volume 16*, pp. 80-88, 2014.
- [16] I. Msiza, F. Nelwamondo and T. Marwala, "Water Demand Prediction using Artificial Neural Networks and Support Vector Regression," *JOURNAL OF COMPUTERS*, vol. 3, no. 11, pp. 1-8, 2008.
- [17] M. Bata, R. Carriveau and D. S.-K. Ting, "Short-term water demand forecasting using hybrid supervised and unsupervised machine learning model," *Smart Water*, vol. 5, no. 2, pp. 1-18, 2020.
- [18] C. Hakan Aladag and E. Eğrioğlu, "Advanced Time Series Forecasting Methods," *Advances in Time Series Forecasting*, pp. 3-10, 2012.
- [19] T. Hong, Short Term Electric Load Forecasting (doctoral dissertation), Raleigh, North Carolina: North Carolina State University, 2010.
- [20] A. Селаков, *Оптимално управљање микро мрежама у карактеристичним радним режимима*, Нови Сад: Универзитет у Новом Саду, Факултет техничких наука, 2017.
- [21] M. E.Khodayar and H. Wu, "Demand Forecasting in the Smart Grid Paradigm: Features and Challenges," *The Electricity Journal*, vol. 28, no. 6, pp. 51-62, 2015.
- [22] G. Gross and F. D. Galiana, "Short-term load forecasting," *Proceedings of the IEEE*, vol. 75, no. 12, pp. 1558-1573, 1987.

- [23] P. Ganguly, A. Zayegh and A. Kalam, "Short Term Load Forecasting Using Fuzzy Logic," in *Conference: International Conference on Research in Education and Science (ICRES), May 18 - 21, 2017, Ephesus-Kusadasi/Turkey*, 2017.
- [24] M. Fagiani, S. Squartini, L. Gabrielli, S. Spinsante and F. Piazza, "A review of datasets and load forecasting techniques for smart natural gas and water grids: Analysis and experiments," *Neurocomputing*, vol. 170, pp. 448-465, 2015.
- [25] M. Fagiani, S. Squartini, R. Bonfigli and F. Piazza, "Short-term load forecasting for smart water and gas grids: A comparative evaluation," in *IEEE 15th International Conference on Environment and Electrical Engineering (EEEIC)*, 2015.
- [26] Z. Janković, A. Selakov, D. Bekut and M. Đorđević, "Day Similarity Metric Model For Short Term Load Forecasting Supported By PSO and Artificial Neural Network," *Electrical Engineering*, 2021.
- [27] N. Amral, C. S. Ozveren and D. King, "Short term load forecasting using Multiple Linear Regression," in *42nd International Universities Power Engineering Conference*, 2007.
- [28] B. Beudet and R. L. Roberts, "A Perspective on the Global Water Marketplace," *Journal of American Water Works Association*, vol. 92, no. 4, pp. 10-12.
- [29] R. Pillay, *Short-term water demand forecasting for production optimisation (dissertation)*, University of Southern Queensland Faculty of Engineering and Surveying, 2005.
- [30] L. Lorente-Leyva, J. Pavón-Valencia, Y. Montero-Santos, I. Herrera-Granda, E. Herrera-Granda and D. Peluffo-Ordóñez, "Artificial Neural Networks for Urban Water Demand Forecasting: A Case Study," *Journal of Physics: Conference Series*, vol. 1284, pp. 1-8, 2019.
- [31] A. J. Abebe and R. Price, "Managing uncertainty in hydrological models using complementary models," *Hydrological Sciences Journal*, vol. 48, no. 5, pp. 679-692, 2003.
- [32] A. Boretti and L. Rosa, "Reassessing the projections of the World Water Development Report," *npj Clean Water*, vol. 2, pp. 1-6, 2019.
- [33] "Natural gas explained," U.S. Energy Information Administration, 9 12 2020. [Online]. Available: <https://www.eia.gov/energyexplained/natural-gas/>. [Accessed 24 4 2021].
- [34] R. Svoboda, V. Kotik and J. Platos, "Short-term natural gas consumption forecasting from long-term data collection," *Energy*, vol. 218, 2021.

- [35] A. Samuel, "Some Studies in Machine Learning Using the Game of Checkers," *IBM Journal of Research and Development*, p. 210–229, 1959.
- [36] S. Gollapudi, *Practical Machine Learning*, Birmingham B3 2PB, UK: Packt Publishing, 2016.
- [37] R. P. Masini, M. C. Medeiros and E. F. Mendes, "Machine Learning Advances for Time Series Forecasting," arXiv.org, 2012.
- [38] G. Bontempi, S. Ben Taieb and Y. Le Borgne, "Machine Learning Strategies for Time Series Forecasting," *Business Intelligence*, vol. 138, pp. 62-77, 2013.
- [39] J. F. Torres, D. Hadjout, A. Sebaa, F. Martínez-Álvarez and A. Troncoso, "Deep Learning for Time Series Forecasting: A Survey," *Big Data*, vol. 9, no. 1, pp. 3-21, 2021.
- [40] Y. Ensafi, S. Hassanzadeh Amin, G. Zhang and B. Shah, "Time-series forecasting of seasonal items sales using machine learning – A comparative analysis," *International Journal of Information Management Data Insights*, vol. 2, no. 1, p. 100058, 2022.
- [41] E. Eğrioğlu and V. R. Uslu, "New Criteria to Compare Interval Estimates in Fuzzy Time Series Methods," *Advances in Time Series Forecasting*, pp. 56-63, 2012.
- [42] J. M. Utts and R. F. Heckard, *Mind on Statistics*, Stamford, CT 06902 USA: Cengage Learning, 2015.
- [43] I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," *SN Computer Science*, vol. 2, no. 3, pp. 1-21, 2021.
- [44] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practices*, OTEXTS, 2018.
- [45] Ü. Çavuş and Ş. Ertekin, "Improving forecasting accuracy of time series data using a new ARIMA-ANN hybrid method and empirical mode decomposition," *Neurocomputing*, vol. 361, pp. 151-163, 2019.
- [46] M. Vrahatis, G. Magoulas, K. Parsopoulos and V. P. Plagianakos, "Introduction to artificial neural network training and applications," in *15th Annual Conference of Hellenic Society for Neuroscience, (Neuroscience 2000)*, Patras, Greece, 2000.
- [47] H. S. Hippert, C. E. Pedreira and R. C. Souza, "Neural Networks for Short-Term Load Forecasting: A Review and Evaluation," *IEEE Transactions on Power Systems*, vol. 16, no. 1, pp. 44-55, 2001.
- [48] A. Pradhan, "Support vector machine-A survey," *International Journal of Emerging Technology and Advanced Engineering*, vol. 2, no. 8, 2012.

- [49] K. -R. Müller, A. J. Smola, G. Rätsch, B. Schölkopf<sup>2</sup>, J. Kohlmorgen and V. Vapnik<sup>3</sup>, "Predicting time series with support vector machines," in *International Conference on Artificial Neural Networks*, Berlin, Heidelberg, 1997.
- [50] C. Elkan, "The paradoxical success of fuzzy logic," *IEEE Expert*, vol. 9, no. 4, pp. 698-703, 1994.
- [51] E. Egrioglu, C. HakanAladag, U. Yolcu, V. R. Uslu and M. A. Basaran, "A new approach based on artificial neural networks for high order multivariate fuzzy time series," *Expert Systems with Applications*, vol. 36, no. 7, pp. 10589-10594, 2009.
- [52] D. Fiorillo, Z. Kapelan, M. Xenochristou, F. De Paola and M. Giugni , "Assessing the Impact of Climate Change on Future Water Demand using Weather Data," *Water Resources Management*, vol. 35, p. 1449–1462, 2021.
- [53] J. Černi, "Temperature Impact on Drinking Water Consumption," in *4th EWaS International Conference: Valuing the Water, Carbon, Ecological Footprints of Human Activities*, Online, 2020.
- [54] M. Xenochristou, Z. Kapelan, C. Hutton and J. Hofman, "Smart Water Demand Forecasting: Learning from the Data," *EPiC Series in Engineering*, vol. 3, pp. 2351-2358, 2018.
- [55] R. P. Timmer and P. J. Lamb, "Relations between Temperature and Residential Natural Gas Consumption in the Central and Eastern United States," *Journal of Applied Meteorology and Climatology*, vol. 46, no. 11, p. 1993–2013, 2007.
- [56] J. Kang and D. M. Reiner, "What is the effect of weather on household electricity consumption?: Empirical evidence from Ireland," *Energy Policy Research Group, University of Cambridge*, 2021.
- [57] I. Staffell and S. Pfenninger, "The increasing impact of weather on electricity supply and demand," *Energy*, vol. 145, pp. 65-78, 2018.
- [58] C. Zhang, H. Liao and Z. Mi , "Climate impacts: temperature and electricity consu," *Natural Hazards*, vol. 99, p. 1259–1275, 2019.
- [59] G. Raidl and J. Puchinger, "Combining (Integer) Linear Programming Techniques and Metaheuristics for Combinatorial Optimization," *Hybrid Metaheuristics*, vol. 114, pp. 31-62, 2008.
- [60] R. Babu and B. Bhattacharyya, "Strategic Placements of PMUs for Power Network Observability Considering Redundancy Measurement," *Measurement*, vol. 134, pp. 606-623, 2019.

- [61] R. Babu and B. Bhattacharyya, "Weak Bus-Oriented Installation of Phasor Measurement Unit for Power Network Observability," *International Journal of Emerging Electric Power Systems*, vol. 18, no. 8, 2017.
- [62] L. Bianchi, M. Dorigo, L. M. Gambardella and W. J. Gutjahr, "A survey on metaheuristics for stochastic combinatorial optimization," *Natural Computing*, vol. 8, p. 239–287, 2009.
- [63] S. Athiyarath, M. Paul and S. Krishnaswamy, "A Comparative Study and Analysis of Time Series Forecasting Techniques," *SN Computer Science*, vol. 1, 2020.
- [64] E. A. Feinberg and D. Genethliou, "Load Forecasting," in *Applied Mathematics for Restructured Electric Power Systems. Power Electronics and Power Systems.*, Boston, MA., Springer, 2005.
- [65] I. Kofi Nti, M. Teimeh, O. Nyarko-Boateng and A. Felix Adekoya, "Electricity load forecasting: a systematic review," *Journal of Electrical Systems and Information Technology*, vol. 7, 2020.
- [66] Z. Wang, J. Li, S. Zhu, J. Zhao, S. Deng, S. Zhong, H. Yin, H. Li, Y. Qi and Z. Gan, "A Review of Load Forecasting of the Distributed Energy System," in *IOP Conference Series: Earth and Environmental Science*, 2019.
- [67] A. A. Mamun, . M. Sohel, M. Naeem, H. Sunny, M. Samiul, D. R. Dipta and E. Hossain, "A Comprehensive Review of the Load Forecasting Techniques Using Single and Hybrid Predictive Models," *IEEE Access*, vol. 8, pp. 134911-134939, 2020.
- [68] S. M. Al-Fattah, "Time Series Modeling for U.S. Natural Gas Forecasting," in *IPTC-10592 paper presented at SPE International Petroleum Technology*, Doha, Qatar, 2005.
- [69] W. Panek and T. Włodek, "Natural Gas Consumption Forecasting Based on the Variability of External Meteorological Factors Using Machine Learning Algorithms," *Energies*, vol. 15, no. 1, 2022.
- [70] B. Soldo, "Forecasting natural gas consumption," *Applied Energy*, vol. 92, pp. 26-37, 2012.
- [71] G. De Souza Groppo, M. Azevedo Costa and M. Libânio, "Predicting water demand: a review of the methods employed and future possibilities," *Water Supply*, vol. 19, no. 8, 2019.
- [72] Y. Zheng, W. Zhang , J. Xie and Q. Liu, "A Water Consumption Forecasting Model by Using a Nonlinear Autoregressive Network with Exogenous Inputs Based on Rough Attributes," *Water*, vol. 14, no. 329, 2022.



- [73] Ying Chen, Peter B. Luh, Che Guan, Yige Zhao, Laurent D. Michel, Matthew A. Coolbeth, Peter B. Friedland and Stephen J. Rourke, "Short-Term Load Forecasting: Similar Day-Based Wavelet Neural Networks," *IEEE Transactions On Power Systems*, Vol. 25, No. 1, 2010.
- [74] Paras Mandal, Tomonobu Senjyu and Toshihisa Funabashi, "Neural networks approach to forecast several hour ahead electricity prices and loads in deregulated market," *Energy Conversion and Management*, Volume 47, Issues 15–16, pp. 2128-2142, 2006.
- [75] Paras Mandal, Tomonobu Senjyu, Naomitsu Urasaki and Toshihisa Funabashi, "A Neural Network Based Several-Hour-Ahead Electric Load Forecasting Using Similar Days Approach," *Electrical Power and Energy Systems* 28, p. 367–373, 2006.
- [76] Tomonobu Senjyu, Hitoshi Takara, Katsumi Uezato and Toshihisa Funabashi, "One-Hour-Ahead Load Forecasting Using Neural Network," *IEEE Transactions On Power Systems*, Vol. 17, No. 1, p. 113, 2002.
- [77] Wei Sun and Yujun He, "Optimal Support Vector Machine Based Short–Term Load Forecasting Model with Input Variables and Samples Selection," in *Third International Conference on Intelligent Computing, ICIC*, Qingdao, China, 2007.
- [78] Che Guan, Peter B. Luh, Matthew Coolbeth and Stephen J. Rourke, "Very short-term load forecasting: Multilevel wavelet neural networks with data pre-filtering," in *IEEE Xplore Conference: Conference: Power & Energy Society General Meeting*, 2009.
- [79] Rahmat-Allah Hooshmand, Habib Amooshahi and Moein Parastegari, "A hybrid intelligent algorithm based short-term load forecasting approach," *Electrical Power and Energy Systems* 45, pp. 313-324, 2018.
- [80] Tomonobu Senjyu, Hirokazu Sakihara, Yoshinori Tamaki and Katsumi Uezato, "Next Day Peak Load Forecasting Using Neural Network With Adaptive Learning Algorithm Based On Similarity," *Electric Machines and Power Systems*, 28, pp. 613-624, 2000.
- [81] Akbar Ebrahimi and Amir Moshari, "Holidays short-term load forecasting using fuzzy improved similar day method," *International Transactions On Electrical Energy Systems*, p. 1254–1271, 2013.
- [82] Irena Koprinska, Mashud Rana, Alicia Troncoso and Francisco Martínez-Álvarez, "Combining Pattern Sequence Similarity with Neural Networks for Forecasting Electricity Demand Time Series," in *International Joint Conference on Neural Networks (IJCNN)*, 2013.
- [83] Amit Jain and M Babita Jain, "Fuzzy Modeling and Similarity based Short Term Load Forecasting using Swarm Intelligence-A step towards Smart Grid," in *Proceedings of*

*Seventh International Conference on Bio-Inspired Computing: Theories and Applications*, 2012.

- [84] Amit Jain, E. Srinivas and Rasmimayee Rauta, "Short Term Load Forecasting using Fuzzy Adaptive Inference and Similarity," *World Congress on Nature & Biologically Inspired Computing (NaBIC)*, 2009.
- [85] X. Kong, C. Li, F. Zheng and C. Wang, "Improved Deep Belief Network for Short-Term Load Forecasting Considering Demand-Side Management," *IEEE Transactions on Power Systems*, vol. 35, no. 2, pp. 1531-1538, 2020.
- [86] Qingqing Mu, Yonggang Wu, Xiaoqiang Pan, Liangyi Huang and Xian Li, "Short-term Load Forecasting Using Improved Similar Days Method," in *Asia-Pacific Power and Energy Engineering Conference*, 2010.
- [87] Ming-Yue Zhai, "A new method for short-term load forecasting based on fractal interpretation and wavelet analysis," in *Electrical Power and Energy Systems, Vol 69*, 2015.
- [88] G. Dudek, "Pattern Similarity-based Methods for Short-Term Load Forecasting – Part 1: Principles," *Applied Soft Computing* 36, pp. 422-441, 2015.
- [89] G. Dudek, "Pattern similarity-based methods for short-term load forecasting –Part 2: Models," *Applied Soft Computing* 36, p. 422–441, 2015.
- [90] M. Karimi, H. Karami, M. Gholami, H. Khatibzadehazad and N. Moslemi, "Priority index considering temperature and date proximity for selection of similar days in knowledge-based short term load forecasting method," *Energy, Vol. 144*, pp. 928-940, 2018.
- [91] Huiting Zheng, Jiabin Yuan and Long Chen, "Short-Term Load Forecasting Using EMD-LSTM Neural Networks with a Xgboost Algorithm for Feature Importance Evaluation," *Energies*, 2017.
- [92] Wei Sun and Chongchong Zhang, "A Hybrid BA-ELM Model Based on Factor Analysis and Similar-Day Approach for Short-Term Load Forecasting," *Energies*, 2018.
- [93] M. Torabi, S. Hashemi, M. R. Saybani, S. Shamshirband and A. Mosavif, "A Hybrid Clustering and Classification Technique for Forecasting Short-Term Energy Consumption," *Environmental Progress & Sustainable Energy*, vol. 38, no. 1, pp. 66-76, 2019.

- [94] H. Liu, Y. Wang, C. Wei, J. Li and Y. Lin, "Two-Stage Short-Term Load Forecasting for Power Transformers Under Different Substation Operating Conditions," *IEEE Access*, vol. 7, pp. 161424-161436, 2019.
- [95] Y. Shen, J. Zhang, L. Jin, P. Zhang, R. Chen and Y. Chen, "Short-term load forecasting of power system based on similar day method and PSO PSO-DBN," in *Conference: 2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2)*, 2018.
- [96] A. Bracale, P. Caramia, P. De Falco and T. Hong, "Multivariate Quantile Regression for Short Term Probabilistic Load Forecasting," *IEEE Transactions on Power Systems*, vol. 35, pp. 628-638, 2020.
- [97] G. Romano, N. Salvati and A. Guerrini, "Estimating the Determinants of Residential Water Demand in Italy," *Water*, vol. 6, pp. 2929-2945, 2014.
- [98] A. Antunes, A. Andrade-Campos, A. Sardinha-Lourenço and M. Oliveira, "Short-term water demand forecasting using machine learning techniques," *Journal of Hydroinformatics*, vol. 20, pp. 1-24, 2018.
- [99] R. Benítez, C. Ortiz-Caraballo, J. C. Preciado, J. M. Conejero, F. Sánchez Figueroa and A. Rubio-Largo, "A Short-Term Data Based Water Consumption Prediction Approach," *Energies*, vol. 12, pp. 1-24, 2019.
- [100] R. B. Cleveland, W. S. Cleveland, J. E. McRae and I. Terpenning, "SLT: A Seasonal-Trend Decomposition Procedure Based on Loses," *Journal of Official Statistics*, vol. 6, no. 1, pp. 3-33, 1990.
- [101] K. Smolak, B. Kasieczka, W. Fialkiewicz, W. Rohm, K. Siła-Nowicka and K. Kopańczyk, "Applying human mobility and water consumption data for short-term water demand forecasting using classical and machine learning models," *Urban Water Journal*, vol. 17, no. 1, pp. 32-42, 2020.
- [102] R. H. Brown, S. R. Vitullo, G. F. Corliss, M. Adya, P. E. Kaefer and R. J. Povinelli, "Detrending daily natural gas consumption series to improve short-term forecasts," in *IEEE Power & Energy Society General Meeting*, Denver, CO, USA, 2015.
- [103] K. Sorensen and F. W. Glover, "Metaheuristics," in *Encyclopedia of Operations Research and Management Science*, S. I. Gass and M. C. Fu, Eds., Boston, MA, Springer US, 2013, pp. 960-970.
- [104] M. Dorigo, M. Birattari and T. Stutzle, "Ant colony optimization," *IEEE Computational Intelligence Magazine*, vol. 1, no. 4, pp. 28-39, 2006.

- [105] M. Dorigo, V. Maniezzo and A. Colorni, "Ant System: Optimization by a colony of cooperating agents," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 26, no. 1, pp. 29-41, 1996.
- [106] T. Stützle and H. H. Hoos, "MAX-MIN Ant System," *Future Generation Computer Systems*, vol. 16, no. 8, pp. 889-914, 2000.
- [107] M. Dorigo and L. Gambardella, "Ant colony system: a cooperative learning approach to the traveling salesman problem," *IEEE Transactions on Evolutionary Computation*, vol. 1, no. 1, pp. 53-66, 1997.
- [108] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, Perth, WA, Australia, 1995.
- [109] G. Lindfield and J. Penny, "Chapter 3 - Particle Swarm Optimization Algorithms," in *Introduction to Nature-Inspired Optimization*, G. Lindfield and J. Penny, Eds., Cambridge, Massachusetts, Academic Press, 2017, pp. 49-68.
- [110] Randy L. Haupt and Sue Ellen Haupt, *Practical Genetic Algorithms*, Second Edition, New York City, United States: John Wiley & Sons, Inc., 2004.
- [111] S. Katoch, S. S. Chauhan and V. Kumar, "A review on genetic algorithm: past, present, and future," *Multimedia Tools and Applications*, vol. 80, no. 5, pp. 8091-8126, 2021.
- [112] K. Jebari, "Selection Methods for Genetic Algorithms," *International Journal of Emerging Sciences, Abdelmalek Essaâdi University*, vol. 3, no. 4, pp. 332-344, 2013.
- [113] L. Murat, A. Volkan, B. Necaattin, P. Huseyin and C. Ertugrul, "Short-Term Fuzzy Load Forecasting Model Using Genetic-Fuzzy and Ant Colony-Fuzzy Knowledge Base Optimization," *Applied Sciences*, vol. 8, no. 6, pp. 1-30, 2018.
- [114] Y.-S. Huang and J.-J. Deng, "Short-Term Load Forecasting Based on Ant Colony Fuzzy Clustering and SVM Algorithm," in *Fourth International Conference on Natural Computation*, Jinan, China, 2008.
- [115] O. C. Ozerdem, E. O. Olaniyi and O. K. Oyedotun, "Short term load forecasting using particle swarm optimization neural network," *Procedia Computer Science*, vol. 120, pp. 382-393, 2017.
- [116] Azzam-ul-Asar, S. R. ul Hassnain and A. Khan, "Short Term Load Forecasting Using Particle Swarm Optimization Based ANN Approach," in *International Joint Conference on Neural Networks*, Orlando, FL, USA, 2007.

- [117] Z. A. Bashir and M. El-Hawary, "Short-Term Load Forecasting Using Artificial Neural Network Based on Particle Swarm Optimization Algorithm," in *Canadian Conference on Electrical and Computer Engineering*, Vancouver, BC, Canada, 2007.
- [118] Y. Wang, Y. Jing, W. Zhao and Y.-e. Mao, "Dynamic neural network based genetic algorithm optimizing for short term load forecasting," in *Chinese Control and Decision Conference*, Xuzhou, China, 2010.
- [119] D.-B. Islam, Z. Baharudin, M. Qamar Raza and P. Nallagownden, "A Hybrid Neural Network and Genetic Algorithm Based Model for Short Term Load Forecast," *Research Journal of Applied Sciences, Engineering and Technology*, vol. 13, pp. 2667-2673, 2014.
- [120] S. M. Moghaddas-Tafreshi and M. Farhad, "A linear regression-based study for temperature sensitivity analysis of Iran electrical load," in *IEEE International Conference on Industrial Technology*, Chengdu, China, 2008.
- [121] L. Donghai, "Mathematical modeling analysis of genetic algorithms under schema theorem," *IOS Press*, vol. 19, pp. 131-137, 2019.
- [122] R. Tinós and S. Yang, "A self-organizing random immigrants genetic algorithm for dynamic optimization problems," *Genetic Programming and Evolvable Machines*, pp. 256-286, 2007.
- [123] "Keras," [Online]. Available: <https://keras.io/>. [Accessed 25 6 2021].
- [124] C. Perlich, "Learning Curves in Machine Learning," in *Encyclopedia of Machine Learning*, Boston, MA, Springer US, 2010, pp. 577-580.
- [125] C. Liu, Z. Jin, J. Gu and C. Qiu, "Short-term load forecasting using a long short-term memory network," in *IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, 2017.
- [126] M. Xenochristou, M. Blokker, I. Vertommen, J. F. Urbanus and Z. Kapelan, "Investigating the Influence of Weather on Water Consumption: a Dutch Case Study," in *1st International WDSA / CCWI 2018 Joint Conference*, Kingston, Ontario, Canada, 2018.



## План третмана података

<b>Назив пројекта/истраживања</b>
Адаптивни модел прогнозе временских серија у паметним инфраструктурним мрежама
<b>Назив институције/институција у оквиру којих се спроводи истраживање</b>
Универзитет у Новом Саду, Факултет техничких наука, Департаман за енергетику, електронику и телекомуникације
<b>Назив програма у оквиру ког се реализује истраживање</b>
Докторске академске студије
<b>1. Опис података</b>
<p>1.1 Врста студије</p> <p><i>Укратко описати тип студије у оквиру које се подаци прикупљају</i></p> <p>Докторска дисертација у оквиру које је приказан напредни адаптивни модел краткорочне прогнозе вредности временских серија у паметним инфраструктурним мрежама. Појаве које модел прогнозира односе се пре свега на потрошњу ресурса као што су електрична енергија, природни гас и вода. Модел користи постојеће методе машинског учења, које унапређује како у погледу тачности прогнозе тако и у погледу перформанси извршења прорачуна.</p>
<p>1.2 Врсте података</p> <p>а) квантитативни</p> <p>б) квалитативни</p>
<p>1.3. Начин прикупљања података</p> <p>а) анкете, упитници, тестови</p> <p>б) клиничке процене, медицински записи, електронски здравствени записи</p> <p>в) генотипови: навести врсту</p>

г) административни подаци: навести врсту **Подаци о сатној потрошњи електричне енергије и природног гаса за дане експеримента, метеоролошки подаци**

д) узорци ткива: навести врсту

ђ) снимци, фотографије: навести врсту

е) текст, навести врсту **преглед литературе**

ж) мапа, навести врсту

з) остало: описати

### 1.3 Формат података, употребљене скале, количина података

#### 1.3.1 Употребљени софтвер и формат датотеке:

а) Ехсел фајл, датотека **.xlsx**

б) SPSS фајл, датотека

в) PDF фајл, датотека **.pdf**

г) Текст фајл, датотека **.docx**

д) JPG фајл, датотека **.jpg**

е) Остало, датотека **.csv**

#### 1.3.2. Број записа (код квантитативних података)

а) број варијабли **велики број**

б) број мерења (испитаника, процена, снимака и сл.) **велики број**

#### 1.3.3. Поновљена мерења

а) да

б) не

Уколико је одговор да, одговорити на следећа питања:

а) временски размак између поновљених мера је \_\_\_\_\_

б) варијабле које се више пута мере односе се на \_\_\_\_\_

в) нове верзије фајлова који садрже поновљена мерења су именоване као \_\_\_\_\_

Напомене: \_\_\_\_\_

Да ли формати и софтвер омогућавају дељење и дугорочну валидност података?

- a) **Да**
- б) *Не*

Ако је одговор не, образложити \_\_\_\_\_

\_\_\_\_\_

## 2. Прикупљање података

### 2.1 Методологија за прикупљање/генерисање података

2.1.1. У оквиру ког истраживачког нацрта су подаци прикупљени?

- a) експеримент, навести тип
- б) корелационо истраживање, навести тип **Анализа утицаја метеоролошких, календарских и инерционих података на потрошњу електричне енергије и природног гаса**
- ц) анализа текста, навести тип
- д) остало, навести шта

2.1.2 Навести врсте мерних инструмената или стандарде података специфичних за одређену научну дисциплину (ако постоје).

### 2.2 Квалитет података и стандарди

2.2.1. Третман недостајућих података

- a) Да ли матрица садржи недостајуће податке? **Да Не**

Ако је одговор да, одговорити на следећа питања:

- a) Колики је број недостајућих података? \_\_\_\_\_
- б) Да ли се кориснику матрице препоручује замена недостајућих података? **Да Не**
- в) Ако је одговор да, навести сугестије за третман замене недостајућих података

2.2.2. На који начин је контролисан квалитет података? Описати



**Квалитет података је контролисан поређењем са јавно доступним подацима и увидом у резултате истраживања.**

2.2.3. На који начин је извршена контрола уноса података у матрицу?

**Контрола уноса података је изведена на основу експертног знања.**

### 3. Третман података и пратећа документација

3.1. Третман и чување података

*3.1.1. Подаци ће бити депоновани у [https://github.com/bpzoran/disertacija\\_podaci.git](https://github.com/bpzoran/disertacija_podaci.git) репозиторијум.*

*3.1.2. URL адреса <https://www.cris.uns.ac.rs/theses.jsf>*

*3.1.3. DOI <https://doi.org/10.1007/s00202-021-01286-6>*

*3.1.4. Да ли ће подаци бити у отвореном приступу?*

*а) Да*

*б) Да, али после ембарга који ће трајати до \_\_\_\_\_*

*в) Не*

*Ако је одговор не, навести разлог \_\_\_\_\_*

*3.1.5. Подаци неће бити депоновани у репозиторијум, али ће бити чувани.*

*Образложење*

---

---

3.2. Метаподаци и документација података

3.2.1. Који стандард за метаподатке ће бити примењен? \_\_\_\_\_

3.2.1. Навести метаподатке на основу којих су подаци депоновани у репозиторијум.

---

---

---

*Ако је потребно, навести методе које се користе за преузимање података, аналитичке и процедуралне информације, њихово кодирање, детаљне описе варијабли, записа итд.*

---

---

---

---

---

---

### 3.3 Стратегија и стандарди за чување података

3.3.1. До ког периода ће подаци бити чувани у репозиторијуму? **На неодређено време**

3.3.2. Да ли ће подаци бити депоновани под шифром? Да **Не**

3.3.3. Да ли ће шифра бити доступна одређеном кругу истраживача? Да **Не**

3.3.4. Да ли се подаци морају уклонити из отвореног приступа после извесног времена?

Да **Не**

Образложити

---

---

## 4. Безбедност података и заштита поверљивих информација

Овај одељак МОРА бити попуњен ако ваши подаци укључују личне податке који се односе на учеснике у истраживању. За друга истраживања треба такође размотрити заштиту и сигурност података.

### 4.1 Формални стандарди за сигурност информација/података

Истраживачи који спроводе испитивања с људима морају да се придржавају Закона о заштити података о личности ([https://www.paragraf.rs/propisi/zakon\\_o\\_zastiti\\_podataka\\_o\\_licnosti.html](https://www.paragraf.rs/propisi/zakon_o_zastiti_podataka_o_licnosti.html)) и одговарајућег институционалног кодекса о академском интегритету.

4.1.2. Да ли је истраживање одобрено од стране етичке комисије? Да **Не**

Ако је одговор Да, навести датум и назив етичке комисије која је одобрила истраживање

---

4.1.2. Да ли подаци укључују личне податке учесника у истраживању? Да **Не**

Ако је одговор да, наведите на који начин сте осигурали поверљивост и сигурност информација везаних за испитанике:

- а) Подаци нису у отвореном приступу
- б) Подаци су анонимизирани
- ц) Остало, навести шта

---

---

## 5. Доступност података

5.1. Подаци ће бити

- а) јавно доступни
- б) доступни само уском кругу истраживача у одређеној научној области
- ц) затворени

Ако су подаци доступни само уском кругу истраживача, навести под којим условима могу да их користе:

---

---

Ако су подаци доступни само уском кругу истраживача, навести на који начин могу приступити подацима:

---

---

5.4. Навести лиценцу под којом ће прикупљени подаци бити архивирани.

---

## 6. Улоге и одговорност

6.1. Навести име и презиме и мејл адресу власника (аутора) података

**Зоран Јанковић, zoran.jankovic@uns.ac.rs**

6.2. Навести име и презиме и мејл адресу особе која одржава матрицу с подацима

**Зоран Јанковић, zoran.jankovic@uns.ac.rs**

6.3. Навести име и презиме и мејл адресу особе која омогућује приступ подацима другим истраживачима

**Зоран Јанковић, zoran.jankovic@uns.ac.rs**