

UNIVERZITET U BEOGRADU
TEHNOLOŠKO-METALURŠKI FAKULTET

Davor Z. Antanasijević

**MODELOVANJE INDIKATORA
KVALITETA ŽIVOTNE SREDINE
PRIMENOM VEŠTAČKIH NEURONSKIH
MREŽA**

doktorska disertacija

Beograd, 2014.

UNIVERSITY OF BELGRADE
FACULTY OF TECHNOLOGY AND METALLURGY

Davor Z. Antanasijević

**ENVIRONMENTAL INDICATOR
MODELING USING
ARTIFICIAL NEURAL NETWORKS**

Doctoral Dissertation

Belgrade, 2014

Podaci o mentoru i članovima komisije

Mentor:

Dr **Viktor Pocajt**, vanredni profesor
Univerzitet u Beogradu, Tehnološko-metalurški fakultet

Članovi komisije:

Dr **Mirjana Ristić**, redovni profesor
Univerzitet u Beogradu, Tehnološko-metalurški fakultet

Dr **Aleksandra Perić-Grujić**, redovni profesor
Univerzitet u Beogradu, Tehnološko-metalurški fakultet

Dr **Antonije Onjia**, naučni savetnik
Univerzitet u Beogradu, Institut za nuklearne nauke „Vinča“

Datum odbrane: 26.05.2014. godine

*“As far as the laws of mathematics refer to reality,
they are not certain;
and as far as they are certain,
they do not refer to reality.”*

Albert Einstein

ZAHVALNICA

Ova doktorska disertacija je rađena na Katedri za zaštitu životne sredine Tehnološko-metalurškog fakulteta Univerziteta u Beogradu, u periodu od 2011. do 2014. godine.

Pre svega bih želeo da se zahvalim mom mentoru, prof. dr Viktoru Pocaјtu. Bila mi čast i zadovoljstvo da radim sa njim i učim od njega. Veoma sam mu zahvalan na izdvojenom vremenu, idejama, savetima i podršci tokom izrade ove disertacije. Njegov entuzijazam me je motivisao tokom svih ovih godina. Takođe sam mu zahvalan za odličan primer koji mi je pružio kao uspešna ličnost i profesor.

Takođe bih želeo da se zahvalim prof. dr Mirjani Ristić, na podršci i savetima tokom naučno-istraživačkog rada u okviru ove doktorske disertacije.

Zahvaljujem se i prof. dr Aleksandri Perić-Grujić na pomoći, primedbama i korisnim sugestijama.

Na kraju bih želeo da se zahvalim mojoj porodici na nesebičnoj podršci, zato što su uvek bili uz mene i verovali u moj uspeh. Mojoj Jeleni dugujem veliku zahvalnost na ljubavi i pažnji koju mi je pružala svih ovih godina, kao i na razumevanju tokom dana kada sam bio isključivo okupiran izradom disertacije. Ona je moja najveća podrška, ali i najveći kritičar.

Modelovanje indikatora kvaliteta životne sredine primenom veštačkih neuronskih mreža

REZIME

Održivi razvoj i očuvanje životne sredine ne mogu se ostvariti bez poznavanja vrednosti indikatora koji opisuju kvalitet životne sredine. Direktive Evropske unije, kao i Ženevska konvencija nalažu predviđanje vrednosti emisija zagađujućih materija više godina unapred, pa je upotreba odgovarajućih modela u tu svrhu neophodna. Predviđanja vrednosti indikatora kvaliteta životne sredine i održivog razvoja se sve češće, poslednjih godina, vrše primenom različitih statističkih modela i simulacijama. Sa napretkom ostvarenim u oblasti računarske tehnologije, veštačke neuronske mreže (ANN) postale su veoma zastupljene kod modelovanja složenih, kao i nelinearnih problema.

U ovoj doktorskoj disertaciji je prikazan razvoj modela zasnovanih na veštačkim neuronskim mrežama za predviđanje vrednosti odabranih indikatora kvaliteta životne sredine, pri čemu su kao ulazne promenljive korišćeni dostupni ekonomski indikatori i indikatori održivog razvoja. Za statističke modele koji se uglavnom primenjuju u oblasti zaštite životne sredine neophodno je obezbediti veliki broj ulaznih podataka koji često nisu poznati ili je njihovo određivanje veoma složeno. Primenom modela razvijenih u okviru ove doktorske disertacije prevaziđen je problem nedostatka podataka neophodnih za primenu postojećih modela, a dobijene su vrednosti indikatora koji opisuju kvalitet životne sredine, koje često nisu poznate na nacionalnom nivou, kako kod razvijenih zemalja, tako i kod zemalja u razvoju. Glavni ciljevi ove disertacije su bili razvoj ANN metodologije i modela za predviđanje sledećih indikatora kvaliteta životne sredine: emisija gasova staklene bašte i PM₁₀ čestica, koncentracija PM₁₀ čestica na nacionalnom nivou, količine generisanog komunalnog čvrstog otpada (KČO) i sadržaja rastvorenog kiseonika u površinskim vodama.

ANN model za predviđanje emisije gasova staklene bašte je najpre optimizovan sa stanovišta parametara arhitekture i obučavanja, dok je izbor ulaznih promenljivih

vršen korelacionom analizom i analizom glavnih komponenti (PCA). Performanse dobijenih ANN modela su upoređene sa odgovarajućim modelima zasnovanim na višestrukoj linearnoj regresiji (MLR) i regresiji glavnih komponenti (PCR). ANN model za koji su ulazne promenljive odabrane korelacionom analizom imao je najbolje performanse i zbog toga je primenjen za predviđanje GHG emisija u 2011. godini. Vrednosti relativnih grešaka ANN modela za 2010. godinu su korišćene za korekciju predviđanja za 2011. godinu. Analiza značajnosti ulaznih promenljivih je pokazala da pojedinačno najveći uticaj na GHG emisiju ima ukupna potrošnja energije na nacionalnom nivou.

Kod predviđanja vrednosti godišnjih emitovanih količina PM_{10} čestica na nacionalnom nivou, eliminacija manje značajnih/uticajnih ulaznih promenljivih vršena je primenom genetskih algoritama. Na ovaj način je smanjen broj ulaznih promenljivih modela, a promenljive/indikatora koje su korišćene u modelu su dostupne na nacionalnom nivou, čime se omogućuje prevazilaženje problema nedostatka vrednosti osnovnih indikatora u mnogim zemljama. Model je razvijen korišćenjem podataka za 26 EU zemalja za period od 1999-2006. godine. Poređenje predviđanja razvijenog ANN modela sa aktuelnim vrednostima PM_{10} emisija je pokazalo da se ovom metodologijom emisije mogu uspešno i precizno odrediti do dve godine unapred. Srednja relativna greška za dvogodišnje predviđanje emisija PM_{10} čestica je bila 10 %, što je tri puta manja greška u odnosu na predviđanja dobijena primenjenim MLR i PCR modelima, koji su razvijeni i testirani istim podacima i ulaznim indikatorima.

ANN model za predviđanje PM_{10} koncentracija na nacionalnom nivou razvijen je korišćenjem istih ulaznih promenljivih/indikatora kao i u slučaju modelovanja emitovanih PM_{10} čestica. Za razvoj modela korišćeni su podaci za odabrane zemlje Evrope, za period 2000-2008. godina. Selekcija najuticajnijih, iz grupe primarno odabranih ulaznih promenljivih, je vršena korelacionom analizom. Srednja relativna greška predviđanja ANN modela je iznosila 13 %. Široka dostupnost odabranih ulaznih promenljivih/indikatora ANN modela omogućuje prevazilaženje problema nedostatka vrednosti osnovnih indikatora u mnogim zemljama.

U okviru ove doktorske disertacije takođe je ispitana i mogućnost primene ANN modela i indikatora održivog razvoja kod predviđanja količina generisanog komunalnog čvrstog otpada. Za razvoj ovog ANN modela korišćeni su podaci zemalja Evropske unije, a pošto su ovakvi modeli posebno značajni za zemlje u razvoju, kod kojih ne postoje podaci o količinama nastalog KČO, razvijeni ANN model je primenjen na Bugarsku i Srbiju. Zaključeno je da se u ovoj doktorskoj disertaciji razvijena ANN metodologija, korišćenjem samo osnovnih indikatora održivog razvoja (bruto domaći proizvod, upotreba materijala i produktivnost resursa), može uspešno primeniti za predviđanje generisanih količina KČO na nacionalnom nivou za zemalje u razvoju, kao i za razvijene i nerazvijene zemalje.

Cilj poslednjeg istraživanja, rađenog u okviru ove disertacije, je bio razvoj ANN modela za predviđanje sadržaja rastvorenog kiseonika u površinskim vodama uz korišćenje nespecifičnih parametara kvaliteta voda, kao ulaznih promenljivih modela (temperatura, pH vrednost, električna provodljivost i protok rečne vode). Primenjeni su podaci dobijeni monitoringom kvaliteta reke Dunav, sa monitoring stanice Bezdan. Poređenje najboljeg ANN modela sa odgovarajućim MLR modelom je pokazalo da je ANN model značajno precizniji, jer je kod testiranja dao pojedinačna predviđanja sa greškom manjom od 10 %.

Na osnovu dobijenih rezultata pokazano je da se ANN metodologija, razvijena u okviru ove disertacije, može koristiti za modelovanje i predviđanje veoma različitih indikatora kvaliteta životne sredine, primenom široko dostupnih ekonomskih indikatora i indikatora održivog razvoja.

Ključne reči: održivi razvoj, nelinearno modelovanje, gasovi staklene bašte, čestično zagađenje, komunalni čvrsti otpad, sadržaj rastvorenog kiseonika

Naučna oblast: Tehničko-tehnološke nauke

Uža naučna oblast: Inženjerstvo zaštite životne sredine

UDK broj: 66.0:581.52

Environmental indicators modeling using artificial neural networks

SUMMARY

Sustainable development and environmental protection cannot be accomplished without knowing the values of basic environmental indicators, which describe the state of environment and effectiveness of applied environmental strategies and policies. European Union directives and the Geneva Convention require the forecasting of the values of pollutant emission and concentration several years in advance, and therefore the use of appropriate models for this purpose is necessary. In recent years, the usage of statistical models and simulation for predicting the values of environmental and sustainability indicators has been increased. Also, with the advent of modern computers, artificial neural networks (ANNs) become common modeling tool, mainly used for complex and nonlinear problem modeling.

The present dissertation has demonstrated that ANNs can be successfully applied in predicting values of different environmental indicators using broadly available sustainability, economical and industrial indicators acting as inputs, in order to overcome the problem of incomplete environmental and sustainability data, which is notable in both developed and developing countries. The main goals of this research were to develop ANN-based methodology and models for the prediction of greenhouse gases (GHGs) and PM10 emissions at the national level, country-level PM10 concentrations, municipal waste generation (MWG) and dissolved oxygen (DO) concentration.

The ANN model architecture and training parameters were optimized, with inputs being selected using correlation analysis and principal component analysis. The developed ANN models were compared with the corresponding multiple linear regression (MLR) model, while the ANN model created using transformed inputs (principal components) was compared with a principal component regression (PCR) model. Since the best results were obtained with the ANN model based on correlation analysis, that particular model was selected for the actual 2011 GHG

emissions forecasting. The relative errors of the 2010 GHG emissions predictions were used to adjust the ANN model predictions for 2011, which subsequently resulted in the adjusted 2011 predictions having a MAPE value of only 3.60%. Sensitivity analysis showed that gross inland energy consumption has the highest sensitivity to GHG emissions.

In the case of annual PM10 emissions modeling at the national level, the inputs for the model were selected and optimized using a genetic algorithm. The wide availability of the input parameters used in this model can overcome a lack of data and basic environmental indicators in many countries, which can prevent or seriously impede PM emission forecasting. The model was trained and validated with the data for 26 EU countries for the period from 1999 to 2006. The ANN model has shown very good performance and demonstrated that the forecast of PM10 emission up to two years can be made successfully and accurately. The mean relative error for two-year PM10 emission prediction was only 10%, which is more than three times better than the predictions obtained from the conventional multilinear regression and principal component regression models that were trained and tested using the same datasets and input variables.

An ANN model for predicting country-level PM10 concentrations was trained and validated with the data for 23 European Union (EU) countries plus the EU27 as a group for the period from 2000 to 2008. The inputs for the model were selected using correlation analyses. Country-level PM10 concentration data that were used as a model output were obtained from the World Bank. The mean relative error for the ANN model prediction, in the case of most of the EU countries, was less than 13%, indicating stable and accurate predictions. The wide availability of input parameters used in this model can overcome the problem of lack of data in many countries, which can in turn prevent the determination of human exposure to PM10 at the national level.

The feasibility of modeling municipal waste generation for countries at different levels of development using ANNs and selected sustainability indicators was also investigated. Data from 26 European countries was used as training, test and vali-

dation datasets for the development of ANN models. Since this kind of modeling is particularly important for developing countries, where MWG data is missing or incomplete, emphasis was placed on the modeling of MWG for Bulgaria and Serbia. Based on comparison of actual MWG data with predictions given by the model, it was shown that ANNs can be applied successfully to the modeling and forecasting of MWG on the national scale. Moreover, the scope for possible application of the model is broad, since it uses generic indicators of sustainability such as gross domestic product, domestic material consumption and resource productivity, and performs well for countries with highly diversified levels of economic development, industrial structure, productivity and output.

The aim of last research made under this dissertation was to create an ANN-based model using non-specific water quality parameters for the prediction of dissolved oxygen (DO) concentration in the Danube River. The input variables used for the ANN model are water flow, temperature, pH and electrical conductivity. The model was trained and validated using the available data from 2004 to 2008 and tested using the data from 2009. The comparison of the ANN model with the MLR model indicates that the ANN model performs much better, since all predictions of the ANN model for the test data were within the error of less than $\pm 10\%$.

Key words: nonlinear modeling, air quality, pollutant emission, sustainable development, GHG, PM10, MWG, dissolved oxygen, European Union, Serbia

Scientific area: Technical-technological sciences

Scientific sub-area: Environmental Engineering

UDC number: 66.0:581.52

Sadržaj

Spisak skraćenica i simbola	iv
Spisak slika	viii
Spisak tabela	x
1. Uvod	1
Teorijski deo	
2. Održivi razvoj i modelovanje indikatora održivog razvoja i kvaliteta životne sredine	4
2.1. Principi održivog razvoja.....	4
2.2. Indikatori održivog razvoja i kvaliteta životne sredine	6
2.3. Prikaz odabranih indikatora održivog razvoja i kvaliteta životne sredine	9
2.3.1. Emisija gasova staklene bašte (GHG) na nacionalnom nivou	9
2.3.2. Emisija PM ₁₀ čestica na nacionalnom nivou	10
2.3.3. Koncentracija PM ₁₀ čestica na nacionalnom nivou	11
2.3.4. Količina generisanog komunalnog čvrstog otpada	13
2.3.5. Sadržaj rastvorenog kiseonika u površinskim vodama	14
2.4. Prikaz odabranih modela koji se koriste za predviđanje emisija i koncentracija zagađujućih materija	15
2.4.1. Inventari emisija	15
2.4.2. RAINS model	16
2.4.3. GAINS model	17
2.4.4. GMAPS model	18
2.4.5. Modeli zasnovani za višestrukoj linearnoj regresiji (MLR)	19
2.4.6. Modeli zasnovani na regresiji glavnih komponenti (PCR)	20
3. Veštačke neuronske mreže (ANN)	21
3.1. Osnove veštačkih neuronskih mreža	21
3.2. Obučavanje neuronskih mreža	28
3.3. Prikaz odabranih arhitektura neuronskih mreža	32
3.3.1. Neuronske mreže sa propagacijom greške unazad (BPNN)	32
3.3.2. Rekurentne neuronske mreže (RNN)	36
3.3.3. Neuronske mreže sa opštom regresijom (GRNN)	37
3.4. Oblasti primene neuronskih mreža	42

4. Hemometrijske metode primenjene kod odabira ulaznih promenljivih modela	44
4.1. Korelaciona analiza.....	44
4.2. Analiza glavnih komponenti (PCA)	44
4.3. Genetski algoritmi (GA)	49

Eksperimentalni deo

5. Metodologija razvoja modela zasnovanog na neuronskim mrežama	52
5.1. Proces razvoja modela zasnovanog na neuronskim mrežama	52
5.2. Prikupljanje relevantnih ulaznih promenljivih modela.....	57
5.3. Statistički pokazatelji performansi modela	58

Rezultati i Diskusija

6. Modelovanje indikatora kvaliteta životne sredine.....	61
6.1. Modelovanje emisije gasova staklene bašte.....	61
6.1.1. Ulazne promenljive i priprema podataka za GHG model	61
6.1.2. Rezultati optimizacije parametara neuronske mreže	66
6.1.3. Rezultati GHG emisionih modela razvijenih odabirom ulaznih promenljivih na osnovu korelacione analize	73
6.1.4. Rezultati GHG modela kreiranih sa ulaznim promenljivim transformisanim primenom analize glavnih komponenti.....	77
6.1.5. Analiza značajnosti pojedinih ulaznih promenljivih CA-GRNN3 modela	80
6.1.6. Ispitivanje mogućnosti primene razvijenog modela za predviđanje emisije gasova staklene bašte u 2011. godini	81
6.2. Modelovanje emisije PM ₁₀ čestica	86
6.2.1. Ulazne promenljive i priprema podataka za PM ₁₀ model	86
6.2.2. Razvoj i optimizacija GRNN modela	90
6.2.3. Rezultati optimizacije GRNN modela	94
6.2.4. Poređenje GRNN10 modela sa linearnim modelima.....	96
6.2.5. Analiza grešaka kod GRNN10 modela.....	100
6.2.6. Analiza značajnosti pojedinih ulaznih promenljivih GRNN10 modela.	102
6.3. Modelovanje koncentracije PM ₁₀ čestica	103
6.3.1. Podaci o PM ₁₀ koncentracijama na nacionalnom nivou.....	103
6.3.2. Razvoj i rezultati optimizacije GRNN modela.....	105
6.3.3. Poređenje GRNN modela za predviđanje PM ₁₀ koncentracije sa linearnim modelom	107
6.4. Modelovanje količina generisanog komunalnog čvrstog otpada	111

6.4.1. Ulazni indikatori i priprema podataka	111
6.4.2. Ulazni podaci za Republiku Srbiju	114
6.4.3. Razvoj ANN modela za predviđanje količine KČO	115
6.4.4. Analiza rezultata razvijenih ANN modela	115
6.4.5. Analiza grešaka GRNN modela	121
6.4.6. Primena GRNN modela na Srbiju	122
6.5. Modelovanje sadržaja rastvorenog kiseonika u površinskim vodama	125
6.5.1. Odabir ulaznih promenljivih modela za predviđanje sadržaja rastvorenog kiseonika	125
6.5.2. Razvoj i rezultati ANN modela za predviđanje sadržaja rastvorenog kiseonika	129
6.5.3. Poređenje RNN modela sa linearnim modelom	133
7. Zaključak	135
Literatura	140
Biografija autora	
Prilozi	

Spisak skraćenica i simbola

α_k – Stopa obučavanja (*learning rate*)

A_k – Hessian matrica drugog izvoda funkcije greške

a_{nm} – Linearni koeficijenti

AAS – Atomska apsorpciona spektroskopija

AES – Atomska emisiona spektroskopija

ANN – Veštačke neuronske mreže (*Artificial neural networks*)

Ar – Stopa aktivnosti (*activity rate*)

β_k – Impuls (*momentum*)

BDP – Bruto domaći proizvod

BDPI – Udeo industrije u BDP-u

BDPP – Udeo poljoprivrede u BDP-u

BP – Algoritam za propagaciju greške unazad (*Backpropagation algorithm*)

BPE – Bruto potrošnja energije

BPNN – Neuronska mreža sa propagacijom greške unazad (*Backpropagation neural network*)

CLRTAP – *Convention on Long-range Transboundary Air Pollution*

$\delta(W)$ – Površina greške

D_i – Rastojanje podataka za obučavanje od vrednosti dobijenih mrežom u multidimenzionom prostoru

DO – Rastvoreni kiseonik (*dissolved oxygen*)

DPM – Domaća potrošnja materijala

EEOI – Električna energija proizvedena iz obnovljivih izvora

E_f – Emisioni faktor (*emission factor*)

EMAS – Broj organizacija sa EMAS (*Eco-Management and Auditing Scheme*) stadardom

E_p – Emisija određenog polutanta

EC – Evropska komisija (*European Comission*)

EU – Evropska unija (*European Union*)

EU15 – Evropska unija sa 15 članica

EU27 – Evropska unija sa 27 članica

EEA – Evropska agencija za zaštitu životne sredine (*European Environment Agency*)

F – Aktivaciona funkcija

FA1.1 – Udeo tačnih predviđanja u odnosu 1:1,1 u odnosu na tačnu vrednost

FA1.25 – Udeo tačnih predviđanja u odnosu 1:1,25 u odnosu na tačnu vrednost

GA – Genetski algoritam

GAINS – *Greenhouse Gas and Air Pollution Interactions and Synergies Model*

GHG – Gasovi staklene bašte (*Greenhouse gases*)

g_k – Gradijent greške modela

GMAPS – *The Global Model of Ambient Particulates*

GRNN – Neuronska mreža sa opštom regresijom (*General regression neural network*)

HPK – Hemijska potrošnja kiseonika

I – Jedinična matrica (*identity matrix*)

IA – Indeks slaganja (*the index of agreement*)

IC – Jonska hromatografija (*ion chromatography*)

ICP-MS – Indukovano spregnuta plazma sa masenom spektroskopijom (*inductively coupled plasma mass spectrometry*)

ICP-OES – Indukovano spregnuta plazma sa atomskom emisionom spektroskopijom (*inductively coupled plasma optical emission spectrometry*)

IASA – *International Institute for Applied Systems Analysis*

ISF – Pojedinačni faktor ujednačavanja (*individual smoothing factor*)

KČO – Komunalni čvrsti otpad

KGO – Količina generisanog otpada

KMO – *Kaiser-Meyer-Olkin* test adekvatnosti podataka

λ – Svojstvena vrednost (*eigenvalue*) korelacione matrice

MAE – Srednja apsolutna greška (*the mean absolute error*)

MAPE – Srednja relativna greška (*the mean absolute percentage error*)

MLP – Višeslojna perceptron mreža (*multi-layer perceptron*)

MLR – Višestruka linearna regresija (*multiple linear regression*)

N_{iz} – Broj neurona u izlaznom sloju

NMVOC – Nemetanska isparljiva organska jedinjenja (*non-methane volatile organic compounds*)

N_{sk} – Broj neurona u skrivenom sloju

n_{tr} – Broj serija podataka tokom obučavanja modela

N_{ul} – Broj neurona u ulaznom sloju

OECD – Organizacija za ekonomsku saradnju i razvoj (*Organization for Economic Co-operation and Development*)

OOD – Količina odloženog otpada na deponije

PC – Glavne komponente (*principal components*)

PCA – Analiza glavnih komponenti (*principal component analysis*)

PCR – Regresija glavnih komponenti (*principal component regression*)

PDG – Proizvodnja drvene građe

p_k – Konjugovani gradijent greške

PM – Čestice (*particulate matter*)

PM_{2,5} – Čestice prečnika manjeg od 2,5 μm

PM₁₀ – Čestice prečnika manjeg od 10 μm

PNN – Probabilistička neuronska mreža (*probabilistic neural networks*)

PPA – Broj putnika prevezenih aviotransportom

PR – Produktivnost resursa

PSV – Prosečna starost vozila

ρ – Veličina iterativnog koraka BP algoritma

$\rho(X_1, X_2)$ – Koeficijent linearne korelacije između dve promenljive X_1 i X_2

r – Parsonov koeficijent korelacije

R^2 – Koeficijent linearne korelacije

RAINS – *Regional Air Pollution Information and Simulation Model*

RMSE – Koren srednje kvadratne greške (*the root mean squared error*)

RNN – Rekurentna neuronska mreža (*Recurrent neural network*)

RPA – Količina robe prevezene aviotransportom

r_q – Procenat ukupne varijanse obuhvaćene odabranim glavnim komponentama

σ_f – Faktor ujednačavanja (*smoothing factor*)

σ_x, σ_y – Standardna devijacija

SEPA – Agencija za zaštitu životne sredine Republike Srbije (*Serbian environmental protection agency*)

toe – Tona ekvivalentne nafte

UEOI – Udeo energije dobijene iz obnovljivih izvora

UN – Ujedinjene nacije (*United Nations*)

UNECE – Komisija Ujedinjenih Nacija za Evropu (*United Nations Economic Commission for Europe*)

UNFCCC – *The United Nation Framework Convention on Climate Change*

UPOP – Udeo površina pod organskom proizvodnjom

US EPA – *United States Environmental Protection Agency*

UV-VIS – Ultra-ljubičasta i vidljiva spektrofotometrija

VIF – *Variance inflation factor*

W – Matrica težinskih koeficijenata

WB – Svetska Banka (*World Bank*)

WCED – Svetska komisija za životnu sredinu i razvoj (*World Commission on Environment and Development*)

WHO – Svetska zdravstvena organizacija (*World Health Organization*)

w_i – Težinski koeficijenti

X_i – Ulazna promenljiva

Y – Izlazna promenljiva

Y_{ac} – Izmerena vrednost izlazne promenljive

Spisak slika

Slika 1. Šematski prikaz a) biološkog i b) veštačkog neurona sa glavnim delovima	22
Slika 2. Osnovne aktivacione funkcije: a) linearna b) odskočna c) sigmoidna	22
Slika 3. Struktura neuronske mreže na primeru troslojne mreže	23
Slika 4. Nadgledano obučavanje neuronske mreže	25
Slika 5. Nenadgledano obučavanje neuronske mreže	25
Slika 6. Arhitekture <i>Backpropagation</i> neuronskih mreža	27
Slika 7. Arhitekture: a) GRNN b) PNN c) Kohonenove mreže	27
Slika 8. Primer modularne neuronske mreže	28
Slika 9. Izgled odziva neuronske mreže u zavisnosti od stepena obučenosti	31
Slika 10. Troslojna BPNN mreža sa standardnom vezom	34
Slika 11. Arhitektura rekurentne neuronske mreže	37
Slika 12. GRNN arhitektura	40
Slika 13. Odziv GRNN mreže u zavisnosti od vrednosti faktora ujednačavanja	41
Slika 14. <i>Cattell scree test</i> - određivanje broja značajnih glavnih komponenti	46
Slika 15. Struktura standardnog genetskog algoritma [142]	50
Slika 16. Šematski prikaz metodologije razvoja neuronske mreže	53
Slika 17. Podela dostupnih podataka tokom obučavanja mreže	55
Slika 18. Sektori GHG emisije i odabrani indikatori [157]	63
Slika 19. Parametri arhitekture i obučavanja GRNN mreže varirani tokom procesa optimizacije	67
Slika 20. Poređenje aktuelnih i modelima predviđenih vrednosti GHG emisija u 2010. godini	73
Slika 21. Poređenje aktuelnih vrednosti GHG emisija sa predviđanjima CA-GRNN3 i CA-MLR modela za 2010. godinu	76
Slika 22. Poređenje aktuelnih vrednosti GHG emisija sa predviđanjima PCA-GRNN3 i PCR modela za 2010. godinu	79
Slika 23. Poređenje aktuelnih i korigovanih modelom predviđenih emisije gasova staklene bašte za 2011. godinu	85
Slika 24. Veza odabranih ulaznih promenljivih i PM ₁₀ emisije	88
Slika 25. Algoritam eliminacione procedure za odabir ulaznih promenljivih	92
Slika 26. Algoritam procedure, za odabir ulaznih promenljivih modela, zasnovane na „glavnim“ ulaznim promenljivim (ulaz = ulazna promenljiva)	93
Slika 27. Poređenje aktuelnih i modelima predviđenih PM ₁₀ emisija za 2005. i 2006. godinu: a) PCR10 b) GRNN10	99
Slika 28. MAPE vrednosti predikcija GRNN10 i PCR10 modela za pojedinačne EU države	100
Slika 29. Poređenje rezultata CA-PCR modela sa aktuelnim vrednostima PM ₁₀ koncentracija	108

Slika 30. Poređenje rezultata CA-GRNN modela sa aktuelnim vrednostima PM ₁₀ koncentracija.....	109
Slika 31. MAPE vrednosti predviđanja CA-GRNN i CA-PCR modela za pojedinačne EU države	110
Slika 32. Poređenje aktuelnih vrednosti generisanog KČO sa predviđanjima modela tokom obučavanja: a) BPNN i b) GRNN	116
Slika 33. Poređenje aktuelnih količina generisanog KČO sa predviđanjima modela tokom provere razvijenih modela (2003 godina): a) BPNN i b) GRNN	118
Slika 34. Poređenje aktuelnih količina generisanog KČO sa predviđanjima modela tokom provere razvijenih modela (2004 godina): a) BPNN i b) GRNN	119
Slika 35. Poređenje aktuelnih vrednosti količina KČO sa predviđanjima GRNN modela za Bugarsku i Srbiju.....	123
Slika 36. Lokacije monitoring stanica na Dunavu (Srbija)	126
Slika 37. Poređenje izmerenog sadržaja rastvorenog kiseonika i vrednosti dobijenih modelima tokom obučavanja: a) GRNN, b) BPNN, c) RNN	130
Slika 38. Poređenje izmerenog sadržaja rastvorenog kiseonika i vrednosti dobijenih modelima tokom provere: a) GRNN, b) BPNN, c) RNN.....	132
Slika 39. Poređenje izmerenog sadržaja rastvorenog kiseonika i rezultata MLR modela za: a) obučavanje b) proveru	133
Slika 40. Poređenje izmerenog i modelima (RNN i MLR) predviđenog sadržaja rastvorenog kiseonika za 2009. godinu.....	134

Spisak tabela

Tabela 1. Lista ulaznih promenljivih GHG modela sa osnovnim statističkim pokazateljima.....	64
Tabela 2. Statistička analiza (minimalne, maksimalne i srednje vrednosti) GHG emisije u državama Evropske unije i po grupama podataka.....	65
Tabela 3. Lista GRNN modela razvijenih tokom optimizacije parametara mreže..	69
Tabela 4. Vrednosti statističkih pokazatelja performansi za GRNN modele razvijene tokom optimizacije prema definisanim parametrima mreže (provera podacima iz 2010. godine).....	70
Tabela 5. Vrednosti statističkih pokazatelja performansi za kreirane GRNN modele tokom procesa optimizacije parametara	71
Tabela 6. Koeficijenti modela sa višestrukom linearnom regresijom (MLR model)	72
Tabela 7. Rezultati korelacione analize ulaznih promenljivih GHG modela	74
Tabela 8. Vrednosti statističkih pokazatelja performansi modela za GRNN i MLR modele kreirane na osnovu rezultata korelacione analize	76
Tabela 9. Lista ulaznih promenljivih na koje je primenjen PCA. Komunaliti vrednosti (<i>communalities extracted</i>), KMO vrednost i <i>Bartlett</i> test statistika.....	77
Tabela 10. PC skor koeficijent matrica, svojstvene vrednosti korelacione matrice (<i>eigenvalues</i>), odgovarajuće varianse i ukupna varijansa definisana glavnim komponentama.....	78
Tabela 11. Koeficijenti PCR GHG modela	79
Tabela 12. Statistički pokazatelji performansi za PCA-GRNN3 i PCR modele.....	79
Tabela 13. Vrednosti ISF-a dobijene tokom obučavanja CA-GRNN3 modela.....	80
Tabela 14. Aktuelne i CA-GRNN3 modelom dobijene GHG emisije za 2010. godinu sa relativnim greškama i odgovarajućim statističkim pokazateljima.....	82
Tabela 15. Aktuelne i CA-GRNN3 modelom dobijene GHG emisija za 2011. godinu sa relativnim greškama i odgovarajućim statističkim pokazateljima.....	83
Tabela 16. Aktuelne i korigovane CA-GRNN3 modelom dobijene GHG emisija za 2011. godinu sa relativnim greškama i odgovarajućim statističkim pokazateljima.....	84
Tabela 17. Spisak dostupnih ulaznih promenljivih i izvora podataka	87
Tabela 18. Podaci za PM ₁₀ emisiju (srednja vrednost za period 1999-2006, vrednosti za 1999., 2004. i 2006.).....	89
Tabela 19. ISF vrednosti ulaznih promenljivih za razvijene GRNN modele eliminacionom procedurom.....	94
Tabela 20. Statistički pokazatelji performansi GRNN modela kreiranih eliminacionom procedurom.....	95
Tabela 21. Statistički pokazatelji performansi GRNN modela kreiranih procedurom sa „glavnim“ ulaznim indikatorima.....	95
Tabela 22. Koeficijenti kreiranih linearnih modela MLR13 i PCR13.....	97

Tabela 23. Koeficijenti kreiranih linearnih modela MLR10 i PCR10	97
Tabela 24. Statistički pokazatelji performansi najboljih linearnih modela (MLR10 i PCR10)	98
Tabela 25. Promena količine emitovanih PM10 čestica za analizirane zemlje.....	101
Tabela 26. Statistička analiza PM ₁₀ koncentracija za period 2000-2008. god.....	104
Tabela 27. Rezultati korelacione analize podataka za razvoj modela za predviđanje PM ₁₀ koncentracija.....	106
Tabela 28. Vrednosti pokazatelja performansi za razvijene GRNN modele za predviđanje PM ₁₀ koncentracija.....	107
Tabela 29. Koeficijenti razvijenog CA-PCR modela.....	108
Tabela 30. Ulazne i izlazne promenljive modela sa jedinicama posle normalizacije	111
Tabela 31. Podaci iz 2000. godine korišćeni za obučavanje modela	113
Tabela 32. Kategorije materijala potrebne za proračun DPM [183].....	114
Tabela 33. Ulazni podaci za Republiku Srbiju.....	115
Tabela 34. Aktuelne količine i predviđanja modela za 2005. godinu (provera modela)	120
Tabela 35. Procenat predviđanja sa relativnom greškom manjom od 10 %.....	121
Tabela 36. Procenjene i GRNN vrednosti količina generisanog KČO sa relativnim greškama	124
Tabela 37. Statistička analiza grupa podataka obučavanje i proveru modela	128
Tabela 38. Grupa podataka, iz 2009. godine, korišćena za proveru modela.....	128
Tabela 39. Vrednosti statističkih pokazatelja performansi kreiranih ANN modela	131
Tabela 40. Koeficijenti kreiranog MLR modela.....	133
Tabela 41. Vrednosti statističkih pokazatelja performansi MLR modela	134

1. Uvod

Očuvanje, unapređenje i zaštita životne sredine, uz održivi društveni i ekonomski razvoj, dominantni su problemi savremenog sveta. Efikasno upravljanje i kontrola navedenih složenih procesa, iziskuje njihovo konstantno praćenje, uz korišćenje adekvatnih indikatora, kojima se kvantifikuje stanje životne sredine i održivog razvoja u različitim delovima sveta. Poznavanje vrednosti određenih indikatora kvaliteta životne sredine od suštinske je važnosti u okviru upravljanja životnom sredinom, jer su oni osnova za donošenje mera i strategija vezanih za poboljšanje stanja u životnoj sredini, ali su u isto vreme i mera ostvarenja zadatih ciljeva i sredstvo za uključivanje javnosti u donošenju odluka vezanih za pitanja zaštite životne sredine.

Osnovni cilj istraživanja opisanih u okviru ove doktorske disertacije je modelovanje emisija različitih zagađujućih materija, čije vrednosti predstavljaju indikatore kvaliteta životne sredine, modelima baziranim na neuronskim mrežama, uz korišćenje dostupnih ekonomskih i industrijskih indikatora i indikatora održivog razvoja, kao ulaznih promenljivih modela. Predmet ove disertacije je razvoj modela za sledeće indikatore kvaliteta životne sredine, koji se prate na nacionalnom nivou: emisija gasova staklene bašte, emisija PM10 čestica, koncentracija PM10 čestica i količina generisanog komunalnog čvrstog otpada, kao i indikator sadržaja rastvorenog kiseonika u površinskim vodama, koji pokazuje kvalitet životne sredine na lokalnom nivou.

Kao polazna hipoteza kod modelovanja navedenih indikatora životne sredine, koji se prate na nacionalnom nivou, pretpostavljena je njihova nelinearna matematička povezanost, kao zavisnih promenljivih, sa ekonomskim indikatorima i indikatorima održivog razvoja, kao nezavisnim promenljivim. Takođe, kod modelovanja sadržaja rastvorenog kiseonika pretpostavljena je nelinearna matematička povezanost specifičnih i nespecifičnih parametara kvaliteta površinskih voda. Na osnovu navedenih pretpostavki, veštačke neuronske mreže su odabrane kao alat za modelovanje, jer su u stanju da vrše predviđanje vrednosti izlaznih parametara, koji su nelinearno povezani sa ulaznim parametrima modela, preciznije u odnosu na odgovarajuće linearne modele.

U skladu sa postavljenim hipotezama i definisanim predmetom i ciljevima istraživanja, doktorska disertacija je podeljena na sedam poglavlja.

Prvo poglavlje disertacije je uvodni deo, sa kratkim opisom održivog razvoja i značaja poznavanja indikatora kvaliteta životne sredine. U okviru ovog poglavlja definisani su predmet, ciljevi i značaj modelovanja indikatora kvaliteta životne sredine. Uz obrazloženje teme opisana je i struktura disertacije.

Naredna tri poglavlja, Održivi razvoj i modelovanje indikatora održivog razvoja i kvaliteta životne sredine, Veštačke neuronske mreže (ANN) i Hemometrijske metode primenjene za odabir ulaznih promenljivih modela, čine Teorijski deo disertacije. U Teorijskom delu istaknut je značaj održivog razvoja i opisani su osnovni principi održivog razvoja, sa posebnim ostvrtom na indikatore održivog razvoja i kvaliteta životne sredine. Nakon toga je dat prikaz odabranih indikatora održivog razvoja i kvaliteta životne sredine, čije modelovanje je predmet ove disertacije: emisija gasova staklene bašte i suspendovanih čestica na nacionalnom nivou, koncentracija suspendovanih čestica u vazduhu na nacionalnom nivou, količina generisanog komunalnog čvrstog otpada na nacionalnom nivou i sadržaj rastvorenog kiseonika u površinskim vodama. Zatim je dat prikaz nekih klasičnih modela koji se koriste za predviđanje emisija i koncentracija zagađujućih materija, koji su predmet ove disertacije. Veštačke neuronske mreže, različite arhitekture neuronskih mreža, kao i obučavanje svake od korišćenih arhitektura takođe su detaljno opisani. Na kraju teorijskog dela su opisane i različite hemometrijske metode primenjene za odabir ulaznih promenljivih modela razvijenih tokom izrade ove doktorske disertacije.

U poglavlju Eksperimentalni deo opisana je metodologija razvoja modela zasnovanog na neuronskim mrežama, a sastoji se iz tri potpoglavlja: Proces razvoja modela zasnovanog na neuronskim mrežama, Pretraživanje relevantnih baza podataka, Statistički pokazatelji performansi modela. U eksperimentalnom delu detaljno je opisana metodologija razvoja modela zasnovanog na neuronskim mrežama, kao i statistički pokazatelji performansi korišćeni prilikom analize performansi modela razvijenih tokom izrade ove disertacije.

Rezultati i diskusija su prikazani u šestom poglavlju, Modelovanje indikatora kvaliteta životne sredine, koje se sastoji iz pet potpoglavlja: Modelovanje emisije gasova staklene bašte, Modelovanje emisije PM10 čestica, Modelovanje koncentracije PM10 čestica, Modelovanje količina generisanog komunalnog čvrstog otpada i Modelovanje sadržaja rastvorenog kiseonika u površinskim vodama. U svakom od navedenih potpoglavlja detaljno je opisana metodologija razvoja ANN modela: odabir ulaznih promenljivih i njihova priprema za modelovanje, optimizacija neuronske mreže i razvoj modela, primena hemometrijskih metoda za odabir ulaznih promenljivih, analiza značajnosti pojedinih ulaznih promenljivih i analiza performansi razvijenih modela.

Poslednje, sedmo poglavlje, sadrži kratki prikaz istraživanja, sa glavnim zaključcima koji su proistekli iz ove doktorske disertacije, kao i pravce budućih istraživanja.

Teorijski deo

2. Održivi razvoj i modelovanje indikatora održivog razvoja i kvaliteta životne sredine

2.1. Principi održivog razvoja

Održivi razvoj definisan je kao razvoj koji izlazi u susret potrebama sadašnjice, a da ne ugrožava sposobnost budućih generacija da zadovolje svoje sopstvene potrebe. Pojam održivog razvoja je definisan na sastanku Svetske komisije za životnu sredinu i razvoj (*World Commission on Environment and Development - WCED*) 1987. godine u izveštaju naslovljenom „Naša zajednička budućnost“ (*Our Common Future*) [1]. Kako bi se praksi ostvarivali, koncept i zacrtani ciljevi održivog razvoja značajno se proširuju u Nacionalnim strategijama održivog razvoja. U Nacionalnoj strategiji održivog razvoja Republike Srbije, održivi razvoj se definiše kao dugoročni koncept razvoja koji podrazumeva stalni ekonomski rast koji, osim ekonomske efikasnosti, tehnološkog napretka, više čistijih tehnologija, inovativnosti celog društva i društveno odgovornog poslovanja, obezbeđuje smanjenje siromaštva, dugoročno bolje korišćenje resursa, unapređenje zdravstvenih uslova i kvaliteta života i smanjenje zagađenja na nivo koji mogu da izdrže činioци životne sredine, sprečavanje novih zagađenja i očuvanje biodiverziteta [2].

U navedenoj Nacionalnoj strategiji održivog razvoja Republike Srbije [2], kao principi održivog razvoja navode se sledeći, opšte prihvaćeni, principi:

- Međugeneracijska solidarnost i solidarnost unutar generacije: zadovoljiti potrebe sadašnjih generacija, ali tako da se ne ugroze prava budućih generacija da zadovolje svoje potrebe. Solidarnost unutar generacije postići demokratski usaglašenom raspodelom raspoloživog prirodnog i stvorenog kapitala, tako da se obezbede osnovne ljudske potrebe za sve društvene grupe;
- Otvoreno i demokratsko društvo (učesće građana u odlučivanju): garantovati građanska prava, obezbediti pristup informacijama i osigurati dostupnost pravde. Obezbediti odgovarajuće konsultacije s građanima i učesće

- građana u donošenju odluka. Braniti stabilnost demokratskih institucija na osnovama mira, bezbednosti i slobode;
- Znanje kao nosilac razvoja: promovisati prosperitetnu, inovativnu, konkurentnu i ekološki efikasnu ekonomiju zasnovanu na znanju, koja obezbeđuje visok životni standard i punu i visokokvalitetnu zaposlenost. Promovisati obrazovanje i razvijanje javne svesti o održivom razvoju;
 - Uključenost u društvene procese: promovisati punu integraciju građana u društvo, podsticati jednake mogućnosti za svakoga tako što će se promovisati ljudska prava, posebno rodna ravnopravnost i boriti protiv svih oblika diskriminacije afirmativnim merama za marginalizovane grupe i smanjenjem siromaštva. Razlike i podeljenost među članovima društva treba svesti na najmanju moguću meru i stalno se boriti protiv socijalne isključenosti ljudi i siromaštva;
 - Integrisanje pitanja životne sredine u ostale sektorske politike: promovisati integraciju ekonomskih, socijalnih i ekoloških pristupa i analiza, te podržati korišćenje instrumenata kao što je strateška procena životne sredine. Podsticati socijalni dijalog, društveno odgovorno poslovanje i javno-privatno partnerstvo;
 - Predostrožnost: zahtevati očuvanje prirodne ravnoteže kada nema pouzdanih informacija o određenom problemu. Svaka aktivnost mora biti planirana i sprovedena tako da prouzrokuje najmanju moguću promenu u životnoj sredini. Preventivno delovati da bi se sprečili mogući znatni negativni uticaji na životnu sredinu, osobito ako bi bila ugrožena dobrobit ljudi i životinja;
 - Zagađivač/korisnik plaća (uključenje troškova vezanih za životnu sredinu u cenu proizvoda): internalizovati troškove vezane za životnu sredinu, tj. uključiti troškove uništavanja životne sredine u ekonomske troškove zagađivača/korisnika, tj. primeniti princip zagađivač/korisnik plaća. Tako se postiže puna ekonomska cena koja pokriva troškove proizvodnje, upotrebe i odlaganja proizvoda tokom čitavog njegovog „životnog“ ciklusa;
 - Održiva proizvodnja i potrošnja: poštovati uravnotežene odnose u eksploataciji prirodnih resursa i obezbediti visok nivo zaštite i poboljšanja kvaliteta

životne sredine. Smanjiti zagađenje životne sredine i promovisati održivu potrošnju i proizvodnju, ali ekonomski rast ne sme da uzrokuje proporcionalni porast degradacije životne sredine.

Kako u mnogim zemljama, razvijenim i zemljama u razvoju, nedostaju podaci o vrednostima različitih indikatora održivog razvoja, na osnovu kojih se prati ispunjavanje definisanih principa u oblasti zaštite životne sredine, postoji konstantna potreba za razvojem matematičkih i statističkih modela. Navedenim modelima vrše se predviđanja vrednosti indikatora održivog razvoja, što omogućava efikasno usklađivanje politike i strategija poboljšanja životne sredine sa budućim potrebama.

2.2. Indikatori održivog razvoja i kvaliteta životne sredine

Indikatorima se nazivaju veličine čije vrednosti opisuju određeni prirodni ili društveni fenomen. U opštem slučaju, indikator se dobija sumiranjem više različitih podataka u jednu vrednost, čime se pojednostavljuje praćenje složenih prirodnih procesa i pojava.

U oblasti zaštite životne sredine i održivog razvoja indikatori se koriste da bi se dobile neophodne informacije o stanju u životnoj sredini i o potrebnim merama za njegovo poboljšanje. Upotreba indikatora održivog razvoja i kvaliteta životne sredine se ogleda kroz:

- dobijanje podataka o ekološkim problemima, kako bi se na adekvatan način odredila njihova težina i značaj,
- pomoć pri određivanju glavnih faktora koji doprinose narušavanju kvaliteta životne sredine,
- praćenje efekata i rezultata ekoloških mera koje se sprovode,
- informisanje stanovništva o ekološkim problemima čije je rešavanje od javnog interesa.

Mnoge države, međunarodne organizacije, instituti i ostale relevantne službe i agencije koriste indikatore održivog razvoja i kvaliteta životne sredine za izveštavanje o stanju u životnoj sredini. Neke od međunarodnih organizacija koje su najaktivnije na prostoru Evrope su: Ujedinjene nacije (UN), Organizacija za ekonomsku saradnju i razvoj (*Organization for Economic Co-operation and Development* - OECD), Svetska banka (*The World Bank* - WB), Evropska agencija za životnu sredinu (*The European Environment Agency* - EEA).

Prateći EEA metodologiju [3], indikatori kvaliteta životne sredine se mogu podeliti na:

- Indikatore stanja (tip A) – opisuju promenu određene promenljive tokom posmatranog perioda (npr. sadržaj rastvorenog kiseonika u vodama ili udeo farmi sa organskom proizvodnjom u ukupnoj poljoprivrednoj proizvodnji, itd.);
- Indikatore performansi (tip B) – predstavljaju indikatore stanja koji su vezani za neku ciljanu vrednost, pri čemu praktično pokazuju razliku između trenutnog i poželjnog stanja u životnoj sredini. Ciljane vrednosti se obično određuju na međunarodnom nivou, pa se ovaj tip indikatora najčešće koristi od strane međunarodnih institucija za procenu stanja u životnoj sredini na nacionalnom ili regionalnom nivou;
- Indikatori efikasnosti (tip C) – pokazuju pritiske na životnu sredinu, kroz informacije o efikasnosti proizvodnih procesa kod kojih dolazi do emisije zagađujućih materija ili generisanja otpada. Oni se najčešće koriste kod donošenja strategija za poboljšanje stanja u životnoj sredini, jer na direktan način ukazuju na ljudske aktivnosti koje narušavaju kvalitet životne sredine;
- Indikatore delotvornosti ekoloških mera (tip D) – prati se poboljšanje kvaliteta životne sredine uzrokovana ekološkim merama koje se sprovode.

U eksperimentalnom delu doktorske teze korišćeni su indikatori održivog razvoja koje Eurostat, statistička agencija Evropske unije, primenjuje da bi se pratilo ispunjavanje ciljeva Evropske strategije održivog razvoja. Eurostatova baza podataka održivog razvoja sadrži preko 100 indikatora koji su podeljeni u 10 grupa [4]:

- Socio-ekonomski razvoj – glavni indikator je „rast bruto domaćeg proizvoda (BDP) po stanovniku“,
- Održiva proizvodnja i potrošnja - glavni indikator je „produktivnost resursa“,
- Socijalna inkluzija - glavni indikator je „rizik od siromaštva“,
- Demografske promene - glavni indikator je „stopa zaposlenosti starijih radnika“,
- Javno zdravlje - glavni indikator je „očekivani srednji ljudski vek“,
- Klimatske promene i energija – glavni indikatori su „emisija gasova staklene bašte“, „udeo obnovljivih izvora u finalnoj potrošnji energije“ i „primarna potrošnja energije“,
- Održivi transport - glavni indikator je „potrošnja energije u transportu u odnosu na BDP“,
- Prirodni resursi - glavni indikatori su „broj ptica“ i „ulov riba“,
- Globalno partnerstvo i
- Odgovorna uprava (vlast).

Indikatori održivog razvoja i kvaliteta životne sredine, koji su tema ove doktorske teze, detaljno su prikazani u sledećem poglavlju.

2.3. Prikaz odabranih indikatora održivog razvoja i kvaliteta životne sredine

2.3.1. Emisija gasova staklene bašte (GHG) na nacionalnom nivou

Emisija gasova staklene bašte (*Greenhouse gas* - GHG) je u centru pažnje već duži niz godina zbog doprinosa globalnom zagrevanju, kao i uticaja na klimatske promene koje su jedan od najvećih problema savremene civilizacije [5]. Da bi se u potpunosti razumeo uticaj čoveka na globalno zagrevanje, i klimatske promene uopšte, neophodno je obezbediti pouzdane podatke o emisijama gasova staklene bašte u atmosferu, a koje su posledica ljudskih aktivnosti [6]. Analizom privrednih aktivnosti, kod kojih se emituju gasovi staklene bašte, može se zaključiti da su sledeći sektori glavni emiteri: energetska sektor (60 % od ukupne emisije na tlu Evrope), saobraćaj (20 %), poljoprivreda (10 %), industrija (7 %) i upravljanje otpadom 3% [7].

The United Nation Framework Convention on Climate Change (UNFCCC) [8] i Kjoto protokol predstavljaju prve aktivnosti na globalnom nivou UN-a i država članica, čiji je cilj bio da se emisije GHG smanje na prihvatljivu meru. Kjoto protokol je stupio na snagu 2005. godine, a jedan od glavnih ciljeva na koji su se zemlje potpisnice obavezale je smanjenje GHG emisija za 5,2 % u periodu od 2008-2012. godine, u poređenju sa emisijama iz 1990. godine. Da bi se ostvarili principi održivog razvoja, takođe je potrebno značajno smanjenje GHG emisija na globalnom nivou, što iziskuje saradnju velikog broja, a praktično svih zemalja sveta.

Indikator održivog razvoja kojim se prate GHG emisije na nacionalnom nivou definisan je sa ciljem kvantifikovanja promena u GHG emisijama i ispunjavanja preuzetih obaveza od strane država. Takođe, pored modela kojim se mogu predviđati vrednosti indikatora, bilo je potrebno razviti i modele kojima bi mogla da se vrši simulacija različitih scenarija GHG emisija i time njihovo efikasnije smanjivanje. Iz navedenog razloga, modelovanje GHG emisije u različitim sektorima njihove emisije bila je tema velikog broja radova, pri čemu su korišćeni različiti pristupi i tehnike [9-18]. Jedan od najznačajnijih model koji se koristi za predviđanje GHG emisija je *Greenhouse Gas and Air Pollution Interactions and Synergies* (GAINS) model [19].

Pored linearnih modela, veštačke neuronske mreže (*Artificial neural networks* - ANNs) se sve češće primenjuju za predviđanje GHG emisija, npr. predviđanje emisije na osnovu potrošnje energije [20] ili na osnovu različitih ekonomskih indikatora [21, 22].

2.3.2. Emisija PM₁₀ čestica na nacionalnom nivou

Čestično zagađenje atmosfere javlja se usled emisije čestica (*Particulate matter* - PM) tokom proizvodnje energije, industrijskih procesa i saobraćaja. Čestice u vazduhu, pogotovo u urbanim sredinama, značajno narušavaju kvalitet vazduha i time ugrožavaju zdravlje velikog broja ljudi u svetu. Mnoge epidemiološke studije su pokazale povezanost između čestičnog zagađenja i disajnih, odnosno srčanih smetnji [23]. Uticaj na zdravlje pre svega zavisi od prečnika samih čestica, kao i od količine i vremena izlaganja zagađenju. Postoje dva tipa uticaja na zdravlje ljudi koji se u literaturi prate [24]:

- dugotrajni efekti, koji su posledica kumulativnog izlaganja i povezani su sa hroničnim posledicama po zdravlje,
- kratkotrajni efekti, koji su posledica izlaganja povišenim koncentracijama čestica na dnevnoj bazi i kao takvi imaju uticaj na akutno stanje respiratornog i kardiovaskularnog sistema.

Iz navedenih razloga, čestice u atmosferi se smatraju kao prioritetni polutant u Evropskoj uniji [25], pri čemu se teži poboljšanju kvaliteta vazduha kako bi se štetni efekti na zdravlje ljudi sveli na minimum.

Kako bi se Evropske direktive o kvalitetu vazduha (2001/81/EC – *NEC directive* [26] i 2008/50/EC – *CAFE (Cleaner air for Europe) directive* [27]) na efikasan način sprovele, neophodno je poznavanje emitovanih količina PM₁₀ čestica, kao i doprinosa pojedinačnih sektora njihovoj emisiji. Takođe, protokolima Ženevske konvencije (*Convention on Long Range Transboundary Air Pollution* - CLRTAP) definisana je obaveza izveštavanja na nacionalnom nivou o emitovanim količinama PM₁₀ čestica. EU direktive, kao i Ženevska konvencija, nalažu i predviđanje emisija zaga-

đujućih materija više godina unapred, pa je upotreba odgovarajućih modela u tu svrhu neophodna.

Od strane Evropske agencije za zaštitu životne sredine (*European Environment Agency* – EEA), RAINS (*Regional Air Pollution Information and Simulation*) model uglavnom je bio korišćen za predviđanje emisije PM₁₀ čestica u atmosferu. Kako se PM₁₀ čestice emituju iz velikog broja različitih izvora, da bi RAINS model mogao biti primenjen, bilo je potrebno razvrstati izvore emisije u 392 kategorije, pri čemu je ukupan broj potrebnih ulaznih promenljivih za pojedine države bio čak 800 [28].

Pored linearnih modela i ANN modeli se koriste za predikciju PM₁₀ emisije. *Kassomenos* sa saradnicima [29] je razvio ANN model za određivanje emisije PM₁₀ čestica iz saobraćaja. Pokazalo se je da je glavna razlika između linearnih (npr. RAINS) i ANN modela u broju ulaznih promenljivih koji su neophodni za predviđanje PM₁₀ emisije. Kako je za ANN modele potreban znatno manji broj ulaznih promenljivih, oni mogu biti primenjeni kao alternativni, u situacijama kada ne postoje svi potrebni podaci za određivanje PM₁₀ emisije u vazduh linearnim modelima, bilo na nacionalnom ili na regionalnom nivou.

2.3.3. Koncentracija PM₁₀ čestica na nacionalnom nivou

Kao što je već navedeno u prethodnom poglavlju, čestično zagađenje vazduha negativno utiče na ljudsko zdravlje, pa se zbog toga čestice smatraju jednom od glavnih zagađujućih materija vazduha, pogotovo u urbanim sredinama. Da bi se smanjile koncentracije čestica u vazduhu doneti su zakoni kojima su definisane maksimalno dozvoljene koncentracije u mnogim zemljama sveta [30]. Zakon o čistom vazduhu (*Clean air act*) u SAD-u i Direktiva 2008/50/EC (*CAFE directive*) [27] Evropske komisije dovele su do smanjivanja koncentracija čestica u vazduhu u urbanim sredinama na prostoru Severne Amerike i Evrope. Međutim, nove epidemiološke studije sprovedene u poslednjoj deceniji, sa poboljšanim metodama analize i detekcije, pokazale su da se zdravstvene smetnje javljaju i pri smanjenim koncentracijama čestica, nastalih uglavnom sagorevanjem goriva, što je dominantan izvor čestičnog zagađenja u gradovima Severne Amerike i Zapadne Evrope [31].

Kvantifikovanje uticaja navedenog zagađenja na stanovništvo u gradovima širom sveta veoma je zahtevno, s obzirom da nema dovoljno informacija o stepenu izloženosti stanovništva u mnogim delovima sveta [32]. Smanjivanjem javnih fondova usled ekonomske krize, dodatno su smanjena sredstva koja su korišćena za monitoring kvaliteta vazduha u razvijenim i zemljama u razvoju. Zbog toga se i u mnogim razvijenim zemljama ne zna sa sigurnošću kolika je izloženost stanovništva čestičnom zagađenju [33]. Kako se u mnogim zemljama upšte ne sprovodi kontinuirani monitoring kvaliteta vazduha, čestično zagađenje se procenjuje korišćenjem različitih statističkih modela [32].

Da bi se poredila izloženost stanovništva čestičnom zagađenju u različitim državama i gradovima sveta, Svetska zdravstvena organizacija (*World Health Organization* - WHO) koristi indikatore kojima se prate srednje godišnje koncentracije PM₁₀ i PM_{2,5} čestica na nacionalnom nivou i u gradovima. Svetska Banka (*World Bank* - WB) takođe koristi vrednosti PM₁₀ koncentracija na nacionalnom nivou kao indikator održivog razvoja u sferi zaštite životne sredine. WHO i WB koriste vrednosti koncentracija PM₁₀ čestica na nacionalnom nivou koje su dobijene GMAPS (*The Global Model of Ambient Particulates*) modelom [34].

Poslednjih godina, ANN se sve više koriste za predviđanje koncentracija čestica u vazduhu u urbanim sredinama. Na primer, *Perez i Reyes* [35] su primenili ANN modele za predikciju maksimalnih vrednosti koncentracija PM₁₀ čestica u Santijagu, dok su *Kukkonen* i saradnici [36] ANN modelima vršili predviđanje vrednosti PM₁₀ koncentracija u Helsinkiju. *Hooyberghs* i saradnici [37] su vršili predviđanje srednjih dnevnih PM₁₀ koncentracija u Belgiji, dok su *Voukantsis* i saradnici [38] kombinovali analizu glavnih komponenti (*principal component analysis* - PCA) i neuronske mreže za predviđanje srednjih dnevnih PM₁₀ i PM_{2,5} koncentracija u urbanim sredinama Grčke i Finske. *Paschalidou* i saradnici [39] su razvili ANN modele za predviđanje časovnih PM₁₀ koncentracija u urbanim sredinama na Kipru.

2.3.4. Količina generisanog komunalnog čvrstog otpada

Upravljanje komunalnim čvrstim otpadom (KČO) prepoznato je kao problem u razvijenim zemljama tokom poslednje decenije XX veka, kada su uočene sve negativne posledice lošeg upravljanja komunalnim čvrstim otpadom po kvalitet života i zdravlje ljudi [40]. Podaci o količinama i sastavu otpada koji se generiše na određenom području su neophodni da bi se vršilo planiranje i optimizacija procesa u sistemu upravljanja otpadom [41]. Neadekvatno upravljanje KČO-om može dovesti do ozbiljnih komunalnih, sanitarnih i ekoloških problema, poput pojave neprijatnih mirisa i rizika od eksplozija u zonama divljih deponija ili kontaminacije podzemnih voda usled izluživanja zagađujućih materija iz neadekvatno odloženog otpada [42].

Upravljanje KČO-om tema je brojnih naučnih radova [43-48], pri čemu se u mnogim ističe nedostatak preciznih podataka u pojedinim evropskim zemljama, koji su neophodni kako bi se u njima upravljanje otpadom adekvatno uredilo [49, 50].

U Srbiji se divlje deponije mogu naći u ruralnim sredinama gotovo u svim opštinama, a posledica su nedovoljnih sredstava koja se izdvajaju za poboljšanje sistema prikupljanja otpada, kao i loše organizacije samog sistema za sakupljanje otpada na lokalnom nivou [51]. Zbog dugogodišnjeg neadekvatnog upravljanja KČO-om na teritoriji Republike Srbije je 2009. godine bilo oko 4400 divljih deponija, i 164 uređene deponije koje postoje na teritoriji Srbije [52].

Potreba za adekvatnim matematičkim modelima koji bi pomogli u planiranju i boljoj organizaciji sistema upravljanja KČO-om je evidentna [53, 54]. Primena modela koji obezbeđuju podatke o generisanim količinama komunalnog čvrstog otpada predstavlja značajno unapređenje same prakse u upravljanju KČO-om [55].

Neuronske mreže se koriste i u ovoj oblasti, pa su tako objavljeni radovi sa njihovom uspešnom primenom za predviđanje: nedeljnih količina generisanog KČO [55], generisanih količina medicinskog otpada [56], izluživanja zagađujućih materija iz solidifikovanog otpada [57].

2.3.5. Sadržaj rastvorenog kiseonika u površinskim vodama

Očuvanje i poboljšanje kvaliteta površinskih voda jedan je od glavnih ekoloških problema u svetu i u Srbiji, s obzirom da njihov kvalitet značajno utiče na zdravlje stanovništva i celokupnog ekosistema. Kvalitet površinskih voda zavisi od prirodnih procesa, poput litologije sliva, atmosferskih padavina, klimatskih uslova, ali i od antropogenih uticaja [58]. Monitoringom kvaliteta površinskih voda određuje se veliki broj fizičko-hemijskih i bioloških parametara. Glavni problem kod monitoringa voda je kompleksnost analiziranja kvaliteta voda, sa velikim brojem promenljivih, koji ga određuju [59]. Kako se kontinuirani monitoring sprovodi na manjem broju većih površinskih voda, i to sa uzorkovanjima na mesečnom ili polumesečnom nivou, modelovanje parametara kvaliteta voda je prepoznato kao praktično veoma značajan alat za dobijanje informacija potrebnih za optimalno upravljanje površinskim vodama [60].

Rastvoreni kiseonik (*dissolved oxygen* - DO) predstavlja sadržaj kiseonika u površinskim vodama. Sadržaj rastvorenog kiseonika se kontinuirano prati u okviru monitoringa kvaliteta voda, jer je on jedan od ključnih parametara kvaliteta voda [61]. Naime, određena količina kiseonika je neophodna za opstanak akvatičnog sveta, pa je tako sadržaj rastvorenog kiseonika mera „zdravstvenog stanja“ akvatičnog ekosistema [62]. Veliki broj faktora, poput temperature, protoka, sedimentacije, reaeracije, nitrifikacije, fotosinteze, utiče na sadržaj rastvorenog kiseonika, koji nelinearno zavisi od vrednosti ostalih parametara kvaliteta površinskih voda.

Primenom ANN efikasno se određuju nelinearne veze između ulaznih i izlaznih promenljivih, pa se u literaturi može naći veliki broj naučnih radova u kojima je predviđanje sadržaja rastvorenog kiseonika bilo vršeno neuronskim mrežama [58, 59, 61-65].

2.4. Prikaz odabranih modela koji se koriste za predviđanje emisija i koncentracija zagađujućih materija

2.4.1. Inventari emisija

Inventari emisija su glavni izvori podataka o emisijama zagađujućih materija u atmosferu i predstavljaju baze podataka koje se koriste kao ulaz kod procene emisija po sektorima. U opštem slučaju, emisije (E_p) se procenjuju množenjem faktora emisije (*emission factor - Ef*) sa odgovarajućom stopom aktivnosti (*activity rate - Ar*) posmatranog zagađivača (*Tier 1 default approach*):

$$E_p = \sum_{aktivnosti} Ar \times Ef \quad (1)$$

U praksi, nijedan od ovih ulaznih parametara (Ef i Ar) nije egzaktno poznat, već se njihove vrednosti određuju tako da budu najbolje procene stvarnih vrednosti [6].

Stopa aktivnosti se najčešće izvodi iz različitih statističkih baza podataka vezanih za ekonomiju i proizvodnju, energetske sektor, itd. Emisioni faktori, s druge strane, moraju biti određeni za svaki pojedinačni sektor, bilo merenjem ili proračunom (masenim bilansom). Čak i u slučaju određivanja emisionih faktora velikim brojem merenja, u velikom broju postrojenja, koja u potpunosti reprezentuju određeni sektor emisije, smatra se da oni imaju nesigurnost od 10 do 30 % [66].

Najčešće korišćeni inventari emisija su:

- EMEP/EEA *guidebook* (ranije poznat kao EMEP CORINAIR) koji daje smernice za proračune emisije, kako iz antropogenih tako i iz prirodnih izvora emisije. Navedeni EMEP/EEA inventar obuhvata izveštaje o emisijama koje države dostavljaju u okviru UNECE programa (*Convention on Long-range Transboundary Air Pollution - CLRTAP*) i EU *National Emission Ceilings Directive* [66];
- AP42 je primarni US EPA inventar emisije od 1972. godine. U njemu se nalaze emisioni faktori za više od 200 različitih kategorija zagađivača [67].

U slučaju da u EMEP/CLRTAP izveštajima nedostaju podaci o emisijama određenih zagađujućih materija, tada su procene emisija vršene RAINS modelom (*Regional Air Pollution Information and Simulation*), odnosno GAINS modelom (*Greenhouse Gas and Air Pollution Interactions and Synergies*) [68]. Detaljne analize inventara emisije, i nesigurnosti podataka koje oni sadrže, prikazane su u odgovarajućoj literaturi [6, 69-71].

2.4.2. RAINS model

The Regional Air Pollution Information and Simulation (RAINS) model je razvijen od strane Međunarodnog instituta za primenjenu sistemsku analizu (*International Institute for Applied Systems Analysis - IIASA*) [72]. RAINS metodologija kombinuje podatke o stepenu ekonomskog razvoja, energetske sektoru, kontroli emisije i njenim troškovima i atmosferskoj disperziji, kako bi se došlo do procene atmosferskog zagađenja [73]. Modelom se vrše predviđanja zagađenja koje utiče na zdravlje ljudi, ali takođe se procenjuju i opasnosti po ekosistem predviđanjima acidifikacije, eutrofikacije, koncentracije ozona, itd. Efekti navedenih ekoloških problema procenjuju se predviđanjem emisija više zagađujućih materija: gasova staklene bašte, sumpor-dioksida, azotnih oksida, amonijaka, nemetanskih isparljivih organskih jedinjenja (*non-methane volatile organic compounds - NMVOC*) i čestica (PM_{2,5} i PM₁₀) [74]. Detaljni opis RAINS modela prikazan je u odgovarajućoj literaturi [73, 75].

Analiza nesigurnosti (*uncertainty analysis*) emisija dobijenih RAINS modelom pokazala je da se tačnost predviđanja razlikuje od države do države [76]. Schöpp i saradnici [77] su zaključili da nekoliko faktora utiče na pojavu varijacije nesigurnosti kod RAINS modela:

- ukoliko se emisija određene zagađujuće materije dominantno vrši u jednom ili dva sektora, eventualna nesigurnost u podacima iz dominantnih sektora ne može biti kompenzovana kvalitetnim podacima iz ostalih sektora;
- kako je RAINS model baziran na pristupu koji se koristi i kod inventara emisija (procene se vrše na osnovu emisionih faktora i stope aktivnosti), nesi-

gurnost emisionih faktora dominantnih izvora emisije u velikoj meri utiče na tačnost predviđenih emisija;

- kod nekih država se velika nesigurnost kod predviđanja emisija ne može izbeći, npr. podaci država centralne i istočne Evrope su nesigurniji u odnosu na podatke EU15 zemalja.

Od 2007. godine RAINS model je inkorporiran u *Greenhouse Gas and Air Pollution Interactions and Synergies* (GAINS) model [19].

2.4.3. GAINS model

GAINS model predstavlja jedan od najznačajnijih modela koji se koriste za predviđanje emisija zagađujućih materija, i takođe je razvijen od strane austrijskog IIASA instituta [72].

GAINS model se koristi za procene trenutnih i budućih emisija zagađujućih materija na osnovu: emisionih faktora, stope aktivnosti, efikasnosti i primene mera za sprečavanje emisije. Predviđanje se vrši na osnovu jednačine [78]:

$$E_{i,p} = \sum_k \sum_m A_{i,k} Ef_{i,k,m,p} x_{i,k,m,p} \quad (2)$$

gde i , k , m , p označavaju države, tip aktivnosti, meru za smanjenje emisije i zagađujuću materiju, redom, $E_{i,p}$ - emisiju određenog polutanta na nacionalnom nivou, $A_{i,k}$ - stopa aktivnosti, $Ef_{i,k,m,p}$ -emisioni faktor određenog polutanta p pri aktivnosti k u državi i uz primenu mere za kontrolu emisije m , $x_{i,k,m,p}$ -udeo primenjenih mera za kontrolu emisije u odnosu na ukupnu emisiju.

Poređenjem jednačina (1) i (2) može se uočiti da GAINS metodologija predstavlja unapređenje inventarskog pristupa u proceni emisija zagađujućih materija. Više detalja o GAINS modelu se može naći na sajtu IIASA instituta [19].

2.4.4. GMAPS model

The Global Model of Ambient Particulates (GMAPS) je model kojim se, na bazi ekonomskih parametara, procenjuje koncentracija čestica (*particulate matter* – PM) u urbanim sredinama, koristeći podatke Svetske zdravstvene organizacije (*World Health Organization* – WHO) i ostalih relevantnih izvora. Jednačina (3) kojom se vrši procena PM koncentracije fokusirana je na prirodne i antropogene izvore zagađujućih materija, kao i na kapacitet životne sredine da uneto zagađenje razblaži i neutrališe. Kao ulazne veličine koriste se podaci o broju stanovnika u državama i gradovima, nacionalna i lokalna gustina naseljenosti, lokalni nivo ekonomske aktivnosti, bruto domaći proizvod po stanovniku, šest različitih indikatora potrošnje goriva/energije, 22 atmosferske i geografske promenljive, kao i binarne promenljive, kako bi se razlikovale koncentracije PM₁₀ čestica od suspendovanih čestica (*total suspended particles* - TSP) [32].

$$\begin{aligned}
 C_{ijkt} = & \sum_{k=1}^K \beta_k Z_k + \sum_{f=1}^F \beta_{E_f} E_{fkt} + \sum_{g=1}^{G2} \beta_{M_g} M_{gik} + \beta_R R_{kt} + \beta_N N_{jkt} + \beta_D D_{jkt} + \beta_{Scale} Scale_{jkt} \\
 & + \beta_Y Y_{kt} + \beta_T Trend_{ijkt} + \beta_{YT} Y_{kt} Trend_{ijkt} + \theta_S S_{ijkt} + \theta_{Scale} S_{ijkt} Scale_{jkt} + \theta_Y S_{ijkt} Y_{kt} \\
 & + \theta_T S_{ijkt} Trend_{ijkt} + \theta_{YT} S_{ijkt} Y_{kt} Trend_{ijkt} + \sum_{g=1}^{G1} \theta_{M_g} M_{gik} S_{ijkt} + \epsilon_{ijkt}
 \end{aligned} \quad (3)$$

gde i, j, k, t, f, g označavaju monitoring stanicu, grad, državu, vreme, tip energetskog izvora, meteorološki/geografski faktor, redom; C – log PM koncentracije; E - log potrošnje energije po stan.; M - log meteorološkog/geografskog faktora; R - log gustine naseljenosti; N - log populacije; D - log gustine naseljenosti u urbanoj sredini; $Scale$ - log nivoa ekonomske aktivnosti; Y - log prihoda po stan.; $Trend$ – vremenski trend (npr. 1985=1, 1986=2, itd.); S – binarna promenljiva koja označava tip PM čestica; β_s i θ_s su parametri koji se procenjuju na osnovu samog modela.

Poređenje izmerenih i GMAPS-om predviđenih PM koncentracija, pokazalo je da linearna korelacija iznosi 0,9 i da se se jednačinom (3) može objasniti 88 % varijacija, što pokazuje da GMAPS model sa zadovoljavajućom preciznošću može da predvidi koncentraciju čestica u atmosferi [33].

Trenutna verzija GMAPS modela može da vrši predviđanja PM koncentracija kako u rezindencijalnim (urbanim) oblastima, tako i u nenaseljenim područjima. Rezultati GMAPS modela se koriste da bi se odredila srednja godišnja ambijentalna PM koncentracija u stambenim i nenaseljenim područjima u 3.226 gradova sa populacijom većom od 100.000 stanovnika, kao i u glavnim gradovima, bez obzira na njihovu veličinu [79]. Podaci o nacionalnim i regionalnim PM koncentracijama se dobijaju na osnovu podataka o koncentracijama u urbanim područjima sa populacijom većom od 100.000 stanovnika. Vrednost PM₁₀ koncentracije na nacionalnom nivou pokazuje prosečnu godišnju izloženost čestičnom zagađenju prosečnog stanovnika u urbanim sredinama [34].

2.4.5. Modeli zasnovani za višestrukoj linearnoj regresiji (MLR)

Modeli zasnovani na višestrukoj linearnoj regresiji (*multiple linear regression* – MLR) najčešće su primenjivani linearni modeli u oblasti zaštite životne sredine i održivog razvoja. Navedeni MLR modeli bazirani su na određivanju linearne zavisnosti izlazne promenljive (Y) od seta ulaznih promenljivih (X_i), prema jednačini:

$$Y = a_0 + a_1X_1 + \dots + a_nX_n \quad (4)$$

gde su a_i ($i=1\dots n$) koeficijenti linearne zavisnosti, koji se najčešće određuju metodom najmanjih kvadrata [80].

Glavni nedostatak MLR modela je činjenica da na rezultate modela znatno utiče eventualna međusobna korelacija nezavisnih promenljivih. Višestruko korelisane ulazne promenljive onemogućavaju identifikovanje najznačajnijih ulaznih promenljivih, a time i adekvatno predviđanje budućih stanja fizičkog procesa (promenljive) koji se modeluje [81].

Problem međusobne povezanosti ulaznih promenljivih može biti eliminisan različitim statističkim tehnikama, npr. *variance inflation factors* (VIF) metodom [82, 83] ili analizom glavnih komponenti (*principal component analysis* - PCA) [84].

2.4.6. Modeli zasnovani na regresiji glavnih komponenti (PCR)

Modeli zasnovani na regresiji glavnih komponenti (*principal component regression* – PCR) predstavljaju kombinaciju dve hemometrijske metode: analize glavnih komponenti (*principal component analysis* – PCA) i višestruke linearne regresije (MLR). Kao što je već navedeno u Poglavlju 2.4.5., uloga PCA je da detektuje i eliminiše međusobno korelisane ulaze MLR modela.

Razvoj PCR modela može se podeliti u dva uzastopna koraka:

- 1) najpre se koristi PCA da bi se originalni ulazi transformisali u međusobno nekorelisane glavne komponente, a zatim
- 2) se primenjuje MLR na vrednosti transformisanih ulaza, pri čemu se dolazi do regresione jednačine.

3. Veštačke neuronske mreže (ANN)

3.1. Osnove veštačkih neuronskih mreža

Neuronske mreže (u širem smislu) mogu se podeliti na biološke neuronske mreže (nervni sistemi živih bića) i na veštačke neuronske mreže (*Artificial neural networks* - ANNs). Termin neuronska mreža, u užem smislu, odnosi se samo na veštačke neuronske mreže i kao takav biće korišćen u daljem tekstu.

Neuronske mreže predstavljaju informaciono-procesne algoritme, kreirane po ugledu na biološke nervne sisteme, koji mogu da uopštavaju (generalizuju) na osnovu učenja (prošlog iskustva), i predviđaju stanja kompleksnih i nelinearnih sistema [57].

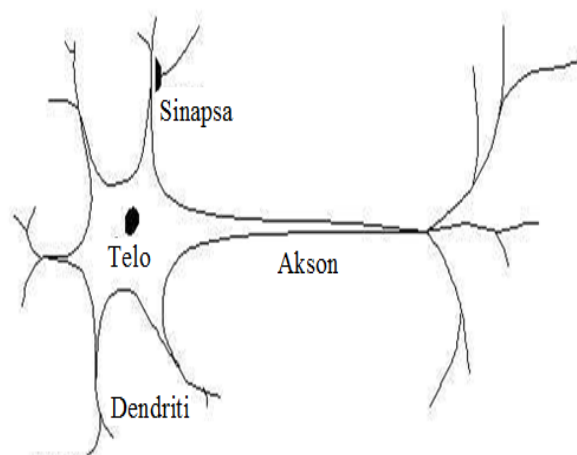
Na Slici 1. prikazani su osnovni delovi biološkog neurona i veštački neuron, osnovna jedinica neuronske mreže, koji se sastoji iz tri elementa:

- težinski koeficijenti (w_i),
- sabirač ulaznih signala (Σ) (ulazni signal je proizvod vrednosti ulazne veličine i težinskog koeficijenta),
- aktivaciona funkcija (F).

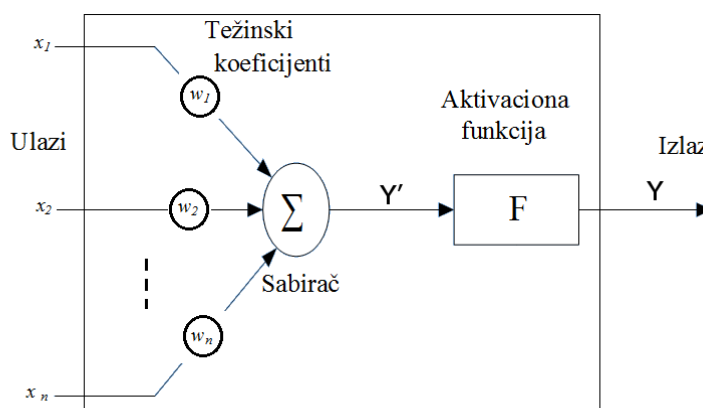
Vrednost izlaznog signala (Y') iz sabirača ulaznih signala može se predstaviti formulom:

$$Y' = x_1 w_1 + x_2 w_2 + \dots + x_n w_n \quad (5)$$

Izlazni signal Y' najčešće se obrađuje aktivacionom funkcijom (F), da bi se dobio izlazni signal neurona (Y). Aktivaciona funkcija je analogna pojačanju kod elektronskih, odnosno regulacionih sistema. To može biti i jednostavna linearna funkcija, odskočna ili neka druga funkcija (Slika 2.).



a)



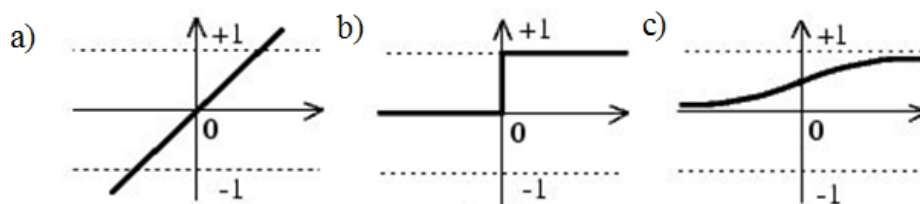
b)

Slika 1. Šematski prikaz a) biološkog i b) veštačkog neurona sa glavnim delovima

Najčešće se primenjuje sigmoidna funkcija, koja daje dobre rezultate i za ulazne signale niskog i za ulazne signale visokog intenziteta i ima oblik:

$$Y = \frac{1}{1 + e^{-Y'}} \quad (6)$$

Uz razne varijante sigmoidne funkcije, povremeno se koriste i hiperbolične funkcije, sinusna i Gausova funkcija.



Slika 2. Osnovne aktivacione funkcije: a) linearna b) odskočna c) sigmoidna

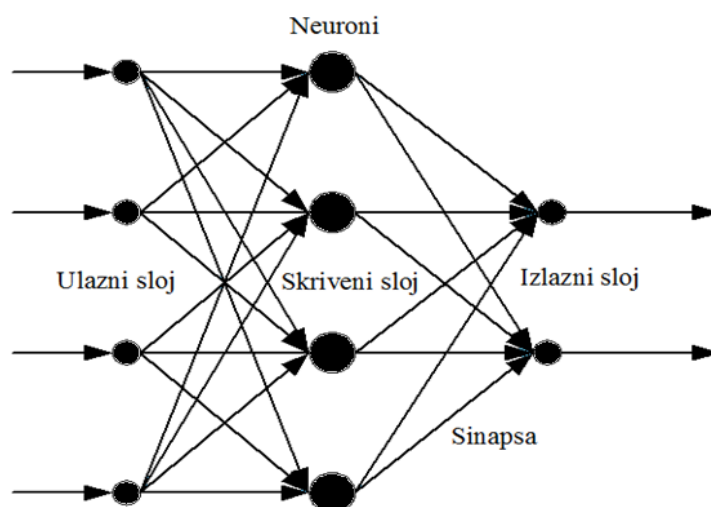
Povezivanjem nezavisnih neurona u slojeve dobija se neuronska mreža, koja može da procesira ulazne veličine, dajući izlaz u vidu klasifikacije ili prognoze.

Kao što je već navedeno, neuronska mreža može da se sastoji od najmanje dva sloja neurona (perceptron), mada se najčešće koriste neuronske mreže sa tri sloja, pošto su mnogi autori pokazali da je jedan skriveni sloj dovoljan da ANN na adekvatan način aproksimira nelinearne funkcije [85].

Sloj u kome se nalaze neuroni koji primaju ulazne podatke naziva se ulazni sloj, dok se sloj u kome se nalaze neuroni koji daju rezultate naziva izlazni sloj. Između njih se može nalaziti jedan ili više unutrašnjih (skrivenih) slojeva (Slika 3.). Neuroni u okviru jednog sloja obično nisu povezani između sebe, već samo sa neuronima ostalih slojeva.

Prevažodne razlike između veštačkih i bioloških neuronskih mreža se ogledaju u:

- različitoj brzini obrade podataka (ANN su i do 10 miliona puta brže),
- različitom broju neurona (ANN imaju do nekoliko stotina neurona u skrivenom sloju za razliku od bioloških neuronskih mreža koje sačinjava nekoliko milijardi neurona), i
- načinu rada (kod bioloških neuronskih mreža neuroni potpuno paralelno obrađuju podatke).



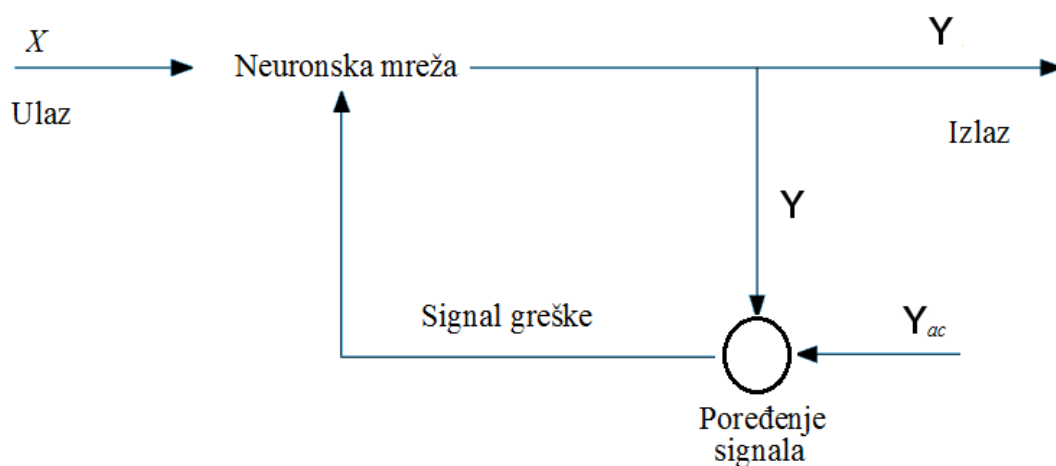
Slika 3. Struktura neuronske mreže na primeru troslojne mreže

U odnosu na tok signala, neuronske mreže mogu se podeliti na:

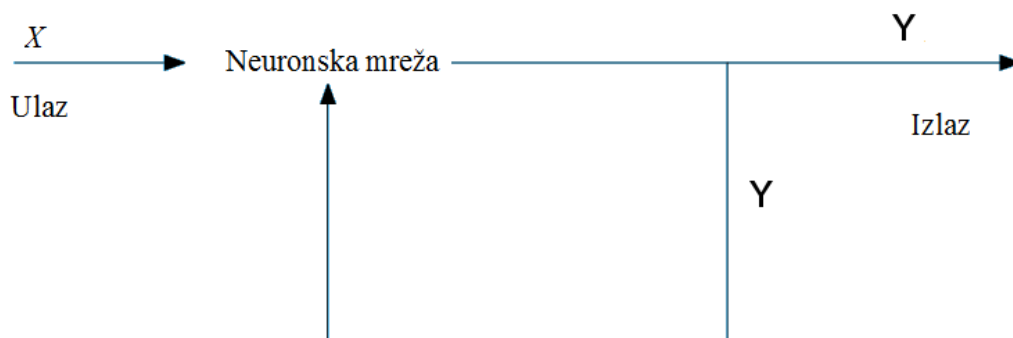
- mreže sa kretanjem signala napred, od ulaznog ka izlaznom sloju (*Feed-forward ANNs*) - signal se prenosi od neurona iz ulaznog sloja ka neuronu u skrivenom sloju, a neuron iz skrivenog sloja signal dalje prenosi ka izlaznom sloju neurona (Slika 3.);
- rekurentne mreže (*Feed-back ANNs*) - kod ovih mreža signal može da se kreće u oba smera, između slojeva neurona.

Prema načinu obučavanja, neuronske mreže se dele na:

- mreže sa nadgledanim obučavanjem (*supervised learning*) - vrši se poređenje vrednosti koje daje mreža (Y) sa tačnom (izmerenom) vrednošću izlazne veličine (Y_{ac}), pri čemu se generiše „signal greške“, na osnovu koga se u sledećoj iteraciji menjaju težinski koeficijenti, a time se menja i vrednost izlaznog signala (obučavanje ovog tipa traje sve dok se ne dobije zadovoljavajuće slaganje vrednosti Y i Y_{ac}) (Slika 4.);
- mreže sa podsticajnim obučavanjem (*reinforcement learning*) - analogni tip nadgledanog treningu, s tom razlikom što „signal greške“ nema neku određenu numeričku vrednost, već je opisnog tipa (npr. dobar, loš);
- mreže sa nenadgledanim obučavanjem, samoobučavanjem (*unsupervised learning*) - kod ovog tipa treninga nema korekcije izlaznog signala mreže kroz poređenje sa tačnom (izmerenom) vrednošću izlazne veličine (Slika 5.).



Slika 4. Nadgledano obučavanje neuronske mreže

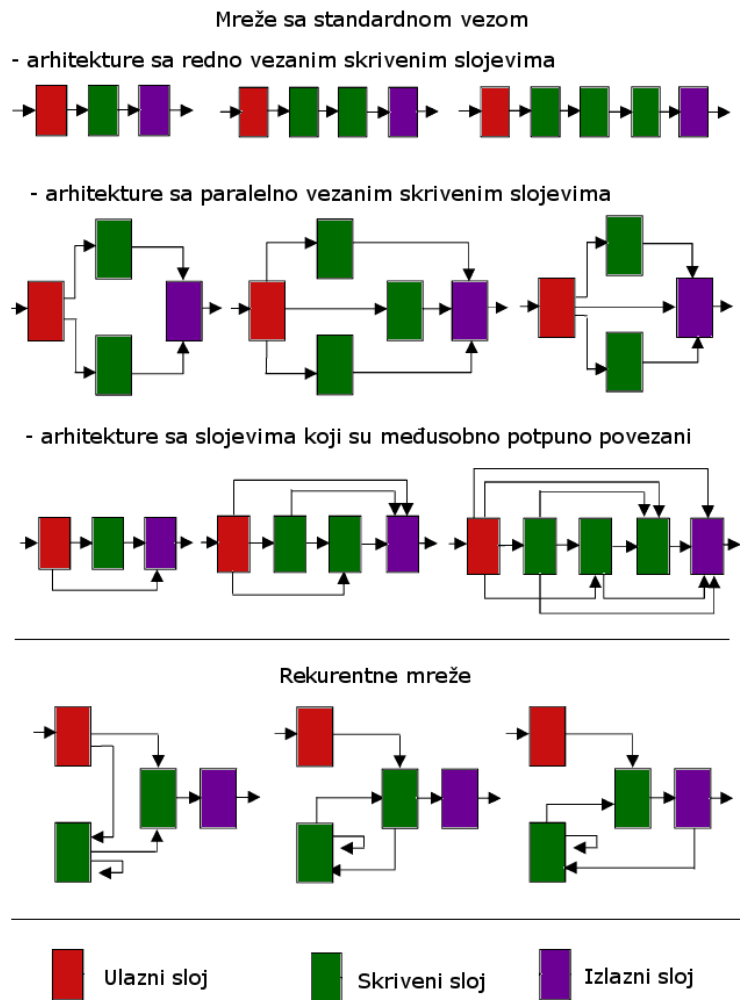


Slika 5. Nenadgledano obučavanje neuronske mreže

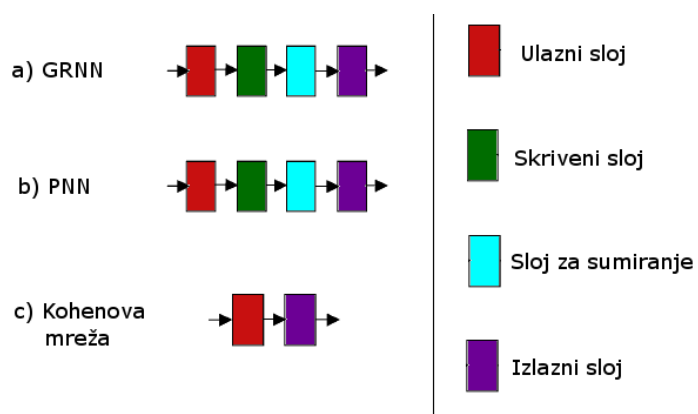
Sa stanovišta primene, jedna od najznačajnijih podela neuronskih mreža je u zavisnosti od arhitekture [86]:

1. Neuronske mreže sa propagacijom greške unazad (*Backpropagation neural networks - BPNNs*) (Slika 6.) su najuniverzalnije i najčešće korišćene, s obzirom da daju dobre rezultate za široki spektar problema. Mreže sa propagacijom greške unazad spadaju u mreže sa nadgledanim obučavanjem. Uglavnom se sastoje se od 3, 4 ili 5 slojeva, mada teorijski mogu da imaju i više slojeva neurona. U zavisnosti od toka signala između slojeva mogu se podeliti na:

- a) mreže sa standardnom vezom (*Feed-forward ANNs*), u kojima je svaki sloj povezan samo sa slojem koji mu neposredno prethodi,
 - b) rekurentne mreže (*Recurrent neural networks – RNN*) sa povratnom spregom u ulazni sloj i to iz samog ulaznog, skrivenog ili izlaznog sloja, ili sa skokovitim vezama, u kojima je sloj povezan sa svim prethodnim slojevima (*Feed-back ANNs*). Povratne veze se ponašaju kao “dugoročna memorija” mreže.
2. Neuronske mreže sa opštom regresijom (*General regression neural network - GRNN*) [87] služe za modelovanje višedimenzionalnog prostora, sa višedimenzionalnim ulaznim podacima. Relativno su jednostavne arhitekture: sastoje se od četiri sloja neurona (Slika 7.), pri čemu je broj neurona po slojevima definisan brojem ulaznih i izlaznih veličina, kao i brojem serija podataka.
 3. Probabilističke neuronske mreže (*Probabilistic neural networks - PNN*) [88] koriste se za klasifikaciju podataka u definisani broj kategorija. Sastoje se od četiri sloja neurona (GRNN su zapravo tip PNN) (Slika 7.),
 4. Kohonenove neuronske mreže (*Kohonen neural networks*) [89, 90] služe za klasifikaciju podataka u definisani broj kategorija i sastoje se od samo dva sloja neurona - ulaznog i izlaznog (Slika 7.). Po tipu su nenadgledane mreže, koje mogu da uče i klasifikuju bez poznavanja izmerenih vrednosti izlazne veličine.

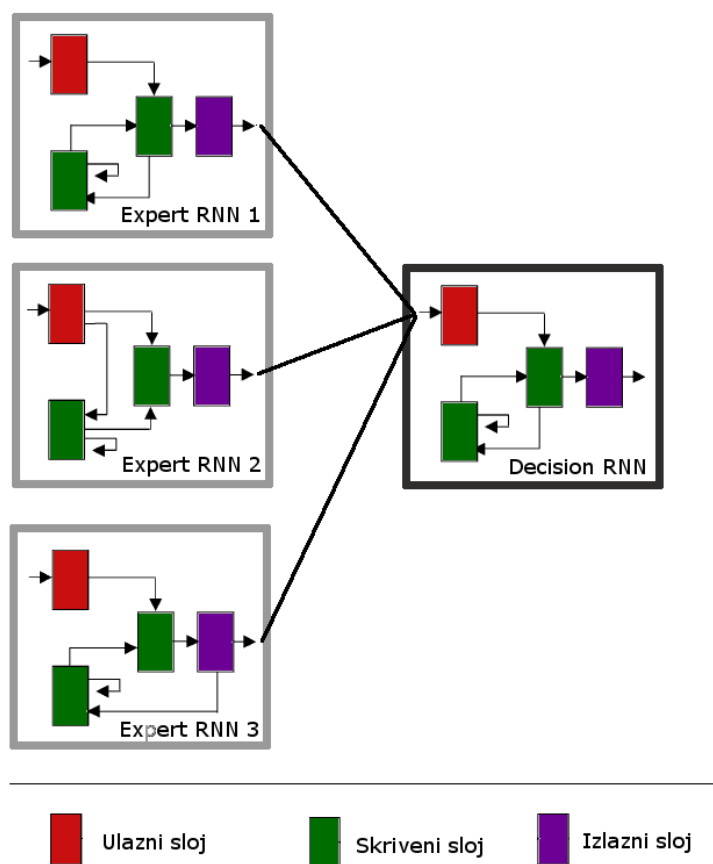


Slika 6. Arhitekture *Backpropagation* neuronskih mreža



Slika 7. Arhitekture: a) GRNN b) PNN c) Kohonene mreže

5. Modularnu mrežu (*Modular ANN*) čini sistem zasebnih rekurentnih mreža (RNN) gde svaka od njih uči na osnovu preuzetog jednog podskupa od ukupnog skupa podataka za obučavanje (Slika 8.). Zbog toga se i koristi kad je moguće ulazni skup podataka prirodno podeliti na više podskupova, koje onda obrađuju takozvane „ekspertske” mreže (*Expert network*). Izlaz iz modularne mreže daje takozvana „mreža koja odlučuje” (*Decision network*), koja obrađuje izlazne signale ekspertske mreže (koristi ih kao ulaz).



Slika 8. Primer modularne neuronske mreže

3.2. Obučavanje neuronskih mreža

Proces obučavanja je fundamentalan za neuronske mreže i u suštini predstavlja određivanje vrednosti težinskih koeficijenata veza između neurona. Cilj ovog procesa je da formiranjem skupa težinskih koeficijenata neuronska mreža da dobre i konzistentne prognoze [91].

Obučavanje neuronske mreže mreže se, u opštem slučaju, sprovodi tako što se ulazni podaci u prvom sloju normalizuju i prosleđuju drugom, skrivenom sloju. Neuron skrivenog sloja proizvode izlaz sumiranjem i ponderisanjem dobijenih vrednosti. Ovaj proces se ponavlja više puta, u zavisnosti od broja slojeva, dok na kraju izlazni sloj ne proizvede rezultat. Algoritmi i način podešavanja težinskih koeficijenata su unapred definisani i zavise od primenjene arhitekture neuronske mreže i odabranih parametara.

Najveći deo algoritama za obučavanje mreža zasniva se na prilagođavanju težinskih koeficijenata, radi postepenog smanjivanja greške (*gradient descent*) koju mreža pravi pri klasifikaciji ili prognozi [92]. Površina greške $\delta(W)$ u višedimenzionalnom prostoru težinskih koeficijenata daje grešku za svaku matricu težinskih koeficijenata (W). Kada je funkcija greške diferencijabilna, može se izračunati višedimenzioni vektor izvoda greške:

$$\nabla \delta = \left\langle \frac{\partial \delta}{\partial w_{i+1,0}}, \dots, \frac{\partial \delta}{\partial w_{i+1,i}} \right\rangle \quad (7)$$

gde su w_i težinski koeficijenti.

Ovaj gradijent pokazuje smer vektora težinskih koeficijenata kojim bi se, u slučaju infinitezimalne promene težinskih koeficijenata, dobio najveći porast greške. Naravno, za obučavanje je interesantan vektor suprotnog smera, kojim se maksimalno smanjuje greška:

$$W^* = W - \rho \nabla \delta(W) \quad (8)$$

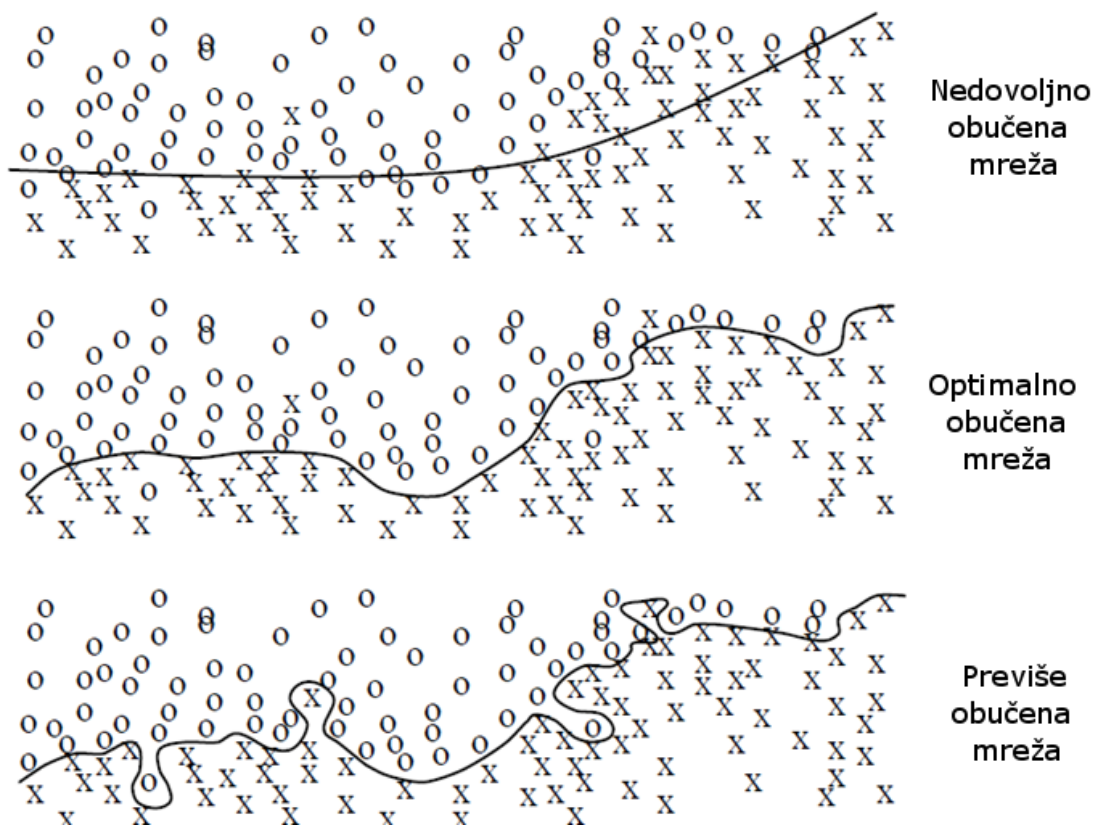
gde je W^* izmenjeni vektor/matrica težinskih koeficijenata, a ρ - veličina koraka.

Ideja ovog pristupa je u tome da se, počevši od proizvoljne matrice težinskih koeficijenata, izračunavanjem gradijenta greške, koeficijenti postepeno prilagođavaju, dok se ne dostigne neki kriterijum za zaustavljanje procesa učenja. Postoji nekoliko

problema vezanih za ovaj pristup, koji se javljaju u većini algoritama za obučavanje neuronskih mreža [92]:

1. Uslov diferencijabilnosti funkcije greške - u opštem slučaju izvod funkcije greške je jednak nuli ili je neodređen, pa je neophodno definisati aproksimativnu funkciju greške koja je diferencijabilna;
2. Zaustavljanje treninga mreže - funkcija greške može asimptotski težiti nekoj minimalnoj vrednosti, pa je potrebno definisati *ad hoc* kriterijum zaustavljanja procesa treniranja mreže;
3. Veličina iterativnog koraka - određivanje veličine koraka kod treninga mreže ne može se algoritamski definisati;
4. Nalaženje globalnog minimuma nije garantovano u opštem slučaju - algoritam može ostati zarobljen u nekom od lokalnih minimuma funkcije greške, ne nalazeći pritom globalni minimum;
5. Period ažuriranja težinskih koeficijenata - takođe se ne može algoritamski definisati (npr. ažuriranje posle svake serije ili celokupnog skupa dostupnih podataka);
6. Brzina treninga mreže - u praktičnim primenama često je izuzetno značajna karakteristika algoritma za obučavanje, a posebno u realnim uslovima.

Prave mogućnosti i upotrebljivost neuronske mreže mogu se sagledati tek kada se mreža testira i pruži dobre rezultate sa ulaznim podacima koje ranije nije imala na raspolaganju. Vođenje procesa obučavanja mreže predstavlja ključnu komponentu u njenom uspešnom razvoju. Uz pravilno definisanje arhitekture i parametara, jedan od najvažnijih problema je dimenzionisanje procesa obučavanja neuronske mreže. Ukoliko proces obučavanja nije dovoljno dug, mreža ne može da „nauči” zakonitosti problema [86]. Ukoliko je obučavanje predugo, mreža ne uči zakonitosti problema već memoriše dati uzorak, uključujući i šumove, i ne može da izvrši adekvatna predviđanja sa novim slučajevima (Slika 9.).



Slika 9. Izgled odziva neuronske mreže u zavisnosti od stepena obučenosti
(x-tačne vrednosti; o- odstupanja od tačnih vrednosti)

Pored vremena trajanja obučavanja mreže, prikazani odziv mreže na Slici 9. može zavistiti i od načina na koji su podaci prikazani neuronskoj mreži (*data representation*) [93], kao i od broja ulaza, pošto preveliki broj ulaza može dovesti do smanjivanja kapaciteta generalizacije kreiranog ANN modela [94]. Jedna od bitnih karakteristika ulaznih podataka je njihova međusobna povezanost, jer međusobno povezane promenljive unose zabunu u proces obučavanja [95]. Problem međusobno povezanih ulaznih promenljivih može se uspešno rešiti primenom različitih statističkih (hemometrijskih) metoda: korelacionom analizom, analizom glavnih komponentata (*Principal component analysis* – PCA) ili genetskim algoritmima (GA) [96, 97]. Navedene hemometrijske metode biće detaljno opisane u Poglavlju 4.

3.3. Prikaz odabranih arhitektura neuronskih mreža

3.3.1. Neuronske mreže sa propagacijom greške unazad (BPNN)

Backpropagation algoritam razvijen je od strane *Rumelhart*-a i saradnika 1986. godine [98]. Po tipu je algoritam sa postepenim smanjivanjem greške (*gradient descent*) kod koga se težinski koeficijenti veza između neurona postepeno, kroz iteracije, menjaju u pravcu poboljšanja tačnosti mreže. Postoje brojne varijante (tipovi) osnovnog algoritma, koje su bazirane na heurističkim (*heuristic*) i standardnim numeričkim optimizacionim metodama [99]:

1. sa opadajućim gradijentom stope obučavanja (*learning rate gradient descent*),
2. sa konjugovanim gradijentom stope obučavanja (*learning rate conjugate gradient*), i
3. sa Njutnovom metodom optimizacije.

Kod prvog tipa *backpropagation* algoritma (BP) implementacijom se težinski koeficijenti (w) menjaju postepeno, u smeru u kome se najbrže poboljšava preciznost modela, tj. smanjuje se razlika između predviđene i tačne vrednosti [100]:

$$w_{k+1} = w_k - \alpha_k g_k \quad (9)$$

gde je w_k vektor trenutnih težinskih koeficijenata, g_k trenutni gradijent greške, α_k stopa obučavanja (*learning rate*). Ova varijanta BP algoritma bazirana je na heurističkoj tehnici [101], za razliku od standardnog BP algoritma je znatno brža i omogućena je promena stope obučavanja tokom procesa.

Kod BP algoritma sa konjugovanim gradijentom, potraga za minimalnom greškom ide u suprotnom smeru, u odnosu na prvi prikazani algoritam:

$$w_{k+1} = w_k + \alpha_k p_k \quad (10)$$

gde je p_k konjugovani gradijent greške koji se računa na osnovu vrednosti prethodnog gradijenta greške:

$$p_k = -g_k + \beta_k p_{k-1} \quad (11)$$

Postoji nekoliko varijanti ovog algoritma, u zavisnosti od načina izračunavanja konstante β_k , koja se često navodi kao impuls (*momentum*) [100, 102]:

- *Fletcher-Reeves* tehnika,
- *Polak-Ribière* tehnika,
- *Powell-Beale* tehnika.

BP algoritam baziran na Njutnovoju metodi optimizacije, važi za znatno brži algoritam od prethodno prikazanih varijanti BP algoritma. Iteracije su zasnovane na sledećem proračunu:

$$w_{k+1} = w_k - A_k^{-1} g_k \quad (12)$$

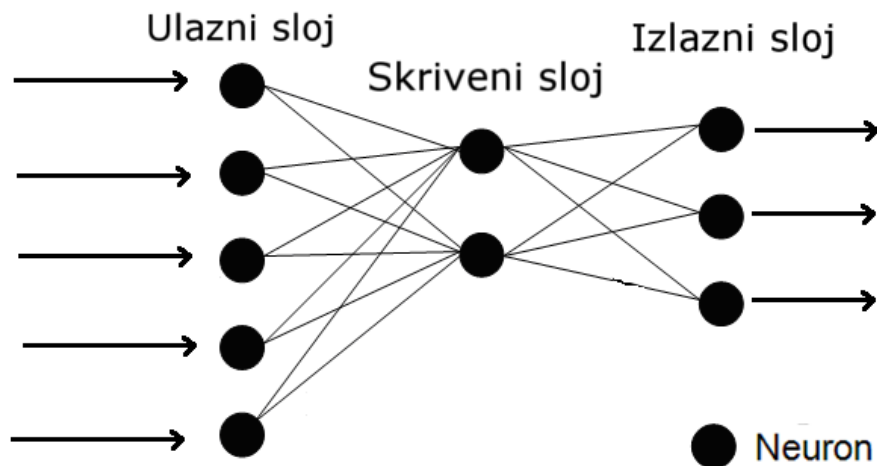
gde je A_k *Hessian* matrica drugog izvoda funkcije greške (pokazatelja performansi modela) za trenutne vrednosti težinskih koeficijenata. Proračun *Hessian* matrice nakon svake iteracije je kompleksan proces i obično se vrši:

- *Quasi-Newton* numeričkom metodom [99], ili
- *Levenberg-Marquardt* algoritmom, koji koristi aproksimaciju *Hessian* matrice [103].

Kao mera (indikator) greške (e) obično se koristi srednja vrednost kvadrata greške (*mean sum of squares* – MSE):

$$MSE = \frac{1}{n} \sum_{i=1}^n e^2 \quad (13)$$

Kako u neuronske mreže sa propagacijom greške unazad spada više tipova različitih arhitektura (Slika 6.), kojima je zajedničko da se obučavanje vrši nekom varijantom *backpropagation* algoritma, u daljem tekstu biće analizirana troslojna BPNN mreža sa standardnom (*Feed-forward*) vezom (Slika 10.).



Slika 10. Troslojna BPNN mreža sa standardnom vezom

Kao što se vidi sa Slike 10. broj neurona zavisi od broja ulaznih i izlaznih promenljivih:

- broj neurona u ulaznom sloju (N_{ul}) jednak je broju ulaznih promenljivih,
- broj neurona u izlaznom sloju (N_{iz}) jednak je broju izlaznih promenljivih,
- broj neurona u skrivenom sloju (N_{sk}) određuje se ili empirijski (metodom probe i greške) ili prema jednačini:

$$N_{sk} = \frac{N_{ul} + N_{iz}}{2} + \sqrt{n_{tr}} \quad (14)$$

gde je n_{tr} broj serija podataka u grupi podataka za obučavanje mreže.

Ostali parametri arhitekture, koje je potrebno definisati kod BPNN, su:

- funkcija za skaliranje,
- aktivacione funkcije po slojevima,
- inicijalne vrednosti težinskih koeficijenata,
- stopa obučavanja (*learning rate*), i
- impuls (*momentum*).

Funkcija za skaliranje predstavlja vezu između ulaznih i skrivenih neurona, i može biti linearna funkcija u različitim opsezima (-1 do 1, 0 do 1 i sl.), sigmoidna funkcija ili hiperbolično-tangetna funkcija. Aktivacione funkcije se definišu za neurone u skrivenom i izlaznom sloju, pri čemu se najčešće koriste linearna, logička, sinusna, hiperbolično-tangetna i Gausova funkcija.

Početne vrednosti težinskih koeficijenata se moraju zadati, da bi se u prvoj iteraciji BP algoritma mogla izračunati matrica težinskih koeficijenata. U praksi se zadaje opseg vrednosti (od -1 do +1), iz koga se početna vrednost nasumično bira.

Stopom obučavanja (α_k u jednačini (9)) se zadaje brzina promene težinskih koeficijenata, npr. ukoliko je stopa obučavanja 0,5 tada se težinski koeficijenti menjaju za polovinu vrednosti greške. Što je stopa obučavanja veća to je brža konvergencija mreže, s tim što ukoliko se zada prevelika vrednost stope obučavanja može doći do oscilacija u performansama ili čak do divergencije mreže (nenalaženje rešenja). Stopa obučavanja najčešće iznosi 0,1.

Impuls (β_k u jednačini (11)) definiše koliki će biti udeo prethodne vrednosti težinskih koeficijenata u njihovoj novoj vrednosti. Vrednost impulsa zavisi od složenosti problema: što je problem kompleksniji to impuls ima nižu vrednost i kreće se od 0,6 do 0,05, dok je obično jednak 0,1 [104].

Po tipu obučavanja, BPNN spada u nadgledano obučavane mreže. Pre samog obučavanja potrebno je definisati nekoliko parametara koji određuju način obučavanja mreže. Najpre je potrebno odabrati varijantu BP algoritma koja će biti korišćena za određivanje težinskih koeficijenata:

- *Vanilla* algoritam je BP algoritam sa opadajućim gradijentom stope obučavanja – koristi se samo stopa obučavanja za određivanje optimalnih težinskih koeficijenata,
- *Momentum* algoritam je BP algoritam sa konjugovanim gradijentom stope obučavanja – pored stope obučavanja koristi se impuls za određivanje optimalnih težinskih koeficijenata.

Kao što je već navedeno, unapređene verzije BP algoritma omogućavaju i promene inicijalnih vrednosti stope obučavanja i impulsa, pri čemu se korak promene navedenih parametara definiše neposredno pre samog obučavanja mreže.

Kako BP algoritam nema mogućnost samozaustavljanja, mreža se obučava sve dok se proces ne zaustavi od strane korisnika. Ovo je uzrok i fenomena "pretreniranosti" (*overtrained*) koji podrazumeva dobijanje mreže sa slabim kapacitetom uopštavanja i bukvalno memorisanje vrednosti koje su korišćene kod obučavanja. Da bi se izbegla pojava "pretreniranosti", poželjno je da se tokom obučavanja vrši evaluacija mreže validacionim podacima, pri čemu je potrebno u određenim vremenskim intervalima (npr. na svakih 30 min.) proveriti tačnost mreže. Ukoliko je veličina greške mreže zadovoljavajuća i ukoliko se ona nije smanjivala tokom većeg broja iteracija, može se smatrati da je mreža konvergirala i da su optimalne vrednosti težinskih koeficijenata određene. U tom slučaju obučavanje mreže se prekida i prelazi se na proveru samog modela uz korišćenje novih – mreži nepoznatih podataka, kako bi se ispitala realna sposobnost uopštavanja razvijenog modela.

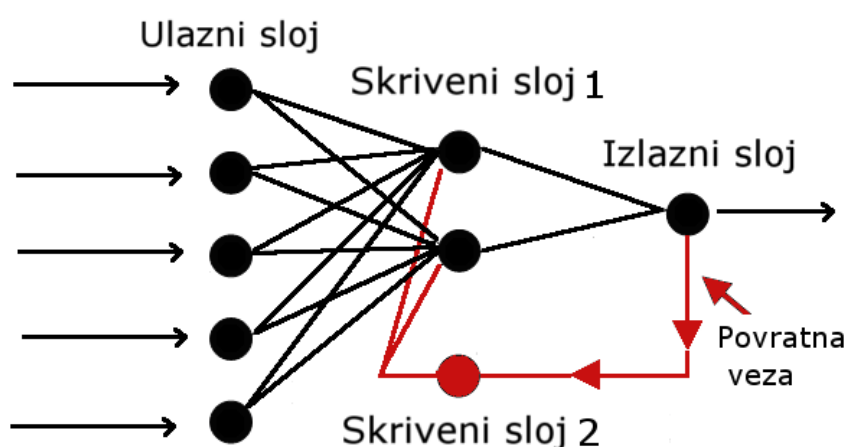
Po završenom obučavanju može se ispitati i značajnost pojedinih ulaznih promenljivih, na osnovu težinskih koeficijenata određenih tokom obučavanja mreže. Što je njihova vrednost veća, to i pojedinačna ulazna promenljiva značajnije utiče na vrednost izlazne promenljive.

3.3.2. Rekurentne neuronske mreže (RNN)

Kako rekurentne neuronske mreže spadaju u BPNN, one se takođe nadgledano obučavaju BP algoritmom koji je prikazan u Poglavlju 3.3.1., pri čemu je i sam proces obučavanja identičan. RNN mreže su odličan izbor kada se radi o podacima koji predstavljaju vremenske serije, pošto imaju sposobnost da uče sekvencijalne zavisnosti u podacima.

Postoji više različitih arhitektura koje su po tipu rekurentne neuronske mreže (Slika 6.). U ovom poglavlju će biti prikazana RNN arhitektura poznata kao *Jordan-Elman* mreža, koja je naziv dobila po istraživačima koji su je razvili [105]. Navedena arhitektura ima četiri sloja, pored ulaznog, skrivenog i izlaznog sloja, postoji i

skriveni sloj koji služi za povratni proračun (*dampened feedback*) i povezan je povratno sa izlaznim slojem, pri čemu signal dalje ide u prvi skriveni sloj (Slika 11.). Ovaj dodatni skriveni sloj se često naziva i „memorijom“ mreže, pošto se u njemu memoriše izlaz mreže iz prethodne iteracije. Broj neurona u ulaznom i izlaznom sloju je jednak broju ulaznih i izlaznih promenljivih, dok se broj neurona u skrivenom sloju računa kao i kod BPNN (jednačina (14)). Broj neurona u drugom skrivenom sloju jednak je broju neurona u izlaznom sloju. Ostali parametri arhitekture se definišu po istim pravilima kao i kod BPNN.



Slika 11. Arhitektura rekurentne neuronske mreže

3.3.3. Neuronske mreže sa opštom regresijom (GRNN)

Razvoj neuronskih mreža sa opštom regresijom (*General regression neural networks* - GRNN) vezuje se za rad *D. Specht*-a [87], publikovan 1991. godine, iako su praktično u isto vreme *Schiøler* and *Hartmann* [106] došli do iste teorije, koja je objavljena 1992. godine [107, 108]. Drugi razlog, zbog kojeg se otkriće GRNN teorije vezuje za *Specht*-a, je činjenica da GRNN zapravo predstavlja tip probabilističke neuronske mreže (PNN), čiju teoriju je *Specht* objavio 1990. godine [88]. Kod PNN se koristi funkcija gustine (raspodele) verovatnoće (*probability density function*) za predviđanje diskretnih vrednosti verovatnoća određenog događaja, dok je kod GRNN ovaj princip proširen i primenjen za određivanje kontinualnih vrednosti promenljivih. GRNN ima veliku prednost kod primene u praksi, jer u poređenju sa

BPNN, može da se obučava mnogo brže, sa znatno manje podataka, kao i sa nepotpunim serijama podataka [87].

GRNN algoritam se može posmatrati kao tip *Nadaraya-Watson kernel* regresije (*Nadaraya-Watson kernel regression*) [109, 110], kod koje se utvrđuje statistička veza i konvergencija velikog broja uzoraka sa optimalnom regresionom površinom u multidimenzionom prostoru [111].

Svako regresiono određivanje zavisne promenljive y u odnosu na nezavisnu promenljivu x , predstavlja određivanje najverovatnije vrednosti y (čija tačnost zavisi od tačnosti primenjene regresione metode). Određivanje nepoznate vrednosti y , za poznatu vrednost x , iziskuje pretpostavku neke funkcijske zavisnosti sa nepoznatim parametrima, npr. kod linearne regresije nepoznati parametar je koeficijent nagiba krive. Kod GRNN algoritma ova funkcijska zavisnost se izražava u vidu funkcije raspodele verovatnoće $f(x, y)$, koja se određuje na osnovu vrednosti x koristeći *Parzen-ovu procenu* (*Parzen window estimation*) [112]. Regresija y za poznatu vrednost x (X) se može prikazati jednačinom:

$$E[y | X] = \frac{\int_{-\infty}^{\infty} y f(X, y) dy}{\int_{-\infty}^{\infty} f(X, y) dy} \quad (15)$$

Tokom obučavanja mreže funkcija raspodele $f(x, y)$ nije poznata, pa se određuje na osnovu poznatih vrednosti promenljivih x i y . Ovim proračunom se zapravo predviđanje nepoznate vrednosti y vrši na osnovu verovatnoće čiji opseg ("širina") zavisi od parametra koji je nazvan faktor ujednačavanja (*smoothing factor* - σ_f), a koji se određuje za svaki par Y i X , pa je konačna verovatnoća jednaka zbiru pojedinačnih verovatnoća [87]. Faktor ujednačavanja je veći od 0, i što je njegova vrednost bliža nuli to su regresione površine glatkije i nelinearnije, a time su i tačnije predikcije koje algoritam daje za nepoznatu vrednost y [111].

U praksi GRNN algoritam radi tako što upoređuje rastojanje između ulaznih podataka (vektora) i predviđenih vrednosti, prema sledećoj jednačini:

$$Y(X) = \frac{\sum_{i=1}^n Y_i \exp\left(\frac{-D_i^2}{2\sigma_f^2}\right)}{\sum_{i=1}^n \exp\left(\frac{-D_i^2}{2\sigma_f^2}\right)} \quad (16)$$

gde je $Y(X)$ vrednost dobijena primenom GRNN mreže za ulaz X , Y_i izmerena (tačna) vrednost, D_i rastojanje vrednosti ulaznih podataka od vrednosti predviđenih mrežom (u multidimenzionom prostoru, čija dimenzionalnost zavisi od broja ulaznih promenljivih), dok je σ_f ranije definisani faktor ujednačavanja.

U prikazanoj jednačini (16), Y_i ima ulogu da obezbedi da predviđanja budu tačnija za veće vrednosti promenljive y , za koje je kritično odrediti što tačnija predviđanja, jer se na taj način efikasno smanjuje srednja apsolutna greška predviđanja.

Parametar D_i određuje vrednost \exp dela jednačine: kada $D_i \rightarrow 0 \Rightarrow \exp \rightarrow 1$, čime se direktno utiče na vrednost izlazne promenljive. Kako je D_i mera rastojanja pojedinačnih vrednosti X i predviđenih Y , on zapravo ima ulogu da izdvoji značajne promenljive od manje značajnih: što je značajnost nekog ulaza manja, to je D_i veće, a vrednost \exp dela jednačine manja, pa samim tim će ulazni podaci navedene promenljive manje uticati na vrednost predviđanja u konačnom zbiru.

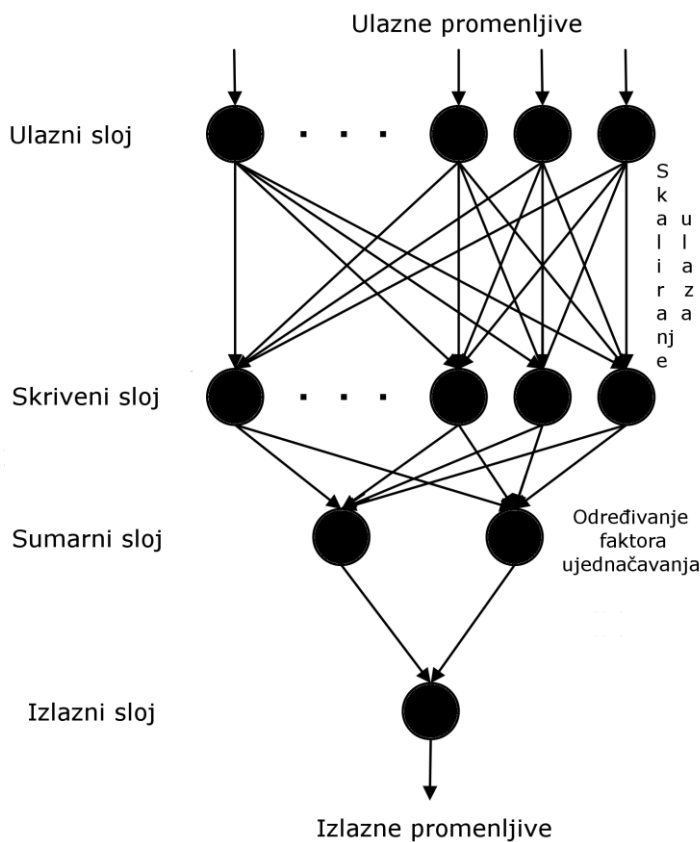
Faktor ujednačavanja zapravo predstavlja standardnu devijaciju i to je jedini nepoznati parametar koji se mora odrediti tokom samog obučavanja mreže. Kako se kod GRNN mreže pretpostavlja normalna raspodela, faktor ujednačavanja se može interpretirati i kao širina Gausove krive za svaku funkciju raspodele verovatnoće koja se tokom obučavanja odredi za svaku ulaznu promenljivu. Faktor ujednačavanja ima vrednosti veće od 0, i obično se kreće od 0,01 do 1, pri čemu manje vrednosti ukazuju na optimalnije obučenu mrežu [113].

Arhitektura GRNN mreže se sastoji od četiri sloja neurona: ulaznog, skrivenog, sumarnog i izlaznog sloja (Slika 12.). Broj neurona zavisi od broja ulaznih i izlaznih promenljivih i broja serija podataka koji su dostupni za obučavanje mreže:

- broj neurona u ulaznom sloju jednak je broju ulaznih promenljivih,

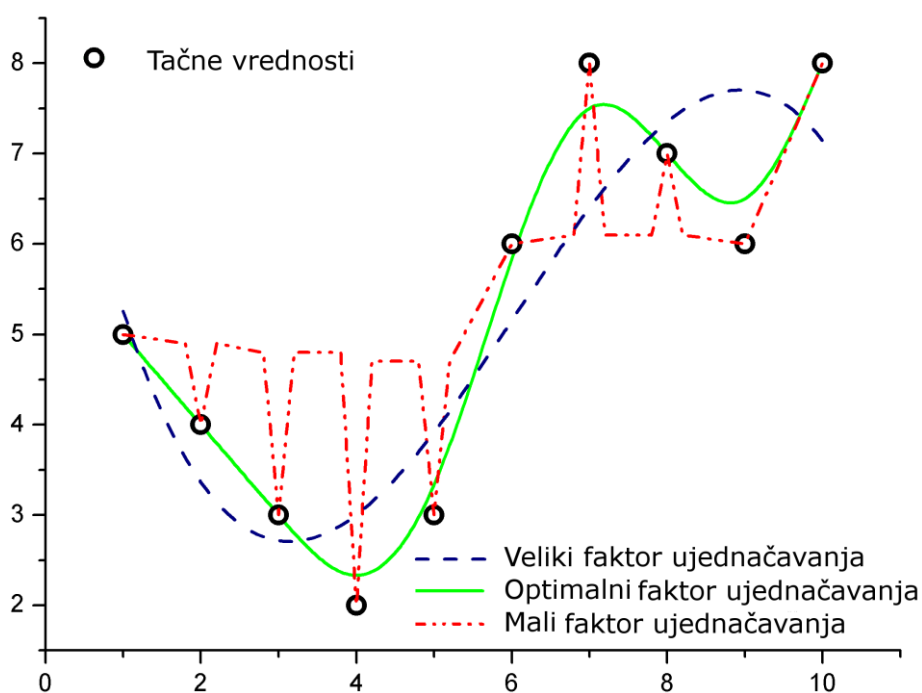
- broj neurona u skrivenom sloju jednak je broju serija podataka u grupi podataka za obučavanje,
- broj neurona u sumarnom sloju jednak je broju izlaznih promenljivih plus jedan, i
- broj neurona u izlaznom sloju jednak je broju izlaznih promenljivih.

Neki autori sumarni sloj neurona prikazuju kao deo skrivenog sloja, pa se u tom slučaju GRNN navodi kao troslojna neuronska mreža. Od parametara arhitekture, jedino se funkcija za skaliranje može odabrati u zavisnosti od tipa problema. Funkcija za skaliranje je veza između ulaznih i skrivenih neurona, i može biti linearna funkcija u različitim opsezima (-1 do 1, 0 do 1 i sl.), sigmoidna funkcija, hiperbolično-tangetna funkcija. Veza između ulaznih i skrivenih neurona se može uspostaviti i direktno (bez skaliranja), ali se u praksi to izbegava da bi se obezbedilo efikasnije obučavanje mreže.



Slika 12. GRNN arhitektura

Po tipu obučavanja, GRNN spada u nadgledane mreže sa jednom iteracijom tokom obučavanja. Zbog ove činjenice, kod GRNN mreže se ne javlja "pretreniranost" u smislu u kojem se ona javlja kod BPNN i RNN mreža. Međutim, u drugom delu obučavanja, u kome se vrši određivanje faktora ujednačavanja, ukoliko se ne odredi optimalan faktor ujednačavanja, mreža se može ponašati kao da je "pretrenirana". Na taj način, faktor ujednačavanja je zapravo kritični parametar obučavanja GRNN mreže, koji direktno određuje njenu preciznost (Slika 13.). Samo optimalna vrednost faktora ujednačavanja obezbeđuje da predviđanja imaju sličnu tačnost u celom opsegu vrednosti izlazne promenljive. GRNN model sa faktorom ujednačavanja većim od optimalnog imaće jako neprecizna predviđanja, dok će faktor ujednačavanja manji od optimalnog imati za posledicu da mreža zapravo samo memoriše izlazne vrednosti iz grupe podataka za obučavanje.



Slika 13. Odziv GRNN mreže u zavisnosti od vrednosti faktora ujednačavanja

Određivanje faktora ujednačavanja može se vršiti:

- manuelnim testiranjem, sve dok se ne dobije zadovoljavajuća preciznost mreže, ili
- automatskim testiranjem, uz korišćenje iterativnog ili genetskog algoritma.

U praksi se preporučuje korišćenje algoritama da bi se faktor ujednačavanja što preciznije odredio [114]. Ostali parametri obučavanja zavise od izbora metoda za određivanje faktora ujednačavanja:

1. kod manuelnog određivanja obučavanje se svodi na primenu prethodno izabranog faktora ujednačavanja, koji se zatim manuelno menja u zavisnosti od ostvarene tačnosti;
2. iterativni algoritam – jedino je potrebno definisati opseg vrednosti faktora ujednačavanja koje je potrebno testirati (od 0 do 0,8, ..., 0 do 12,8), pri čemu se najčešće faktor ujednačavanja određuje u opsegu 0-1,6;
3. genetski algoritam –potrebno je definisati *Genetic Breeding Pool Size* koji se može interpretirati kao opseg broja faktora ujednačavanja koji će se genetskim algoritmom testirati, a potrebno je definisati i kriterijum za zaustavljanje obučavanja koji se najčešće definiše kao 20 generacija (iteracija) bez poboljšanja tačnosti od 1 %.

Određivanje faktora ujednačavanja genetskim algoritmom daje, pored globalnog faktora ujednačavanja (*overall smoothing factor*), i pojedinačne faktore ujednačavanja (*Individual Smoothing Factor - ISF*), koji su karakteristični za svaku od ulaznih promenljivih. Vrednost pojedinačnih faktora ujednačavanja se kreće od 0 do 3 i pokazuje koliko pojedinačni ulaz ima uticaja na izlaznu promenljivu kreiranog modela. Što je vrednost pojedinačnog faktora ujednačavanja veća to je veći i njegov uticaj na model. Zbog ove činjenice, pojedinačni faktori ujednačavanja mogu se koristiti za analizu značajnosti i kao kriterijum za eliminaciju manje značajnih i beznačajnih ulaza, što će biti prikazano u Poglavlju 6.

3.4. Oblasti primene neuronskih mreža

U ovom delu su prikazane glavne oblasti u kojima se ANN već duži niz godina uspešno primenjuju, pri čemu broj tipova problema na koje se ANN primenjuju svake godine sve više raste. Značajnija praktična primena neuronskih mreža počinje tek razvojem računarskih tehnologija, kada su algoritmi postali pristupačni za opštu upotrebu. Naročito dobre rezultate pokazuju na području klasifikacije i

aproksimacije, na problemima mapiranja, kao i prilikom predviđanja i modelovanja sistema kod kojih fizički procesi nisu jasni ili su veoma kompleksni. Između ostalog primenjuju se za: prepoznavanje oblika, rukopisa i govora; finansijske i ekonomske modele; predviđanje kretanja cena na tržištu; upravljanje sistemima; upravljanje proizvodnim procesima; analizu električnih kola; psihijatrijske procene; kompresovanje podataka; naftna istraživanja; kriminološka istraživanja; analizu medicinskih testova; ispitivanje EEG i EKG signala; pronalaženje optimalnog rešenja; upravljanje robotima; analiziranje podataka pri pirolizi i spektroskopiji; vremensku prognozu.

Poslednjih godina su neuronske mreže intenzivno primenjivane za simulacije i predviđanja u oblasti zaštite životne sredine. Korišćene su za modelovanje atmosfere disperzije [86], procesa eutrofikacije vodenih tokova u Turskoj [115], kod predviđanja učinka postrojenja za prečišćavanje otpadnih voda [116], za modelovanje ukupne količine azota u rekama Japana [117], za klasifikaciju mahovina i lišajeva [118], za predviđanje nastalih količina medicinskog otpada [56] i čvrstog otpada [55, 85], za modelovanje izluživanja materija iz čvrstog otpada [57], za predviđanje toplotne moći komunalnog čvrstog otpada [119], za predviđanje količina nastalog biogasa digestijom otpada [120], za predviđanje emisije gasova sa efektom staklene bašte [21, 22], za predviđanje časovnih koncentracija zagađujućih materija vazduha i to NO_x , CO, O_3 i PM_{10} [38, 84, 121-126], za predviđanje dugoročnih koncentracija istih zagađujućih materija vazduha u urbanim područjima [127-129], za modelovanje i kontrolu procesa sagorevanja [130] i za modelovanje potrošnje energije [20, 131, 132].

4. Hemometrijske metode primenjene kod odabira ulaznih promenljivih modela

4.1. Korelaciona analiza

Korelacionom analizom se utvrđuje linearna zavisnost dve promenljive, pri čemu se Parsonovim koeficijentom korelacije (r) kvantifikuje navedena linearna zavisnost. Koeficijent korelacije se kreće od -1 do +1, gde vrednost +1 predstavlja maksimalnu pozitivnu linearnu povezanost (npr. ako vrednost jedne promenljive raste, raste i vrednost promenljive sa kojom je linearno povezana), dok -1 predstavlja maksimalnu negativnu linearnu korelaciju (npr. rast vrednosti jedne promenljive se podudara sa opadanjem vrednosti promenljive sa kojom je linearno povezan). Kada je $r=0$, radi se o paru potpuno nekorelisanih promenljivih. Koeficijent korelacije se izračunava na osnovu formule:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n-1)\sigma_x\sigma_y} \quad (17)$$

gde su x i y vrednosti promenljivih čija korelacija se ispituje, \bar{x} i \bar{y} srednje vrednosti ispitivane populacije ($i=1\dots n$) i σ_x , odnosno σ_y , standardne devijacije (jednačina (18)) [133].

4.2. Analiza glavnih komponenti (PCA)

Analiza glavnih komponenti (*Principal Component Analysis* – PCA) predstavlja statističku metodu koja ima najveću primenu kada je potrebno izvršiti obradu podataka sa više promenljivih. Mogućnost analize glavnih komponenti prvi je opisao *Karl Pearson* još 1901. godine [134], ali je tek 30 godina kasnije *Hotelling* razradio praktične računске aspekte metode, dok je šira primena PCA počela 80-tih godina prošlog veka, razvojem računarske tehnologije [135]. Upravo razvojem statističkih softverskih paketa, PCA je postala jedna od najpopularnijih tehnika za obradu podataka i redukovanje dimenzija, sa velikom primenom u inženjerstvu, biološkim, ekonomskim i društvenim naukama.

Suština PCA tehnike je u identifikaciji i izračunavanju varijanse i korelacije između podataka. Varijansa promenljive X jednaka je kvadratu standardne devijacije (σ_x) i

pokazuje koliko se podaci međusobno razlikuju:

$$\sigma_{\bar{X}}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \quad (18)$$

Prilikom izračunavanja u PCA se pretpostavlja linearna korelacija između promenljivih, a kvantifikuje se koeficijentom linearne korelacije (između dve promenljive X_1 i X_2):

$$\rho(X_1, X_2) = \frac{\frac{1}{n} \sum_{i=1}^n (X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2)}{\sigma_{X_1} \sigma_{X_2}} \quad (19)$$

Navedenom jednačinom korelacija se kvantifikuje u opsegu od -1 (maksimalna negativna korelacija) do +1 (maksimalna pozitivna korelacija).

Nakon određivanja varijanse i korelacije među podacima, PCA analizom utvrđuju se glavne komponente, PC_1, PC_2, \dots, PC_n , koje predstavljaju linearne kombinacije originalnih promenljivih X_1, X_2, \dots, X_m , pri čemu linearni koeficijenti (a_{11}, \dots, a_{nm}) nisu međusobno korelisani:

$$PC_1 = a_{11}X_1 + a_{12}X_2 + \dots + a_{1m}X_m \quad (20)$$

$$PC_2 = a_{21}X_1 + a_{22}X_2 + \dots + a_{2m}X_m \quad (21)$$

...

$$PC_n = a_{n1}X_1 + a_{n2}X_2 + \dots + a_{nm}X_m \quad (22)$$

Prema tome, ukoliko između originalnih promenljivih postoji značajna korelacija, broj glavnih komponenti će biti mnogo manji od broja originalnih promenljivih. Ukoliko je 20-30 originalnih promenljivih visoko pozitivno ili negativno korelisano, broj neophodnih glavnih komponenti može biti svega 2 ili 3 [135].

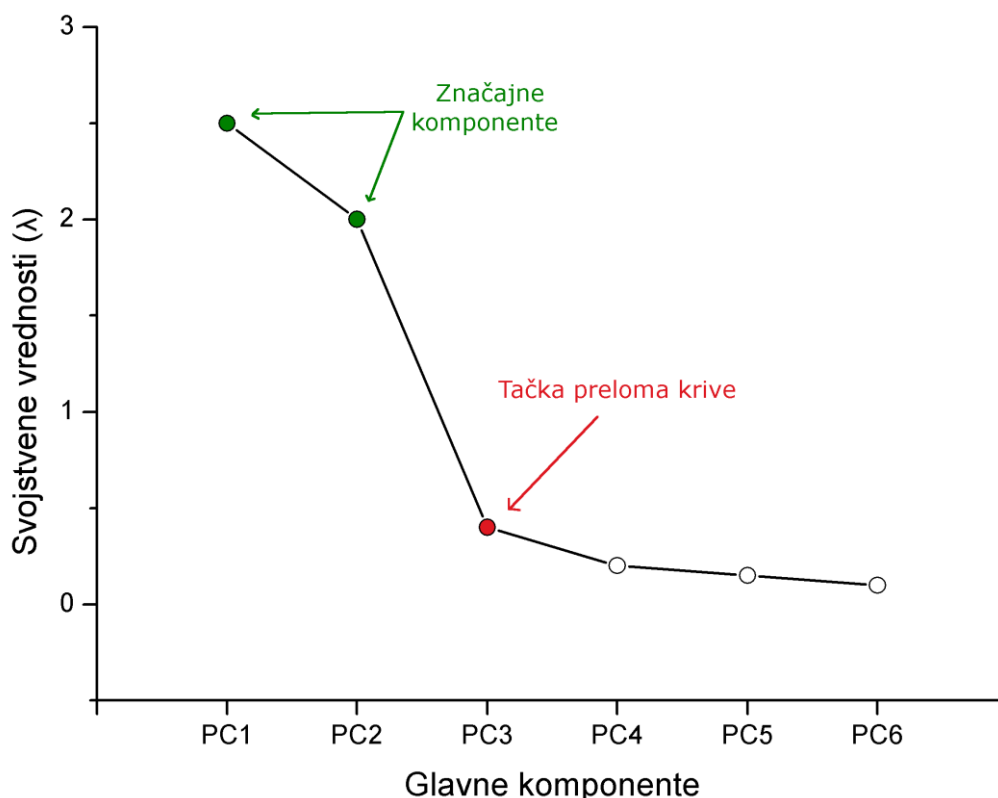
Glavne komponente se definišu prema veličini varijanse u seriji podataka koje predstavljaju: prva (PC_1) predstavlja najveće odstupanje, druga (PC_2) predstavlja drugo odstupanje, po veličini, itd. Kvantifikacija varijanse koju predstavljaju pojedinačne komponente vrši se na osnovu svojstvene vrednosti λ korelacione matrice (C), prema jednačini:

$$|C - \lambda I| = 0 \quad (23)$$

gde je I jedinična matrica (*identity matrix*) [80].

Svojtvene vrednosti korelacione matrice (λ) se takođe koriste da bi se odredio broj glavnih komponenti koje je potrebno izdvojiti, pri čemu se koristi više različitih metoda. Prema *Kaiser*-ovoj metodi, sve glavne komponente koje imaju λ vrednost manju od 1 treba zanemariti. Glavni nedostatak ove metode određivanja broja glavnih komponenti je što bi se pri doslovnoj primeni *PC* sa λ vrednošću od 0,99 bila zanemarena, dok bi *PC* sa $\lambda=1,01$ bila zadržana, iako između njih praktično nema razlike u varijansi koju definišu.

Drugi metod je takozvani *Cattell scree test*, kod koga se opadajuće λ vrednosti nose na dijagram (Slika 14.). Broj značajnih komponenti se određuje sa dijagrama, pri čemu se za značajne komponente uzimaju one komponente koje se pojavljuju pre „preloma“ krive (pre PC_3 na Slici 14.) [136].



Slika 14. *Cattell scree test* - određivanje broja značajnih glavnih komponenti

U nekim situacijama je teško odrediti u kom delu krive dolazi do preloma, pa je neophodno dodati još jedan kriterijum, da bi se došlo do broja značajnih glavnih komponenti. Kao dodatni kriterijum često se primenjuje procenat od ukupne varijanse, koji je definisan odabranim brojem glavnih komponenti. Kako se svojstvenim vrednostima korealacione matrice (λ) zapravo kvantifikuje obuhvaćena varijansa određenom glavnom komponentom, procenat ukupne varijanse (r_q) obuhvaćene odabranim glavnim komponentama, može se izračunati kao količnik zbira λ vrednosti odabranih glavnih komponenti sa zbirom λ vrednosti svih glavnih komponenti:

$$r_q = \frac{\sum_{j=1}^q \lambda_j}{\sum_{j=1}^p \lambda_j} \cdot 100 \quad (24)$$

gde je p ukupan broj glavnih komponenti, a q broj izdvojenih glavnih komponenti. Na ovaj način se broj potrebnih komponenti određuje na osnovu udela varijanse koji je potrebno izraziti. U opštem slučaju potrebno je da procenat obuhvaćene varijanse bude bar 60 %. U praksi, kada je PCA pravilno primenjena sa dve glavne komponente, moguće je obuhvatiti potrebnih 60 %, dok se ostale komponente smatraju za „šum” [137].

Pored navedenog kriterijuma za ispitivanje da li je PCA pravilno primenjena (60 % varijanse definisane sa dve komponente), često se primenjuju i posebni statistički testovi, kako bi se utvrdilo da li je PCA uopšte primenljiva na određenu grupu podataka. U tu svrhu se najčešće koriste *Kaiser-Meyer-Olkin* (KMO) test adekvatnosti podataka i *Bartlett-ov test sferičnosti* (*Bartlett's Sphericity test*). KMO test adekvatnosti podataka poredi vrednosti posmatranih koeficijenata korelacije sa parcijalnim vrednostima koeficijenata korelacije. Vrednost KMO testa se kreće od 0 do 1, pri čemu male vrednosti testa ukazuju da se radi o korelisanim parovima promenljivih, koje ne mogu biti predstavljene nekom drugom promenljivom, glavnom komponentom, pa samim tim PCA nije primenljiva.

Vrednosti KMO testa se često interpretiraju na sledeći način:

- vrednosti preko 0,9 se opisuju kao „izvanredne“,
- vrednosti od 0,8 do 0,9 se opisuju kao „dobre“,
- vrednosti od 0,7 do 0,8 se opisuju kao „osrednje“,
- vrednosti od 0,6 do 0,7 se opisuju kao „prosečne“,
- vrednosti od 0,5 do 0,6 se opisuju kao „loše“,
- vrednosti ispod 0,5 su neprihvatljive.

Bartlett-ov test sferičnosti poredi dobijenu korelacionu matricu sa matricom bez korelacija (*the identity matrix*) [138]:

$$\chi_k^2 = \left[n - k - \frac{2(p-k)+7+\frac{2}{p-k}}{6} + \sum_{j=1}^k \left(\frac{\bar{\lambda}}{\lambda_j + \bar{\lambda}} \right)^2 \right] \times \left[-\ln \prod_{j=k+1}^p \lambda_j + (p-k) \ln \bar{\lambda} \right] \quad (25)$$

gde je p broj glavnih komponenti, λ_j svojstvena vrednost k glavne komponente, n broj serija podataka; $\bar{\lambda}$ se računa prema formuli (26).

$$\bar{\lambda} = \sum_{j=k+1}^p \frac{\lambda_j}{p-k} \quad (26)$$

Nulta hipoteza ovog testiranja je da su sve promenljive u dobijenoj korelacionoj matrici međusobno nekorelisane, i ukoliko je značajnost testa (p) manja od 0,0001, hipoteza se prihvata, pa se može nastaviti sa primenom PCA.

PCA se često koristi kao statistička tehnika u analizi zagađenja vazduha i životne sredine uopšte. Kod modelovanja neuronskim mrežama, PCA se koristi za smanjenje broja ulaznih promenljivih, kao i za njihovu transformaciju u nove promenljive, takozvane glavne komponente, koje predstavljaju linearnu kombinaciju originalnih promenljivih, uz očuvanje maksimalno moguće varijanse originalnih podataka [80].

Geiger i *Kubin* [139] su analizirali pojavu relativnog gubitka informacija primenom PCA za redukciju dimenzionalnosti podataka (redukcija broja ulaza modela) i zaključili da primena PCA ne dovodi do značajnog gubitka informacija koje nose ori-

ginalni podaci. Do značajnog gubitka informacija ne dolazi ni u slučaju kada se PCA koristi za transformisanje ulaza rotacijom bez smanjivanja dimenzionalnosti, a takođe je primećeno da relativni gubitak informacija do koga dolazi opada sa porastom broja serija podataka na koje se PCA primenjuje [139].

Više detalja o primeni PCA za pripremu podataka kod ANN modelovanja emisije gasova staklene bašte biće prikazano u Poglavlju 6., dok se primeri primene PCA za modelovanje ostalih zagađujućih materija (koncentracije ozona, NO, NO₂, PM₁₀ i PM_{2,5}, generisanja čvrstog otpada itd.) mogu pronaći u relevantnoj literaturi [15, 38, 80, 84, 85].

4.3. Genetski algoritmi (GA)

Koncept genetskih algoritama (GA) zasnovan je na analogiji sa Darwinovom teorijom evolucije bioloških sistema. U skladu sa Darwinovom teorijom i GA je baziran na borbi različitih jedinki (rešenja) da „prežive“ sledeću generaciju (iteraciju), pri čemu se ukrštanjem (*crossover*), razmenom genetskog materijala, i mutacijom omogućava dobijanje boljih jedinki (rešenja) u svakoj narednoj generaciji (iteraciji). Za tvorca genetskih algoritama najčešće se uzima Holand (*J. H. Holland*) koji ih je teorijski definisao u knjizi „Adaptacija u prirodnim i veštačkim sistemima” iz 1975. godine [140].

GA se sastoje od tri osnovne genetske operacije koje se izvršavaju u svakoj iteraciji: selekcije, ukrštanja i mutacije (nalik na biološke sisteme). Navedene operacije se koriste da bi se početno rešenje problema modifikovalo tokom iteracija, čime se dolazi do optimalnog rešenja problema. Za razliku od konvencionalnih - matematičkih iterativnih metoda, genetski algoritmi simultano analiziraju više rešenja, čime se brže dolazi do konvergencije na globalnom nivou, pri čemu se smanjuje verovatnoća da se iterativni proces zaustavi na nekom od lokalnih minimuma, ne dajući pritom najbolje rešenje problema [141].

Jedan od glavnih parametara genetskih algoritama, kao i kod klasičnih optimizacionih tehnika, je kvantifikator tačnosti/ispravnosti dobijenog rešenja posle svake iteracije. U slučaju GA, ispravnost rešenja se proverava takozvanom funkcijom gre-

ške (*fitness* funkcija). Funkcija greške se definiše od slučaja do slučaja i može biti neka fizička, ekonomska ili statistička mera. Operacija selekcije se bazira na vrednosti funkcije greške i vrši kako bi se obezbedilo da u narednoj iteraciji broj rešenja bude srazmeran sa njihovom ispravnošću. Sledeća operacija u iteraciji je ukrštanje koje je ujedno i najvažnija operacija genetskih algoritama. Ukrštanje predstavlja razmenu genetskog materijala (parametara) između jedinki (rešenja), pri čemu se dolazi do novih, poboljšanih rešenja. Dok se ukrštanje vrši kombinovanjem više različitih rešenja, operacija mutacije se vrši na pojedinačnom rešenju, pri čemu se na slučajan način menjaju parametri rešenja. Iako se mutacijom mogu dobiti rešenja koja su manje tačna/ispravna u odnosu na početno rešenje, ona je neophodna, jer se na taj način sprečava da se evolucija završi suviše brzo (rana konvergencija). Naime, ranom konvergencijom se favorizuje postizanje rešenja koja predstavljaju lokalni minimum, a ne globalni, koji predstavlja optimalno rešenje. Detaljna struktura standardnog genetskog algoritma prikazana je na Slici 15.

```

Begin (1)
  t = 0 [start with an initial time]
  Initialize Population P(t) [initialize a usually random population of individuals]
  Evaluate fitness of Population P(t) [evaluate fitness of all individuals in
                                     population]
  While (Generations < Total Number) do begin (2)
    t = t + 1 [increase the time counter]
    Select Population P(t) out of Population P(t-1) [select sub-population for
                                                    offspring production]

    Apply Crossover on Population P(t)
    Apply Mutation on Population P(t)
    Evaluate fitness of Population P(t) [evaluate new fitness of population]
  end (2)
end (1)

```

Slika 15. Struktura standardnog genetskog algoritma [142]

Kao što se na Slici 15. može videti, proces optimizacije genetskim algoritmima počinje nasumičnim odabirom/kreiranjem početnog rešenja. Dalji proces optimizacije sledi navedene korake [143]:

1. Provera početnog rešenja odabranom funkcijom greške;

2. Na osnovu dobijene vrednosti funkcije greške, biraju se rešenja na kojima će biti vršene ostale operacije;
3. Vršiti se ukrštanje parametara različitih rešenja (krossover) pri čemu se dobijaju nova rešenja;
4. Zatim se vrši nasumična mutacija nad rešenjima dobijenim u prethodnom koraku;
5. Ponavljaju se koraci od drugog do četvrtog, pri čemu se vrši naredna iteracija i dobija se nova generacija rešenja;
6. Genetski algoritmi nemaju kriterijum za samozaustavljanje, već se završetak optimizacije bazira najčešće na dostizanju određenog, prethodno definisanog, broja iteracija.

Prednost GA kod optimizacionih procesa se upravo bazira na njihovom paralelnom/simultanom pretraživanju skupa mogućih rešenja, za razliku od klasične iterativne optimizacije kod koje se analizira jedno rešenje za drugim. Još jedna od prednosti je i brže pronalaženje globalnog minimuma, bez opasnosti da se optimizacija zaustavi na nekom od lokalnih minimuma. Takođe, GA nisu ograničene fenomenom diferencijabilnosti, poput ostalih matematičkih tehnika [144].

Genetski algoritmi se često koriste kao tehnika za odabir ulaza kod linearnih i nelinearnih modela, čime se eliminišu promenljive koje nisu značajne, a time se smanjuje i greška koju navedeni beznačajni ulazi unose u model [145]. Kod neuronskih mreža, kao funkcija greške najčešće se koristi srednja kvadratna greška (*mean squared error* - MSE) dobijenih vrednosti izlazne promenljive za definisanu validacionu grupu podataka [130].

Eksperimentalni deo

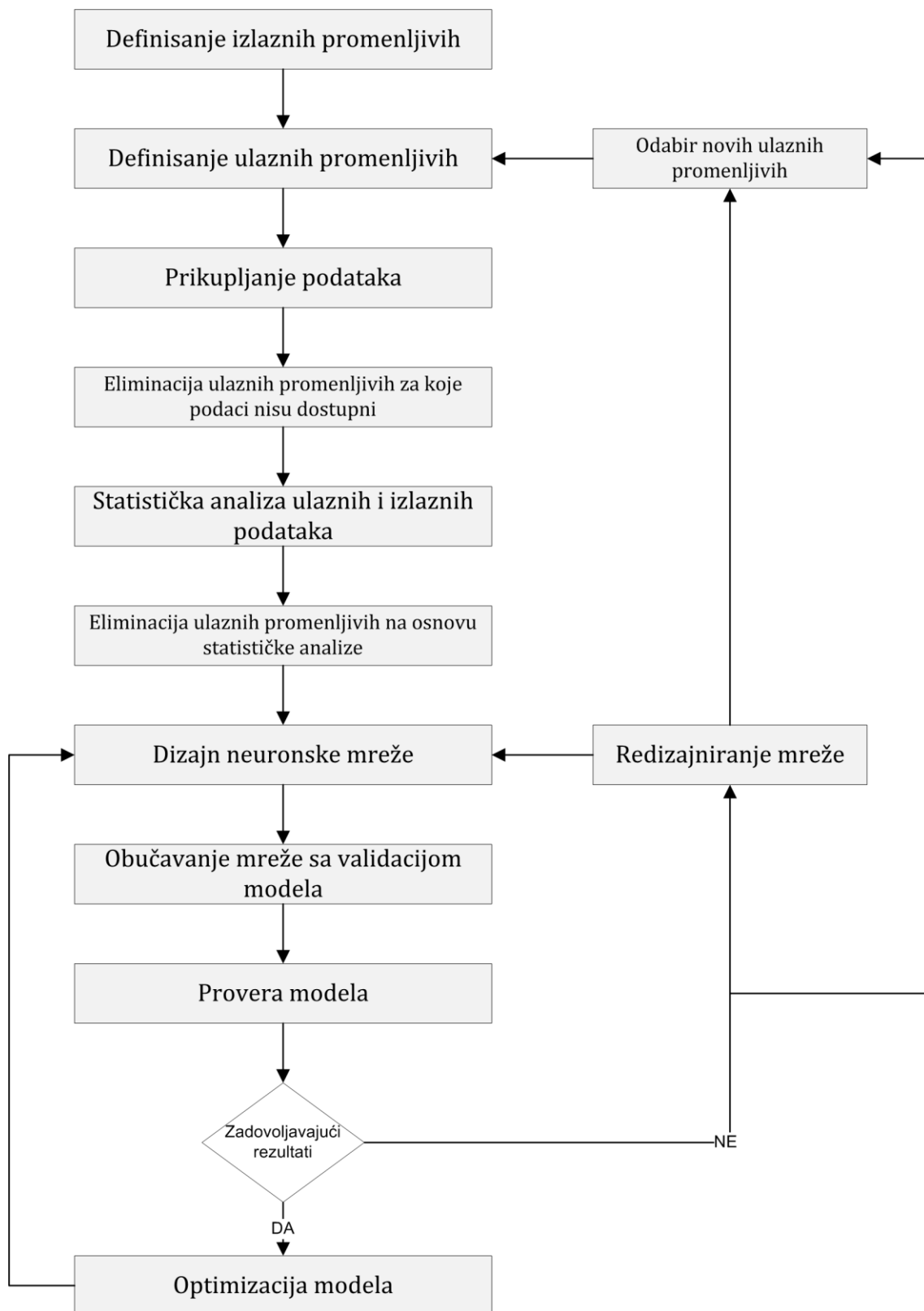
5. Metodologija razvoja modela zasnovanog na neuronskim mrežama

5.1. Proces razvoja modela zasnovanog na neuronskim mrežama

Proces razvoja neuronske mreže se ne može u potpunosti algoritamski definisati. U nekim fazama razvoja, posebno pri vođenju obučavanja i izboru i definisanju arhitekture, potrebno je osloniti se na empirijska pravila i metodu probe i greške [86]. Da bi se proces razvoja neuronske mreže učinio efikasnijim, pogodno ga je podeliti definisanjem faza u razvoju modela. Metodologija razvoja modela zasnovanog na neuronskim mrežama, koja je primenjena za modelovanje indikatora kvaliteta životne sredine, algoritamski je prikazana na Slici 16.

Prvi korak u razvoju modela zasnovanog na neuronskim mrežama je definisanje izlazne promenljive modela. Kako se pojedini aspekti kvaliteta životne sredine mogu pratiti različitim pristupima i indikatorima, u ovom koraku je važno definisati izlaznu promenljivu čije predviđanje omogućava kvantifikovanje ili opisivanje što konkretnijeg problema. Jednom kada se definiše izlazna promenljiva, ili više njih, moguće je definisanje ulaznih promenljivih koje utiču na definisanu izlaznu promenljivu modela. Ovaj korak je od ključnog značaja za uspešan razvoj modela, pri čemu je važno da se neuronskoj mreži prikaže što više promenljivih (po mogućstvu sve) koje imaju uticaj na odabranu izlaznu promenljivu, s obzirom da neuronska mreža ima sposobnost da uoči zakonitosti i veze među ovim promenljivima, koje kasnije koristi za uopštavanje.

U sledećoj fazi vrši se prikupljanje podataka definisanih ulaznim i izlaznim veličinama. Zatim se vrši eliminacija primarno definisanih ulaznih promenljivih, za koje odgovarajući podaci nisu dostupni ili poznati. Najvažnije pravilo u ovoj fazi je da je potrebno ceo domen razmatranog problema „pokriti” primerima, što se kod kompleksnih problema, sa velikim brojem ulaznih promenljivih, svodi na složeni zadatak ravnomernog pokrivanja N-dimenzionog prostora [86].



Slika 16. Šematski prikaz metodologije razvoja neuronske mreže

Jedno od često korišćenih empirijskih pravila je i da broj serija podataka za obučavanje neuronske mreže mora biti bar deset puta veći od broja ulaznih promenljivih. U ovoj fazi, često je potrebno rešavati i problem nekompletnih podataka. Mogućnost rada sa nekompletnim podacima je jedna od najvažnijih prednosti neuronskih mreža, kada su već oformljene. Čak je i u procesu obučavanja moguće koristiti nepotpune ulazne podatke, pri čemu neuronska mreža ima određene opcije za rad sa njima. Naravno, bolje je, ukoliko je moguće, podatke koji nedostaju upotpuniti dobro pretpostavljenim (na primer interpolisanim) vrednostima [86].

Statistička analiza podataka omogućava smanjenje broja primarno definisanih ulaznih veličina, kroz eliminaciju beznačajnih ili korelisanih ulaznih promenljivih, kao i formiranje odnosa između promenljivih, što za posledicu ima efikasniji (brži) proces obučavanja, uz kreiranje modela sa boljim karakteristikama. Dodatno se mogu vršiti različite vrste normalizacija, kojima se uprošćava prostor u kome neuronska mreža radi i time se ubrzava obučavanje mreže, kao i kroz logaritmovanje promenljivih kod kojih se očekuju razlike u redovima veličine, čime se dodatno poboljšava osetljivost mreže.

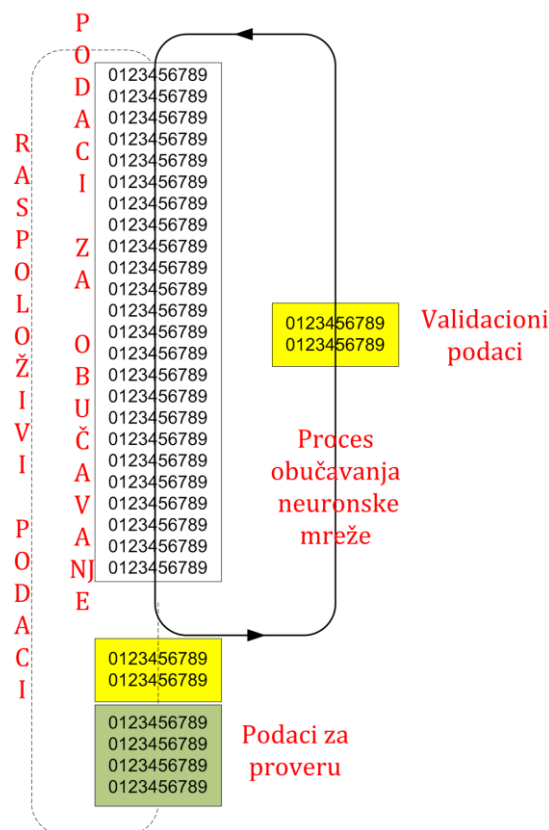
Razvoj neuronske mreže obuhvata i definisanje arhitekture i podešavanje odgovarajućih parametara odabrane arhitekture (npr. broj slojeva i neurona, povratne veze i drugo), kao i definisanje odgovarajućih parametara obučavanja mreže (brzina obučavanja, impuls (*momentum*) i faktor ujednačavanja).

Neka empirijska pravila vezana za izbor arhitekture i područje primene neuronskih mreža su [104]:

- troslojne mreže sa propagacijom greške unazad (BPNN) su verovatno najefikasnije za većinu problema (koriste se u 95 % slučajeva). Vreme obučavanja im je znatno kraće u odnosu na četvoroslojne i petoslojne (približno za red veličine po sloju), pri čemu imaju zadovoljavajući kapacitetu upštavanja;
- rekurentne mreže (RNN) su veoma dobre za vremenske serije podataka;
- probabilističke mreže (PNN) su pogodne za klasifikaciju podataka;

- neuronske mreže sa opštom regresijom (GRNN) veoma kratko i efikasno uče, i mogu da se obučavaju sa oskudnim skupom podataka. Za razliku od PNN, one daju kontinuirani izlaz i pokazale su se korisnim za aproksimiranje kontinualnih funkcija;
- Kohonenove mreže su našle primenu u grupisanju i klasifikaciji podataka.

Obučavanje neuronske mreže vrši se na osnovu prethodno definisane arhitekture – algoritma i parametara. Tokom samog procesa potrebno je dostupne serije podataka podeliti na skup podataka koji će biti korišćen za obučavanje mreže i skup podataka za proveru razvijenog modela (10-15 % od svih podataka). Takođe iz definisanog skupa za obučavanje poželjno je odgovarajućim algoritmom izdvojiti podatke koji će služiti za validaciju rezultata mreže tokom samog obučavanja (10-20 % podataka) (Slika 17.).



Slika 17. Podela dostupnih podataka tokom obučavanja mreže

Provera modela predstavlja simulaciju njegove primene u realnim uslovima, tj. predviđanje vrednosti izlazne promenljive na osnovu do tada mreži nepoznatih ulaznih podataka, definisanih kroz poseban skup podataka za proveru modela. Kao što je već navedeno, provera modela se vrši podacima (ulaznim i izlaznim) koji nisu korišćeni u procesu obučavanja. Provera modela obavlja se upoređivanjem rezultata koje daje neuronska mreža i stvarnih rezultata različitim statističkim pokazateljima performansi modela.

U slučaju da dobijeni rezultati tokom provere modela nisu zadovoljavajući, potrebno je podvrgnuti analizi celokupan proces razvoja modela. Izmene razvijenog modela mogu se izvesti na više načina [104]:

- redefinisanjem ulaznih promenljivih, s obzirom da se nezadovoljavajući rezultati neuronske mreže najčešće dobijaju zbog pogrešnog izbora ili neadekvatne prezentacije ulaznih podataka;
- izborom druge arhitekture;
- povećanjem broja skrivenih neurona i/ili slojeva, čime se mreži daje više stepeni slobode, ali i (najčešće) povećava vreme obučavanja.
- izmenom načina odabira validacionih podataka u procesu obučavanja;
- promenom parametara mreže (stope obučavanja, impulsa ili faktora ujednačavanja);
- smanjenjem veličine neuronske mreže, razčlanjivanjem problema koji se razmatra na nekoliko delova, pri čemu se zatim za svaki od njih razvija zasebna neuronska mreža.

U trenutku kada se modelom dobiju zadovoljavajući rezultati, onda se može pristupiti poslednjoj fazi razvoja modela i izvršiti optimizacija parametara razvijenog modela u cilju dodatnog poboljšanja njegove tačnosti.

5.2. Prikupljanje relevantnih ulaznih promenljivih modela

Sveobuhvatno i sistematično pretraživanje baza podataka i prikupljanje relevantnih ulaznih promenljivih predstavlja osnov za razvoj kvalitetnih modela baziranih na neuronskim mrežama. Kako je u okviru ove teze vršeno modelovanje dva tipa indikatora kvaliteta životne sredine, na nacionalnom i lokalnom nivou, pretraživanje je vršeno u dva smera:

- kod modelovanja emisije gasova staklene bašte, emisije i koncentracije čestica u atmosferi i količine generisanog komunalnog čvrstog otpada, vršeno je pretraživanje međunarodnih baza podataka (Eurostat, Svetska banka i dr.) koje sadrže indikatore vezane za stepen ekonomskog i industrijskog razvoja evropskih zemalja,
- kod modelovanja sadržaja rastvorenog kiseonika vršena je pretraga podataka koji su dobijeni redovnim monitoringom kvaliteta reke Dunav na teritoriji Republike Srbije.

Za modelovanja sadržaja rastvorenog kiseonika korišćeni su podaci iz Hidroloških godišnjaka Hidrometerološkog zavoda Republike Srbije [146-150] koji sadrže vrednosti fizičko-hemijskih, saprobioloških, bakterioloških i bioloških parametara kvaliteta površinskih voda: organoleptičkih osobina, kiselosti, alkalnosti, kiseoničnog režima, mineralizacije, biogenih sadržaja, sadržaja opasnih i štetnih materija (fenola, anjonaktivnih deterdženata, mineralnih ulja, teških metala, organohlornih pesticida i herbicida), nivoa radioaktivnosti, indeks saprobnosti po *Pantle-Bucku*, hlorofil a, najverovatniji broj koliformnih bakterija i ukupan broj aerobnih mezofilnih bakterija, taksonomske liste sa saprobnim vrednostima indikatorskih organizama, relativnu abundancu taksona i rezultate saprobioloških analiza po dva parametra (indeksu saprobnosti po *Pantle-Bucku* i stepenu saprobnosti po *Liebmannu*). Vrednosti navedenih parametara kvaliteta površinskih voda dobijeni su primenom različitih instrumentalnih tehnika, između ostalih atomskom apsorpcijom spektroskopijom - AAS, atomskom emisionom spektroskopijom - AES, jonском hromatografijom - IC, ultra-ljubičastom i vidljivom spektrofotometrijom - UV-VIS, indukovano spregnutom plazmom sa atomskom emisionom spektroskopijom - ICP-OES, indukovano spregnutom plazmom sa masenom spektroskopijom -

ICP-MS, kao i klasičnim tehnikama, volumetrija - Morov postupak, gasne analize – Vinklerov postupak, u periodu od 2004. do 2009. godine, u sklopu redovnog monitoringa kvaliteta površinskih voda na teritoriji Republike Srbije. Određivanje fizičko-hemijskih parametara kvaliteta voda vršeno je u skladu sa važećim domaćim i međunarodnim standardima (ISO, SRPS ISO, EPA i APHA AWWA WEF).

5.3. Statistički pokazatelji performansi modela

Određivanje performansi modela predstavlja složen postupak kojim se ispituje sposobnost razvijenog modela da da tačne prognoze u realnim uslovima. Kod neuronskih mreža realni uslovi se simuliraju primenom modela na nepoznate podatke, tj. podatke koji nisu bili korišćeni prilikom obučavanja mreže. Do danas ne postoji definisan set statističkih pokazatelja performansi, kao ni procedura kojom bi se definisao postupak obučavanja. Različiti autori koriste različite grupe statističkih pokazatelja koje primenjuju od slučaja do slučaja, pri čemu je jedino pravilo da pri proveru modela treba koristiti više statističkih pokazatelja koji opisuju različite aspekte ponašanja modela.

Tokom razvoja ANN modela, analiza i ocena performansi modela vršena je statističkim pokazateljima performansi modela kroz:

- poređenje rezultata ANN i konvencionalnih modela,
- poređenje rezultata razvijenih modela sa aktuelnim (izmerenim) vrednostima modelovanih indikatora održivog razvoja.

Kao reprezentativni statistički pokazatelji performansi modela odabrani su neki od najčešće korišćenih statističkih pokazatelja kod razvoja ANN modela u oblasti zaštite životne sredine [35, 151-155]:

- koeficijent linearne korelacije (R^2) između predviđenih i tačnih vrednosti,
- indeks slaganja (*the index of agreement - IA*)

$$IA = 1 - \frac{\overline{(C_p - C_o)^2}}{\left[\overline{|C_p - \bar{C}_o|} + \overline{|C_o - \bar{C}_o|} \right]^2} \quad (27)$$

- srednja apsolutna greška (*the mean absolute error - MAE*)

$$MAE = \frac{1}{n} \sum |C_p - C_o| \quad (28)$$

- koren srednje kvadratne greške (*the root mean squared error - RMSE*)

$$RMSE = \left[\frac{1}{n} \sum (C_p - C_o)^2 \right]^{1/2} \quad (29)$$

- srednja relativna greška (*the mean absolute percentage error - MAPE*)

$$MAPE = 100 \frac{1}{k} \sum \frac{|C_o - C_p|}{C_o} \quad (30)$$

- udeo tačnih predviđanja (*the percent of predictions within a factor the observed values - FA*):

- 1:1,1 u odnosu na tačnu vrednost (FA1.1) $0,9 < \frac{C_p}{C_o} < 1,1$ (31)

- 1:1,25 u odnosu na tačnu vrednost (FA1.25)

$$0,8 < \frac{C_p}{C_o} < 1,25 \quad (32)$$

Kako svaki od navedenih statističkih pokazatelja ima nedostatke i ograničenja u pogledu informacija koje pruža o performansama modela, neophodno je vršiti višestruku statističku analizu performansi modela [156]. Takođe, ni pojedinačno gledano značaj svakog od pokazatelja nije isti.

Koeficijent linearne korelacije (R^2) daje najmanje informacija, a time je i najmanje značajan, od prikazanih statističkih pokazatelja, jer on predstavlja stepen linearne korelacije predviđenih i tačnih vrednosti, ali se veoma često koristi u međufazama razvoja ANN modela zbog jednostavnog i brzog izračunavanja od strane softverskih paketa koji se koriste za crtanje dijagrama.

Indeks slaganja (IA) je veoma važan indikator performansi, jer on daje stepen srednjeg slaganja tačnih i predviđenih vrednosti, uključujući grešku između njih.

To je relativan i normiran statistički indikator performansi sa vrednostima od 0 (loše slaganje) do 1 (dobro slaganje).

MAE i *RMSE* su odabrani zato što se najčešće koriste za kvantifikaciju rezidualne greške, pokazujući apsolutnu grešku između predviđenih i izmerenih vrednosti. Kako *RMSE* predstavlja „kvadratnu“ grešku, on je bolji pokazatelj velikih (ekstremnih) grešaka modela, u odnosu na *MAE* [38].

MAPE pokazuje srednju vrednost relativne greške koju napravi model primenjen na određen broj slučajeva, i kao takav veoma je značajan kod razvoja ANN modela, jer u slučaju modelovanja nacionalnih vrednosti, može biti indikator za ocenu performansi modela za svaku od država (regiona) obuhvaćenih modelom.

FA indikator, pored *IA*, jedan je od najznačajnijih statističkih pokazatelja performansi modela, jer on pokazuje sposobnost modela da da tačna predviđanja za određenu rezoluciju greške, za svaki pojedinačni ispitivani slučaj. Osobina *FA* pokazatelja da može imati različite rezolucije greške/tačnosti (npr. *FA1.1* pokazuje udeo predviđanja sa relativnom greškom manjom od $\pm 10\%$, dok *FA1.25* pokazuje udeo predviđanja sa relativnom greškom manjom od $+25\%$ i -20%) omogućava da se modeli efikasno klasifikuju po performansama, u zavisnosti od složenosti problema i zahtevane tačnosti.

Rezultati i Diskusija

6. Modelovanje indikatora kvaliteta životne sredine

6.1. Modelovanje emisije gasova staklene bašte

U ovom poglavlju opisan je razvoj, optimizacija, evaluacija i simulacija ANN modela za predviđanje emisije gasova staklene bašte na nacionalnom nivou, sa odabranim ekonomskim pokazateljima i pokazateljima održivog razvoja, kao ulaznim promenljivim. Odabir najznačajnijih ulaznih promenljivih je vršen različitim hemometrijskim metodama: korelacionom analizom i analizom glavnih komponenti. Nakon kreiranja odgovarajućeg GRNN modela, razvijeni su i modeli zasnovani na višestrukoj linearnoj regresiji (MLR) i regresiji glavnih komponenti (PCR), radi poređenja sa rezultatima GRNN modela. Mogućnost primene GRNN modela u praksi ispitana je predviđanjem emisije GHG gasova za 2011. godinu. Navedenim ispitivanjem došlo se je zaključaka da su glavne prednosti u primeni ANN modela, u odnosu na linearne modele (inventari emisije, GAINS model i sl.): potreba za manjim brojem ulaznih promenljivih i činjenica da vrsta i broj ulaznih promenljivih kod ANN modela mogu biti prilagođeni stanju u pojedinačnim zemljama. Na taj način, ANN metodologija može biti primenjena u skladu sa dostupnim podacima, pri čemu za svaku pojedinačnu zemlju može biti razvijen odgovarajući model, bez potrebe da se vrše aproksimacije vrednosti ulaznih veličina i time unose greške u predviđanjima, što bi bio slučaj kad bi se predviđanja vršila odgovarajućim linearnim modelima [157].

6.1.1. Ulazne promenljive i priprema podataka za GHG model

Za uspešno predviđanje emisije gasova staklene bašte neophodno je za ulazne promenljive modela odabrati dostupne indikatore održivog razvoja, koji kvantifikuju sve sektore njihove emisije: proizvodnja i prenos energije, saobraćaj, poljoprivreda, industrija i upravljanje otpadom.

Na Slici 18. prikazani su udeli navedenih sektora u ukupnoj emisiji gasova staklene bašte u Evropi [158], kao i odabrani ekonomski, industrijski i indikatori održivog razvoja koji su korišćeni kao ulazni parametri modela za predviđanje emisije gasova staklene bašte.

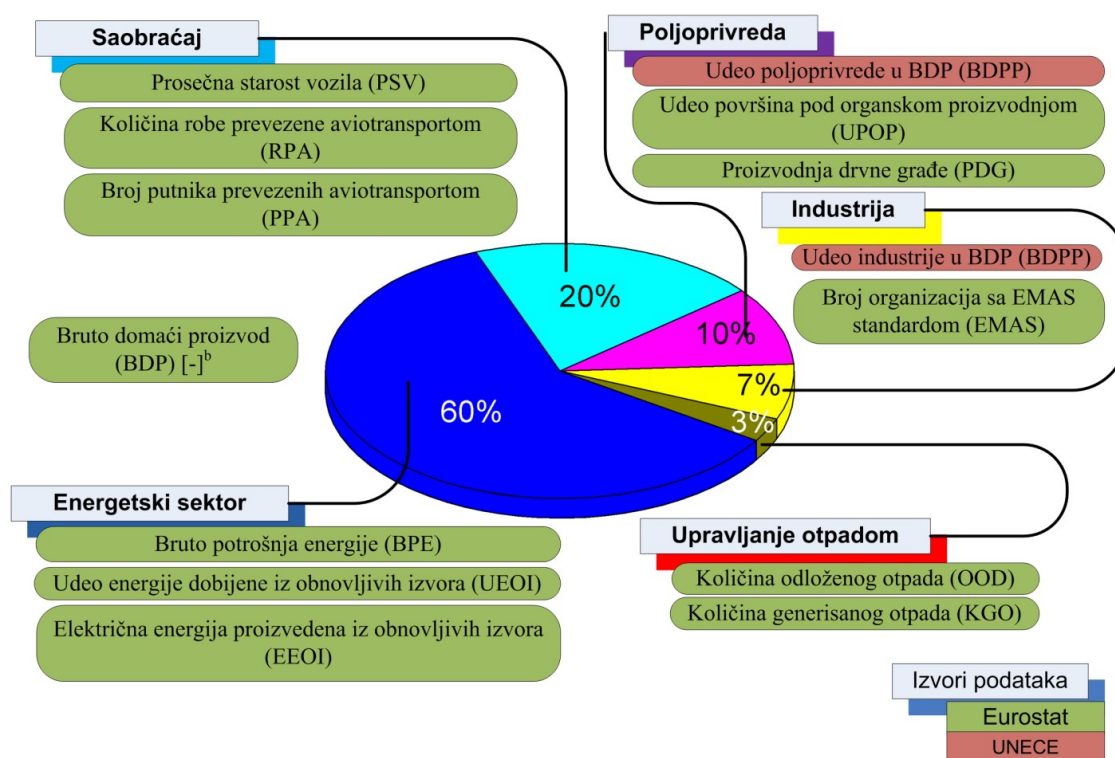
Za obučavanje, validaciju i proveru modela zasnovanog na neuronskim mrežama korišćeni su podaci za 28 evropskih zemlja za period od 2004. do 2010. godine. Odabrane ulazne promenljive su normalizovane po stanovniku i/ili sa vrednošću BDP-a Evropske unije kako bi se omogućilo poređenje država različite veličine. Vrednosti navedenih ulaznih promenljivih i podaci o emisiji gasova staklene bašte su preuzeti iz baza podataka Eurostata [4] i Komisije Ujedinjenih Nacija za Evropu (*United Nations Economic Commission for Europe - UNECE*) [159].

U slučaju indikatora prosečna starost vozila (PSV), koje Eurostat prati kao broj vozila u različitim starosnim intervalima, vršeno je i statističko svodenje više Eurostatovih izveštaja na jedan indikator.

Eurostat daje izveštaje o starosti vozila u sledećim starosnim indikatorima: broj vozila mlađih od 2 godine (N_{0-2}), broj vozila starosti od 2 do 5 godina (N_{2-5}), broj vozila starosti od 5 do 10 godina (N_{5-10}), broj vozila starijih od 10 godina (N_{10+}). Svodenje navedena četiri izveštaja na jedan indikator, starost vozila, vršeno je pod pretpostavkom da je prosečna starost vozila u svakom od zatvorenih intervala jednak srednjoj vrednosti tih starosnih intervala, dok je u slučaju otvorenog intervala (N_{10+}) pretpostavljeno da je većina vozila starosti od 10 do 15 godina, a kako postoji i određen broj vozila starijih od 15 godina, prosečna starost vozila je procenjena na 13 godina.

U skladu sa navedenim pretpostavkama, indikator prosečna starost vozila (PSV), koji je korišćen kao ulazna promenljiva, proračunat je na osnovu jednačine (33):

$$PSV = \frac{1 \cdot N_{0-2} + 3,5 \cdot N_{2-5} + 7,5 \cdot N_{5-10} + 13 \cdot N_{10+}}{N_{0-2} + N_{2-5} + N_{5-10} + N_{10+}} \quad (33)$$



Slika 18. Sektori GHG emisije i odabrani indikatori [157]

Dostupni podaci za ulazne promenljive, prikazane na Slici 18., su podeljeni u dve grupe: podaci od 2004. do 2009. godine su korišćeni za obučavanje i podaci iz 2010. godine, koji su korišćeni za proveru modela. Validacioni podaci su izdvojeni kao podskup podataka za obučavanje, tokom optimizacije parametara modela. Statistička analiza ulaznih podataka data je u Tabeli 1., dok je statistička analiza emisije gasova staklene bašte data u Tabeli 2. Da bi se proverom utvrdile realne performanse modela neophodno je da srednje vrednosti istih ulaznih promenljivih imaju slične vrednosti u obe grupe podataka (Tabele 1. i 2.). Takođe je poželjno da se i ekstremne vrednosti podataka kod provere modela (minimumi i maksimumi) nalaze u okviru definisane grupe podataka za obučavanje, što je u ovom slučaju ispunjeno za svaki ulazni indikator (Tabela 1.).

Tabela 1. Lista ulaznih promenljivih GHG modela sa osnovnim statističkim pokazateljima

Ulazni indikator	Podaci za obučavanje (2004-2009)			Podaci za proveru (2010)		
	Minimalna vrednost	Maksimalna vrednost	Srednja vrednost	Minimalna vrednost	Maksimalna vrednost	Srednja vrednost
Bruto domaći proizvod ^a (BDP)	0,341	2,740	1,008	0,437	2,661	1,013
Bruto potrošnja energije (BPE) [toe ^b po stan.]	1,653	10,34	3,906	1,663	9,188	3,841
Električna energija proizvedena iz obnovljivih izvora (EEOI)	0,000	1,094	0,192	0,000	0,900	0,232
Udeo energije dobijene iz obnovljivih izvora (UEOI)	0,001	0,651	0,142	0,004	0,611	0,168
Prosečna starost vozila (PSV) [god.]	5,290	12,39	8,367	6,327	11,97	8,682
Količina robe prevezene aviotransportom (RPA) [t po stan.]	0,000	1,613	0,070	0,001	1,392	0,072
Broj putnika prevezenih aviotransportom (PPA) [putnika po stan.]	0,147	9,101	2,546	0,348	8,376	2,655
Udeo poljoprivrede u BDP (BDPP)	0,003	0,141	0,028	0,003	0,066	0,025
Udeo površina pod organskom proizvodnjom (UPOP)	0,000	0,185	0,047	0,002	0,172	0,060
Proizvodnja drvne građe (PDG) [t po stan]	0,000	10,88	1,639	0,000	9,500	1,707
Udeo industrije u BDP (BDPP)	0,069	0,399	0,214	0,074	0,351	0,206
Broj organizacija sa EMAS ^c standardom (EMAS)	0,000	1641	133,4	0,000	1402	162,2
Količina odloženog otpada na deponije (OOD) [t po sta.]	0,002	0,675	0,252	0,002	0,590	0,226
Količina generisanog otpada (KGO) [t po stan.]	0,256	0,794	0,506	0,303	0,689	0,488

^anormalizovan po stanovniku i vrednosti BDP-a Evropske Unije

^btona ekvivalentne nafte

^cEMAS - *Eco-Management and Auditing Scheme*

Tabela 2. Statistička analiza (minimalne, maksimalne i srednje vrednosti) GHG emisije u državama Evropske unije i po grupama podataka

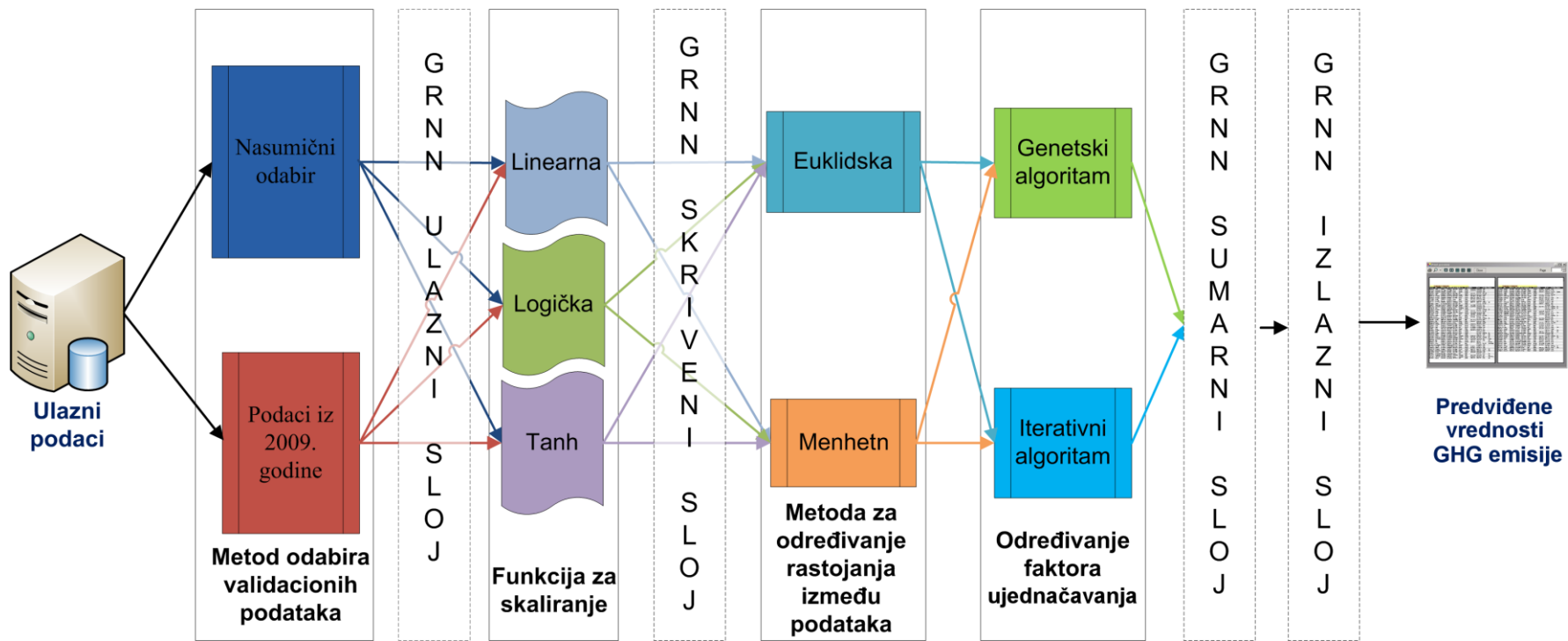
Država	GHG emisija [t po stan.]		
	Minimalna vrednost	Maksimalna vrednost	Srednja vrednost
Belgija	11,60	14,15	12,87
Bugarska	7,76	9,26	8,59
Češka	12,85	14,46	13,91
Danska	10,99	13,17	11,91
Nemačka	11,13	12,35	11,85
Estonija	12,23	15,75	14,20
Irska	13,57	16,72	15,42
Grčka	10,46	12,22	11,61
Španija	7,72	10,03	9,12
Francuska	7,98	9,01	8,52
Italija	8,17	9,92	9,17
Kipar	13,07	15,09	14,33
Letonija	4,81	5,59	5,11
Litvanija	5,98	7,54	6,75
Luksemburg	23,13	27,84	25,67
Mađarska	6,67	7,88	7,38
Malta	7,29	7,64	7,42
Holandija	12,03	13,32	12,65
Austrija	9,53	11,29	10,56
Poljska	10,01	10,68	10,37
Portugal	6,66	8,20	7,48
Rumunija	5,66	7,08	6,59
Slovenija	9,53	10,60	10,06
Slovačka	8,16	9,63	9,08
Finska	12,38	15,39	13,99
Švedska	6,42	7,79	7,17
Velika Britanija	9,26	11,01	10,29
Norveška	10,66	11,95	11,40
Grupa podataka za	Minimalna vrednost	Maksimalna vrednost	Srednja vrednost
Obučavanje	4,811	27,84	10,93
Proveru	5,587	23,81	10,26

6.1.2. Rezultati optimizacije parametara neuronske mreže

Posle odabira ulaznih indikatora i definisanja podataka za obučavanje i proveru, sledeći korak u razvoju je kreiranje samog modela, tj. definisanje parametara arhitekture i obučavanja neuronske mreže. Odabrana je ANN arhitektura sa opštom regresijom (GRNN) kojom su već razvijeni precizni modeli u oblasti zaštiti životne sredine [160]. Glavni parametri arhitekture mreže (broj slojeva i broj neurona po slojevima) kod GRNN mreže su definisani brojem ulaznih i izlaznih promenljivih i brojem serija podataka u grupi podataka za obučavanje. U konkretnom slučaju, kako se model kreira sa 14 ulaza i jednim izlazom i kako je dostupno ukupno 168 serija podataka (28 država x 6 godina), broj neurona po slojevima je:

- ulazni sloj - 14 neurona (jednak broju ulaza),
- skriveni sloj - 168 neurona (jednak broju serija podataka u trening datase-tu),
- sloj za sumiranje - 2 neurona (jednak broju izlaza plus 1),
- izlazni sloj - 1 neuron (jednak broju izlaza).

Ostali parametri arhitekture i obučavanja mreže, prikazani na Slici 19., su bili varirani tokom obučavanja, kako bi se dobili optimalni parametri za modelovanje emisije gasova staklene bašte.



Slika 19. Parametri arhitekture i obučavanja GRNN mreže varirani tokom procesa optimizacije

Kao što se može videti na Slici 19., ispitan je uticaj različitih tehnika za odabir validacionih podataka:

- nasumični odabir - 12 % nasumično odabranih serija podataka iz grupe podataka za obučavanje modela;
- podaci iz određene (poslednje) godine za obučavanje - može se primeniti pošto se radi o vremenskim serijama podataka.

Takođe su testirane i tri različite funkcije za skaliranje ulaznih podataka, koje ujedno predstavljaju i vezu između ulaznih i skrivenih neurona:

– linearna funkcija $f(x) = x$ (34)

– sigmoidna (logička) funkcija $f(x) = \frac{1}{1 + \exp(-x)}$ (35)

– hiperboličko-tangentna (tanh) funkcija $f(x) = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)}$ (36)

Testirane su i dve različite metode za određivanje međusobnog rastojanja podataka:

- Euklidski metod (*The Euclidean distance metric*) - rastojanje između dve tačke, a i b , sa j brojem dimenzija se određuje jednačinom

$$\sqrt{\sum_{i=1}^j (a_i - b_i)^2} \quad (37)$$

- Menhetn (L_1) metod (*The City block distance metric*) - rastojanje između dve tačke, a i b , sa j brojem dimenzija, određuje se jednačinom

$$\sum_{i=1}^j |a_i - b_i| \quad (38)$$

kao i dve metode za određivanje faktora ujednačavanja: genetski i iterativni algoritam.

Kako bi se ispitale sve kombinacije svih navedenih parametara mreže, razvijena su 24 GRNN modela sa svim raspoloživim ulazima (Tabela 3).

Tabela 3. Lista GRNN modela razvijenih tokom optimizacije parametara mreže

Oznaka modela	Metod odabira validacionih podataka	Funkcija za skaliranje	Metod za određivanje rastojanja podataka	Algoritam za određivanje faktora ujednačavanja
GRNN1	Nasumični odabir 12 % podataka za obučavanje	Linearna	Euklidska	Genetički Iterativni
GRNN2			Menhetn	Genetički Iterativni
GRNN3			Euklidska	Genetički Iterativni
GRNN4			Menhetn	Genetički Iterativni
GRNN5		Logička	Euklidska	Genetički Iterativni
GRNN6			Menhetn	Genetički Iterativni
GRNN7			Euklidska	Genetički Iterativni
GRNN8			Menhetn	Genetički Iterativni
GRNN9		Tanh	Euklidska	Genetički Iterativni
GRNN10			Menhetn	Genetički Iterativni
GRNN11			Euklidska	Genetički Iterativni
GRNN12			Menhetn	Genetički Iterativni
GRNN13	Podaci iz 2009. godine	Linearna	Euklidska	Genetički Iterativni
GRNN14			Menhetn	Genetički Iterativni
GRNN15			Euklidska	Genetički Iterativni
GRNN16			Menhetn	Genetički Iterativni
GRNN17		Logička	Euklidska	Genetički Iterativni
GRNN18			Menhetn	Genetički Iterativni
GRNN19			Euklidska	Genetički Iterativni
GRNN20			Menhetn	Genetički Iterativni
GRNN21		Tanh	Euklidska	Genetički Iterativni
GRNN22			Menhetn	Genetički Iterativni
GRNN23			Euklidska	Genetički Iterativni
GRNN24			Menhetn	Genetički Iterativni

Modeli su razvijeni softverskim paketom *Neuroshell 2* [161]. Kao što je već navedeno, svi testirani modeli su imali 14 ulaznih, 168 skrivenih, jedan izlazni neuron i dva neurona za sumiranje. Vrednosti statističkih pokazatelja performansi modela za različite parametre prikazani su u Tabeli 4.

Tabela 4. Vrednosti statističkih pokazatelja performansi za GRNN modele razvijene tokom optimizacije prema definisanim parametrima mreže (provera podacima iz 2010. godine)

GRNN parametar	Statistički indikator performansi modela			
	IA	FA1,1 [%]	MAE [t po stan.]	RMSE [t po stan.]
<i>Metod odabira validacionih podataka</i> <i>srednje vrednosti za 12 modela</i>				
Nasumični odabir¹	0,991	88,1	0,465	0,701
Podaci iz 2009. godine	0,985	68,9	0,753	0,930
<i>Funkcija za skaliranje</i> <i>srednje vrednosti za 8 modela</i>				
Linearna funkcija	0,987	78,5	0,611	0,812
Logička funkcija	0,989	78,6	0,615	0,805
Tanh funkcija	0,988	78,4	0,602	0,831
<i>Metode za određivanje rastojanja</i> <i>srednje vrednosti za 12 modela</i>				
Euklidska	0,988	79,4	0,607	0,813
Menhetn	0,988	77,6	0,611	0,818
<i>Određivanje faktora ujednačavanja</i> <i>srednje vrednosti za 12 modela</i>				
Genetički algoritam	0,989	80,5	0,595	0,787
Iterativni algoritam	0,987	76,5	0,624	0,845

¹bold slova označavaju optimalne parametre mreže

GRNN modeli koji su razvijeni sa nasumičnim odabirom validacionih podataka i kod kojih je faktor ujednačavanja određivan genetičkim algoritmom imali su znatno bolje rezultate prilikom testiranja sa nepoznatim podacima tokom provere modela, dok su ostali modeli, koji su imali različite funkcije za skaliranje i određivanje međusobnog rastojanja podataka, imali slične performanse (Tabela 4.).

Vrednosti statističkih pokazatelja performansi za svaki od razvijenih GRNN modela prikazani su u Tabeli 5. zajedno sa rezultatima MLR modela, koji je prikazan radi poređenja i čiji koeficijenti su dati u Tabeli 6.

Tabela 5. Vrednosti statističkih pokazatelja performansi za kreirane GRNN modele tokom procesa optimizacije parametara

Model	Statistički indikator performansi modela			
	<i>IA</i>	<i>FA1,1</i> [%]	<i>MAE</i> [t po stan.]	<i>RMSE</i> [t po stan.]
GRNN1	0,995	89	0,422	0,541
GRNN2	0,989	86	0,447	0,750
GRNN3	0,995	89	0,399	0,521
GRNN4	0,990	89	0,450	0,735
GRNN5	0,993	93	0,469	0,627
GRNN6	0,988	86	0,512	0,780
GRNN7	0,993	89	0,466	0,636
GRNN8	0,989	86	0,489	0,770
GRNN9	0,988	86	0,501	0,793
GRNN10	0,989	89	0,463	0,750
GRNN11	0,989	86	0,489	0,761
GRNN12	0,989	89	0,476	0,753
GRNN13	0,983	75	0,778	0,999
GRNN14	0,987	68	0,748	0,853
GRNN15	0,975	71	0,862	1,183
GRNN16	0,985	61	0,784	0,910
GRNN17	0,987	71	0,729	0,889
GRNN18	0,985	68	0,775	0,959
GRNN19	0,989	75	0,669	0,803
GRNN20	0,984	61	0,807	0,978
GRNN21	0,987	71	0,692	0,877
GRNN22	0,985	71	0,753	0,941
GRNN23	0,989	71	0,659	0,810
GRNN24	0,985	64	0,780	0,960
MLR	0,950	43	1,343	1,660

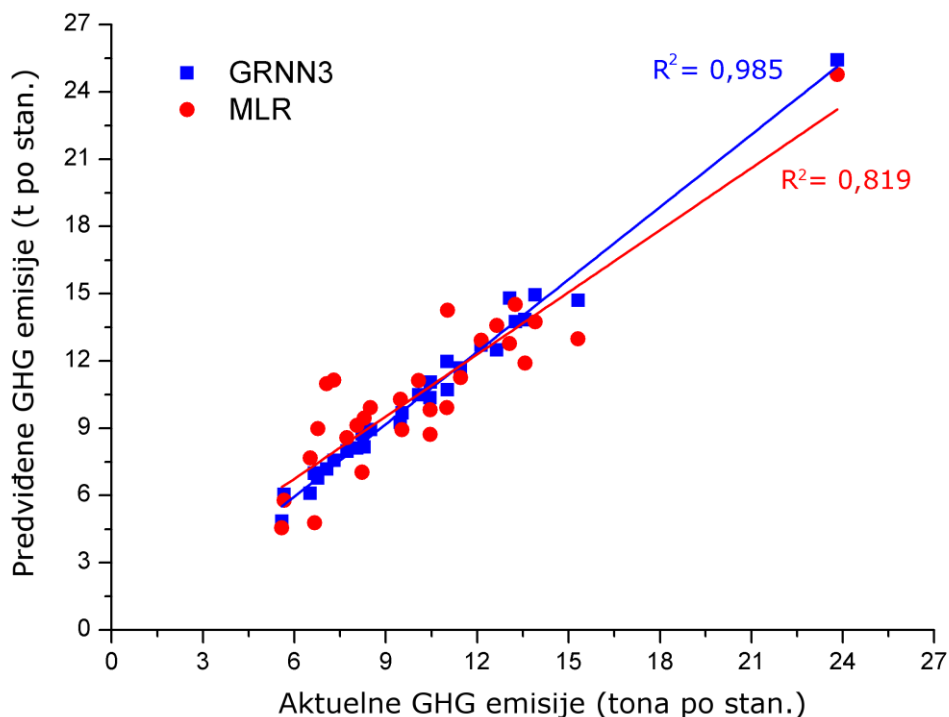
Tabela 6. Koeficijenti modela sa višestrukom linearnom regresijom (MLR model)

Promenljive modela	Koeficijenti	Standardna greška
Konstanta	-2,386	2,126
OOD	0,763	1,358
UEOI	-1,951	3,341
PDG	-0,253	0,105
EMAS	0,000	0,000
KGO	-2,407	2,085
BRE	1,391	0,254
BDP	3,121	1,420
EEOI	-8,351	1,951
UPOP	30,114	4,376
PPA	0,620	0,131
RPA	2,715	1,206
BDPP	48,872	11,368
BDPI	15,217	2,855
PSV	0,032	0,141

Kao što se može videti u Tabeli 5., GRNN model (sa oznakom GRNN3) koji je kreiran sa nasumično odabranim validacionim podacima, linearnom funkcijom, Euklidskom metodom i genetskim algoritmom, dao je najbolje rezultate kod predviđanja emisije gasova staklene bašte za evropske zemlje u 2010. godini. Jedino je GRNN5 model imao bolju vrednost za jedan od statističkih pokazatelja (*FA1.1*), u odnosu na GRNN3 model. To znači da je GRNN3 model imao samo jedno predviđanje sa greškom od 10 % više u odnosu na GRNN5 model. Kako je GRNN3 imao bolje vrednosti ostalih statističkih pokazatelja performansi, navedeni model se može uzeti za najbolje optimizovan model, čije se performanse mogu dalje poboljšati adekvatnim odabirom ulaznih promenljivih. Takođe se može primetiti (Tabele 3. i 5.) da su modeli sa *tanh* funkcijom dali tačnija predviđanja u kombinaciji sa iterativnim algoritmom, dok su modeli sa logičkom funkcijom dali tačnija predviđanja u kombinaciji sa genetskim algoritmom.

Poređenje GRNN3 modela sa rezultatima modela sa linearnom regresijom (MLR model), koji je kreiran i proveren sa istim podacima kao i GRNN3 model, pokazalo je da GRNN model ima značajno bolje performanse (Tabela 5.). MLR model je kreiran korišćenjem IBM softverskog paketa SPSS 19 [162]. Poređenje predviđanja

emisije gasova staklene bašte za 2010. godinu primenom GRNN3 i MLR modela sa aktuelnim vrednostima emisija prikazano je na Slici 20.



Slika 20. Poređenje aktuelnih i modelima predviđenih vrednosti GHG emisija u 2010. godini

6.1.3. Rezultati GHG emisionih modela razvijenih odabirom ulaznih promenljivih na osnovu korelacione analize

Korelaciona analiza se često koristi kao metod odabira i/ili redukcije ulaznih promenljivih koje nisu dovoljno značajne ili koje su korelisane sa ostalim ulaznim promenljivim, zbog čega unose zabunu tokom obučavanja mreže. Rezultati korelacione analize za ulazne promenljive koje su korišćene kod modelovanja emisije gasova staklene bašte prikazani su u Tabeli 7.

Tabela 7. Rezultati korelacione analize ulaznih promenljivih GHG modela

	OOD	UEOI	PDG	EMAS	KGO	BRE	BDP	EEOI	UPOP	PPA	RPA	BDPP	BDPI	PSV
OOD	1													
UEOI	-0,35	1												
PDG	-0,21	0,62	1											
EMAS	-0,24	-0,09	-0,13	1										
KGO	0,08	-0,25	-0,26	0,20	1									
BPE	-0,49	0,16	0,31	-0,03	0,30	1								
BDP	-0,41	0,10	-0,01	0,10	0,60	0,84¹	1							
EEOI	-0,40	0,91¹	0,38	0,01	-0,09	0,15	0,23	1						
UPOP	-0,37	0,43	0,41	0,18	-0,13	0,10	0,07	0,45	1					
PPA	0,34	-0,05	-0,16	-0,04	0,72	0,16	0,44	0,05	-0,18	1				
RPA	-0,16	-0,20	-0,10	-0,07	0,32	0,70	0,71	-0,16	-0,13	0,11	1			
BDPP	0,35	0,01	-0,02	-0,22	-0,46	-0,52	-0,67	-0,07	-0,25	-0,44	-0,27	1		
BDPI	-0,24	0,43	0,22	0,02	-0,45	-0,04	-0,19	0,42	0,11	-0,34	-0,39	0,20	1	
PSV	0,31	0,19	0,21	-0,26	-0,64	-0,48	-0,69	-0,03	-0,06	-0,31	-0,37	0,54	0,25	1
GHG	-0,15²	-0,27²	-0,08	-0,04	0,43²	0,77²	0,73²	-0,23²	-0,06	0,31²	0,74²	-0,41²	-0,18²	-0,47²

¹međusobno značajno povezanih ulaznih promenljivih

²ulazne promenljive značajno povezane sa GHG emisijom

OOD-Količina odloženog otpada na deponije, UEOI-Udeo energije dobijene iz obnovljivih izvora, PDG Proizvodnja drvne građe, EMAS-broj organizacija sa EMAS standardom, KGO-Količina generisanog otpada, BPE-Bruto potrošnja energije, BDP-Bruto domaći proizvod, EEOI-Električna energija proizvedena iz obnovljivih izvora, UPOP-udeo površina pod organskom proizvodnjom, PPA-Broj putnika prevezenih aviotransportom, RPA-Količina robe prevezene aviotransportom, BDPP- Udeo poljoprivrede u BDP, BDPI-Udeo industrije u BDP, PSV-Prosečna starost vozila

Kao što se može videti u Tabeli 7., za dva para ulaznih promenljivih dobijene su visoke vrednosti koeficijenata međusobne korelacije ($r > 0,80$):

- Bruto domaći proizvod (BDP) i Bruto potrošnja energije (BPE),
- Udeo energije iz obnovljivih izvora (UEOI) i Udeo električne energije proizvedene iz obnovljivih izvora (EEOI).

Međusobna korelacija navedenih indikatora je očekivana, pošto je dobro poznata međusobna povezanost BDP-a i potrošnje energije, kao i činjenica da se električna energija uglavnom dobija iz obnovljivih izvora. Na osnovu navedenih rezultata korelacione analize razvijena su 4 GRNN modela (označena sa IV1-IV4), sa međusobno značajno nekoreliranim ulaznim promenljivim ($r < 0,80$). Takođe je kreiran i GRNN model (CA-GRNN3) sa ulaznim promenljivim koje su značajno povezane sa emisijom gasova staklene bašte, sa koeficijentom korelacije većim od 0,10. U skladu sa tim, CA-GRNN3 model je razvijen bez sledećih ulaznih indikatora: proizvodnja drvne građe, udeo površina pod organskom proizvodnjom i broj organizacija sa EMAS standardom.

Vrednosti statističkih pokazatelja performansi za svih 5 GRNN modela, razvijenih tokom odabira ulaznih promenljivih korelacionom analizom, prikazane su u Tabeli 8. Svi GRNN modeli prikazani u Tabeli 8. imaju iste parametara arhitekture i obučavanja kao i GRNN3, koji se pokazao kao najbolje optimizovan model. Jedino je broj ulaznih neurona morao da se menja zato što su određene ulazne promenljive bile eliminisane.

CA-GRNN3 model je dao najbolje predikcije emisije gasova staklene bašte za 2010. godinu, u poređenju sa IV-GRNN modelima sa nezavisnim ulaznim promenljivim. Takođe ima i značajno bolje performanse u poređenju sa početnim GRNN3 modelom. Navedeni rezultati pokazuju da se eliminacijom ulaznih promenljivih, koje nisu značajnije povezane sa izlaznom promenljivom, mogu znatno poboljšati performanse GRNN modela. U Tabeli 8. su takođe date vrednosti statističkih pokazatelja performansi za MLR model (CA-MLR) koji je razvijen sa istim podacima kao i

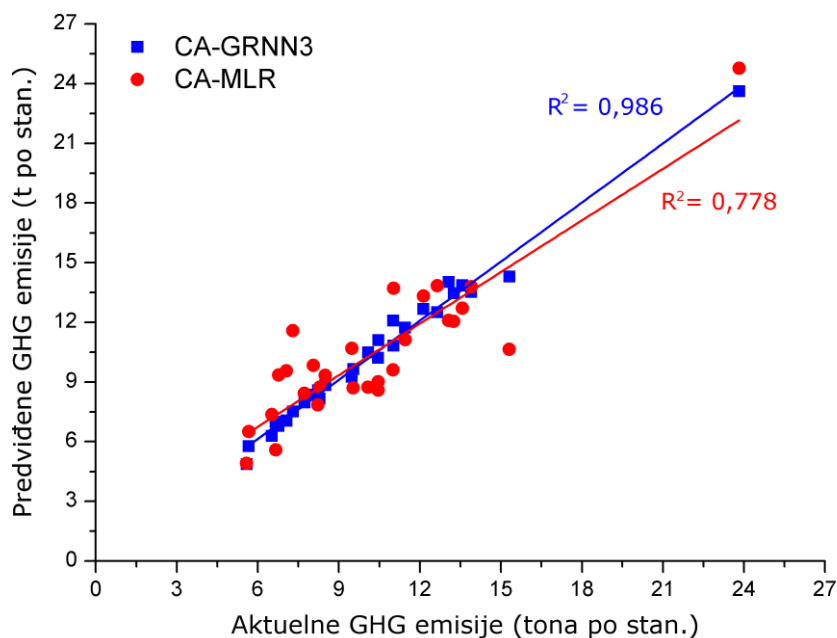
CA-GRNN3 model. Poređenjem rezultata CA-GRNN3 i CA-MLR (Slika 21.) može se konstatovati da je CA-GRNN3 model dao znatno tačnije predikcije vrednosti emisija gasova staklene bašte u 2010. godini.

Tabela 8. Vrednosti statističkih pokazatelja performansi modela za GRNN i MLR modele kreirane na osnovu rezultata korelacione analize

Model	Eliminirani ulazi	Statistički pokazatelji performansi modela			
		IA	FA1,1 [%]	MAE [t po stan.]	RMSE [t po stan.]
IV1-GRNN3 ¹	BDP i BPE	0,995	93	0,387	0,536
IV2-GRNN3	BDP i EEOI	0,994	96	0,433	0,557
IV3-GRNN3	BRE i UEIOI	0,989	86	0,498	0,750
IV4-GRNN3	UEIOI i EEOI	0,989	89	0,460	0,739
CA-GRNN3²	PDG, EMAS i UOPO	0,996	96	0,353	0,450
CA-MLR	PDG, EMAS i UOPO	0,940	46	1,401	1,759

¹IV1-IV4 modeli imaju po 12 ulaznih neurona

²CA-GRNN3 model ima 11 ulaznih neurona



Slika 21. Poređenje aktuelnih vrednosti GHG emisija sa predviđanjima CA-GRNN3 i CA-MLR modela za 2010. godinu

6.1.4. Rezultati GHG modela kreiranih sa ulaznim promenljivim transformisanim primenom analize glavnih komponenti

Analiza glavnih komponenti (PCA) je jedna od često korišćenih hemometrijskih tehnika za transformaciju promenljivih i smanjenje broja ulaznih promenljivih. PCA je sprovedena na redukovanoj grupi podataka za obučavanje (podaci koji su korišćeni za razvoj CA-GRNN3 modela) primenom IBM-ovog softverskog paketa SPSS 19 [162]. Kao što je opisano u Poglavlju 4.2., najpre je potrebno proveriti da li se PCA može primeniti na navedenu grupu podataka. U Tabeli 9. date su vrednosti kriterijuma na osnovu kojih se može zaključiti da li je PCA moguće primeniti. *Kaiser* [163] preporučuje da se PCA primeni na određenu grupu podataka samo ukoliko je vrednost *Kaiser-Meyer-Olkin* (KMO) testa adekvatnosti podataka veća od 0,5, a kako je u ovom slučaju vrednost KMO 0,616 (Tabela 9.), može se konstatovati da je navedeni uslov za primenu PCA ispunjen. Takođe, Barlettov test sferičnosti pokazuje da je značajnost podataka velika ($p < 0,0001$), pa je samim tim PCA primenljiva na navedenu grupu podataka.

Tabela 9. Lista ulaznih promenljivih na koje je primenjen PCA. Komunaliti vrednosti (*communalities extracted*), KMO vrednost i *Bartlett* test statistika

Kaiser-Meyer-Olkin (KMO)		0,616
Bartlett test	Chi-kvadrat test	2291,8
	Stepeni slobode	55
	Značajnost (<i>p</i>)	<0,0001
Ulaz	Komunalitet	
UEOI	0,965	
KGO	0,842	
BPE	0,909	
BDP	0,971	
PPA	0,935	
RPA	0,941	
BDPP	0,700	
PSV	0,766	
OOD	0,899	
EEOI	0,947	
BDPI	0,964	

Komunaliti vrednosti, koje se kreću od 0 do 1 i pokazuju koliki deo varijanse podataka je zadržan iz svake pojedinačne ulazne promenljive, prikazane su takođe u Tabeli 9., i kao što se može videti komunalitet za većinu ulaznih promenljivih, sem za BDPP i PSV, ima vrednosti veće od 0,8. Kao rotacioni metod tokom PCA korišćen je *Promax* sa *Kaiser* normalizacijom. PC skor koeficijent matrica (*PC score coefficient matrix*), svojstvene vrednosti korelacione matrice (*eigenvalues*) i odgovarajuće varijanse, kao i ukupna varijansa koja je ekstrahovana pojedinačnim komponentama prikazane su u Tabeli 10. Kao što se može videti, ukupna varijansa ekstrahovana sa 5 komponenti iznosi gotovo 90 %. Pet prikazanih komponenti korišćene su kao ulazne veličine GRNN modela (PCA-GRNN3), koji je u skladu sa tim imao 5 ulaznih neurona, dok su ostali parametri mreže bili podešeni kao i u slučaju GRNN3 modela. Radi poređenja, kreiran je još jedan MLR model sa PCA transformisanim ulazima, a kako kombinacija PCA ulaza sa linearnim MLR predstavlja regresiju glavnih komponenti (PCR), navedeni PCA-MLR model je označen kao PCR model. Koeficijenti PCR modela prikazani su u Tabeli 11.

Tabela 10. PC skor koeficijent matrica, svojstvene vrednosti korelacione matrice (*eigenvalues*), odgovarajuće varijanse i ukupna varijansa definisana glavnim komponentama

Ulaz	Komponenta				
	1	2	3	4	5
UEOI	0,056	0,034	0,449	0,002	0,007
KGO	-0,185	-0,022	-0,027	0,293	-0,155
BPE	-0,054	0,394	0,033	-0,012	0,202
BDP	-0,143	0,284	0,074	0,088	0,039
PPA	-0,011	0,006	0,041	0,547	0,038
RPA	0,110	0,461	-0,058	-0,024	-0,183
BDPP	0,321	0,012	0,001	-0,056	-0,090
PSV	0,367	0,046	0,034	0,058	0,062
OOD	0,308	-0,005	-0,176	0,433	0,014
EEOI	-0,044	-0,033	0,452	0,012	-0,014
BDPI	-0,002	0,004	-0,002	0,005	0,876
Svojstvene vrednosti	4,258	2,714	1,469	0,843	0,553
Varijansa po komponenti [%]	38,71	24,68	13,35	7,67	5,02
Kumulativna varijansa [%]	38,71	63,38	76,74	84,41	89,43

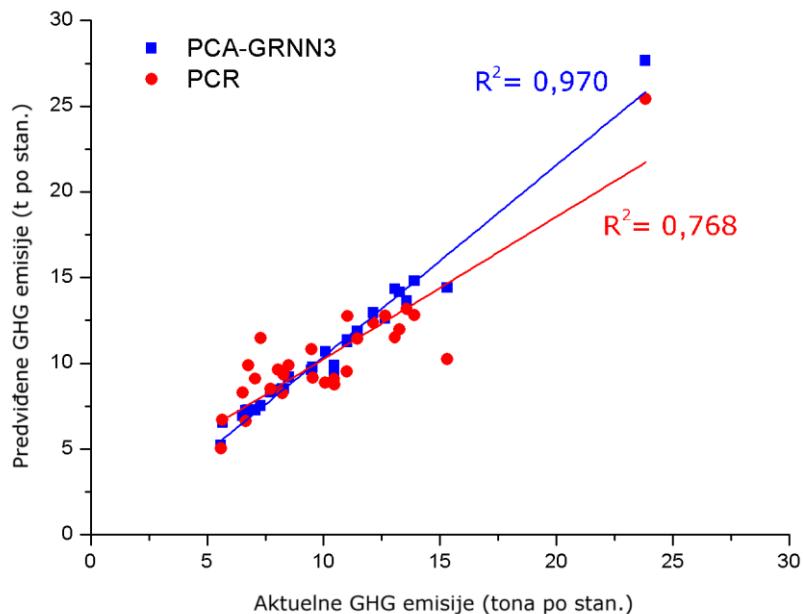
Tabela 11. Koeficijenti PCR GHG modela

Promenljive modela	Koeficijenti	Standardna greška
Konstanta	10,81	0,138
PC1	-0,373	0,175
PC2	3,373	0,173
PC3	-1,607	0,169
PC4	0,729	0,151
PC5	1,153	0,175

Vrednosti statističkih pokazatelja performansi za razvijene PCA-GRNN3 i PCR modele prikazane su u Tabeli 12., dok je poređenje dobijenih rezultata sa podacima za proveru (2010. godina) prikazano na Slici 22.

Tabela 12. Statistički pokazatelji performansi za PCA-GRNN3 i PCR modele

Model	Statistički pokazatelji performansi modela			
	<i>IA</i>	<i>FA1,1</i> [%]	<i>MAE</i> [t po stan.]	<i>RMSE</i> [t po stan.]
PCA-GRNN3	0,986	89	0,637	0,942
PCR	0,936	39	1,362	1,782



Slika 22. Poređenje aktuelnih vrednosti GHG emisija sa predviđanjima PCA-GRNN3 i PCR modela za 2010. godinu

PCA-GRNN3 model je dao značajno bolje predikcije u poređenju sa PCR modelom, ali je takođe primećena i slabija sposobnost uopštavanja u poređenju sa CA-GRNN3 modelom (Tabela 8., Slika 21.). Iako neki autori sugerišu da smanjivanje dimenzija primenom PCA može dovesti do gubitka informacija tokom navedenog procesa [15], što za posledicu može imati dobijanje modela lošijih karakteristika (npr. model daje slabije rezultate od očekivanih), u konkretnom slučaju, na osnovu dobijenih rezultata kod obučavanja, može se pretpostaviti da je došlo do "pretreniranja" PCA-GRNN3 modela, zbog čega model ima slabiju sposobnost uopštavanja u odnosu na CA-GRNN3, koji tokom obučavanja nije "pretreniran".

6.1.5. Analiza značajnosti pojedinih ulaznih promenljivih CA-GRNN3 modela

Analizom značajnosti (*sensitivity analysis*) dolazi se do podataka o uticaju pojedinačnih ulaznih promenljivih na razvijeni model. U slučaju GRNN modela, koji je obučavan sa genetičkim algoritmom, analiza značajnosti se može vršiti na osnovu vrednosti pojedinačnih faktora ujednačavanja (*individual smoothing factor* - ISF), koji su određeni tokom obučavanja za svaku pojedinačnu ulaznu promenljivu modela. ISF može imati vrednosti od 0 do 3, pri čemu, što je određena ulazna promenljiva značajnija, ona ima i veću vrednost faktora ujednačavanja. Analiza značajnosti je izvršena na osnovu ISF-a dobijenih tokom obučavanja najboljeg GRNN modela (CA-GRNN3) za predviđanje emisije gasova staklene bašte (Tabela 13.).

Tabela 13. Vrednosti ISF-a dobijene tokom obučavanja CA-GRNN3 modela

Ulaz	ISF
BPE	3,00
PPA	2,07
EEOI	0,91
BDPP	0,73
UEOI	0,41
KGO	0,19
RPA	0,05
PSV	0,05
BDP	0,02
BDPI	0,02
OOD	0,02

Dobijeni rezultati pokazuju da bruto potrošnja energije (BPE) ima najveći uticaj na CA-GRNN3 GHG model (ISF=3), dok po značajnosti dalje slede broj putnika prevezenih aviosaobraćajem (2,07), udeo električne energije dobijene iz obnovljivih izvora (0,91), udeo poljoprivrede u BDP-u (0,73) i udeo energije dobijene iz obnovljivih izvora (0,41). Relativno mali uticaj na model imaju ostale ulazne promenljive prikazane u Tabeli 13. Činjenica da BDP, koji je značajno povezan sa emisijom gasova staklene bašte, spada u grupu ulaznih promenljivih koje imaju neznatan uticaj na rezultate modela, znači da je GRNN „odlučio” da model zasnuje na BPE, koji je prema rezultatima korelacione analize (Tabela 7.) najznačajnije povezan sa GHG emisijom.

6.1.6. Ispitivanje mogućnosti primene razvijenog modela za predviđanje emisije gasova staklene bašte u 2011. godini

U trenutku modelovanja emisije gasova staklene bašte (februar 2013. god.) vrednosti GHG emisija za posmatrane evropske zemlje za 2011. godinu nisu bile dostupne na Eurostat-u, pa samim tim, dostupne vrednosti ulaznih promenljivih za 2011. godinu nisu mogle biti iskorišćene kod razvoja modela, što je prožilo mogućnost da se izvrši ispitivanje primene CA-GRNN3 modela u realnim uslovima tj. za predviđanje nepoznatih vrednosti emisija gasova staklene bašte za 2011. godinu. Da bi navedeno ispitivanje moglo da se izvrši najpre je bilo potrebno definisati relativnu grešku rezultata dobijenih tokom provere modela za svaku pojedinačnu državu. Poznavajući pojedinačne relativne greške za prethodnu godinu, u ovom slučaju 2010., mogu se korigovati predviđanja modela za narednu godinu (2011.) i dobiti tačnija predviđanja. U Tabeli 14. date su aktuelne i CA-GRNN3 modelom predviđene vrednosti GHG emisija za 2010. godinu, sa relativnim greškama i odgovarajućim statističkim pokazateljima. Kao što se može videti u Tabeli 14., samo u slučaju Letonije CA-GRNN3 model ja napravio relativnu grešku veću od 10 %. Srednja relativna greška (MAPE) za sva predviđanja za 2010. godinu ima odličnu vrednost, od svega 3,6 %.

Vrednosti emisije GHG gasova za evropske zemlje za 2011. godinu dobijene CA-GRNN3 modelom, zajedno sa vrednostima emisija gasova staklene bašte preuzete

od Evropske agencije za zaštitu životne sredine (*European Environment Agency - EEA*) [158], date su u Tabeli 15. Podaci o emisiji gasova staklene bašte za 2011. godinu nisu bili dostupni za Maltu i Kipar.

Tabela 14. Aktuelne i CA-GRNN3 modelom dobijene GHG emisije za 2010. godinu sa relativnim greškama i odgovarajućim statističkim pokazateljima

Zemlja	Aktuelne vrednosti [Mt]	Predviđanje modela [Mt]	Relativna greška [%]
Belgija	132,5	138,4	4,47
Bugarska	61,4	64,1	4,37
Češka	139,2	141,3	1,53
Danska	61,1	67,0	9,78
Nemačka	936,5	959,7	2,47
Estonija	20,5	19,2	-6,55
Irska	61,3	62,6	2,13
Grčka	118,3	125,6	6,15
Španija	355,9	367,6	3,28
Francuska	522,4	541,7	3,69
Italija	501,3	494,4	-1,39
Kipar	10,8	11,6	7,31
Letonija	12,1	10,5	-12,8
Litvanija	20,8	20,1	-3,61
Luksemburg	12,1	12,0	-0,84
Mađarska	67,7	67,9	0,31
Malta	3,00	3,10	3,02
Holandija	210,1	207,8	-1,08
Austrija	84,6	88,0	4,02
Poljska	400,9	391,8	-2,25
Portugal	70,6	74,2	5,13
Rumunija	121,4	123,9	2,10
Slovenija	19,5	19,8	1,31
Slovačka	46,0	47,8	3,99
Finska	74,6	72,6	-2,65
Švedska	66,2	66,2	-0,09
Velika Britanija	590,2	576,9	-2,26
Norveška	53,9	53,0	-1,74
Statistički pokazatelji performansi	<i>MAPE</i> [%]		3,60
	Maksimalna rel. greška [%]		-12,8
	Minimalna rel. greška [%]		-0,09
	<i>FA1,1</i> [%]		96
	<i>MAE</i> [Mt]		4,7
		<i>RMSE</i> [Mt]	7,4

Tabela 15. Aktuelne i CA-GRNN3 modelom dobijene GHG emisija za 2011. godinu sa relativnim greškama i odgovarajućim statističkim pokazateljima

Zemlja	Aktuelne vrednosti [Mt] [158]	Predviđanje modela [Mt]	Relativna greška [%]
Belgija	121,3	130,4	7,50
Bugarska	67,9	65,8	-3,09
Češka	141,1	136,2	-3,47
Danska	56,1	64,5	15,0
Nemačka	917	918,9	0,21
Estonija	20,9	18,7	-10,5
Irska	57,3	63,3	10,5
Grčka	118,5	125,8	6,16
Španija	356,1	369,3	3,71
Francuska	497,5	533,5	7,24
Italija	493,7	496,0	0,47
Kipar	-	12,4	-
Letonija	12,1	10,6	-12,4
Litvanija	21,4	20,1	-6,07
Luksemburg	12,3	12,0	-2,44
Mađarska	65,6	67,7	3,20
Malta	-	4,9	-
Holandija	195,8	207,6	6,03
Austrija	81,9	88,3	7,81
Poljska	409,3	388,8	-5,01
Portugal	70	74,0	5,71
Rumunija	123,7	125,5	1,46
Slovenija	19,5	20,8	6,67
Slovačka	45,9	45,7	-0,44
Finska	67,3	70,6	4,90
Švedska	62,8	65,3	3,98
Velika Britanija	553,8	581,1	4,93
Norveška	52,7	53,2	0,95
Statistički pokazatelji performansi	<i>MAPE</i> [%]		5,40
	Maksimalna rel. greška [%]		15,0
	Minimalna rel. greška [%]		0,21
	<i>FA1,1</i> [%]		85
	<i>MAE</i> [Mt]		6,8
		<i>RMSE</i> [Mt]	11,0

Poređenje aktuelnih i modelom predviđenih GHG emisija pokazuje da je maksimalna relativna greška modela bila 15 %, dok je samo u slučaju četiri zemlje ona iznosila više od 10 %: Danska, Estonija, Irska i Letonija. Srednja relativna greška

(MAPE) za sva predviđanja za 2011. godinu porasla je u odnosu na 2010. godinu na 5,4 %.

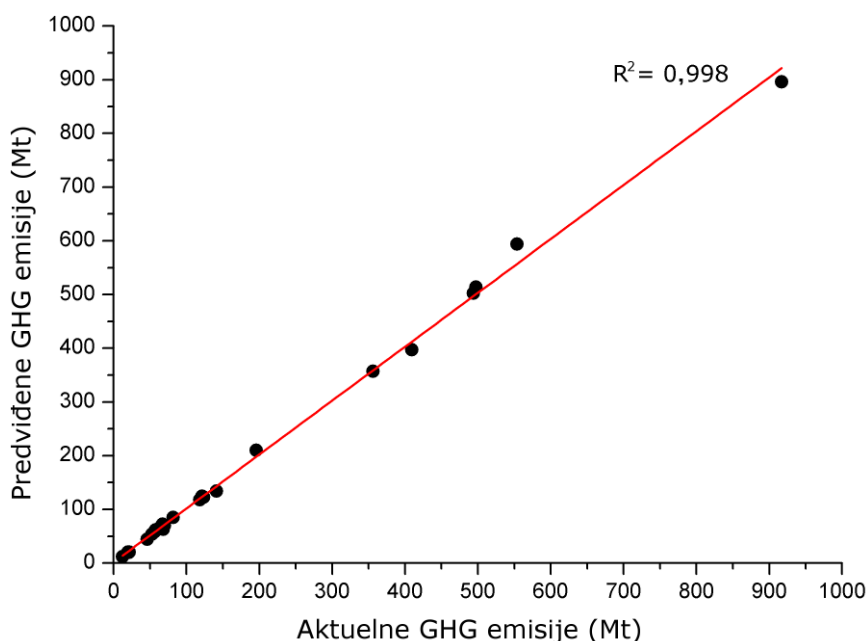
Tabela 16. Aktuelne i korigovane CA-GRNN3 modelom dobijene GHG emisija za 2011. godinu sa relativnim greškama i odgovarajućim statističkim pokazateljima

Zemlja	Aktuelne vrednosti [Mt] [158]	Korigovana predviđanja modela [Mt]	Relativna greška [%]
Belgija	121,3	124,6	2,72
Bugarska	67,9	62,9	-7,36
Češka	141,1	134,1	-4,96
Danska	56,1	58,2	3,74
Nemačka	917	896,1	-2,28
Estonija	20,9	19,9	-4,78
Irska	57,3	62,0	8,20
Grčka	118,5	118,0	-0,42
Španija	356,1	357,2	0,31
Francuska	497,5	513,8	3,28
Italija	493,7	502,8	1,84
Kipar	-	11,5	-
Letonija	12,1	12,0	-0,83
Litvanija	21,4	20,8	-2,80
Luksemburg	12,3	12,1	-1,63
Mađarska	65,6	67,5	2,90
Malta	-	4,7	-
Holandija	195,8	209,8	7,15
Austrija	81,9	84,8	3,54
Poljska	409,3	397,5	-2,88
Portugal	70,0	70,2	0,29
Rumunija	123,7	122,9	-0,65
Slovenija	19,5	20,5	5,13
Slovačka	45,9	43,9	-4,36
Finska	67,3	72,5	7,73
Švedska	62,8	65,3	3,98
Velika Britanija	553,8	594,3	7,31
Norveška	52,7	54,1	2,66
Statistički pokazatelji performansi	<i>MAPE</i> [%]		3,60
	Maksimalna rel. greška [%]		8,20
	Minimalna rel. greška [%]		0,20
	<i>FA1,1</i> [%]		100
	<i>MAE</i> [Mt]		6,0
<i>RMSE</i> [Mt]		10,6	

Kada se modelom dobijene predikcije za 2011. godinu koriguju vrednostima relativnih grešaka kod modelom predviđenih emisija za 2010. godinu dobijaju se znatno tačnije vrednosti emisija gasova staklene bašte (Tabela 16.). Korigovana predviđanja GHG emisija imaju relativne greške manje od 10 %, pri čemu je maksimalna relativna greška bila samo 8 %. Takođe vrednost srednje relativne greške za predviđanja za 2011. godinu, posle korekcije, iznosi samo 3,6 %, identična je srednjoj relativnoj grešci kod provere modela, što govori da je greška modela na adekvatan način kontrolisana.

Neophodno je istaći da je razvijeni GRNN model posebno precizan za jednogodišnje predviđanje GHG emisija, kada su poznati podaci o njihovoj emisiji iz prethodne godine. Na ovaj način, podaci iz poslednje godine kod obučavanja modela imaju veću težinu od ostalih, jer se koriste i za razvoj modela i za korekciju dobijenih predviđanja.

Poređenje korigovanih predikcija emisije gasova staklene bašte za 2011. godinu i aktuelnih vrednosti emisija dato je na Slici 23. Kao što se može videti, vrednost koeficijenta korelacije je gotovo jednak 1.



Slika 23. Poređenje aktuelnih i korigovanih modelom predviđenih emisija gasova staklene bašte za 2011. godinu

6.2. Modelovanje emisije PM₁₀ čestica

U ovom poglavlju opisan je razvoj i primena ANN modela za predviđanje godišnjih emisija PM₁₀ čestica na nacionalnom nivou, sa odabranim ekonomskim i industrijskim indikatorima i indikatorima održivog razvoja kao ulaznim promenljivim. Odabir najznačajnijih ulaznih promenljivih je vršen genetičkim algoritmom. Kao i u slučaju modelovanja emisije gasova staklene bašte, arhitektura neuronskih mreža sa opštom regresijom (GRNN) odabrana je za kreiranje modela za emisiju PM₁₀ čestica. Nakon razvoja odgovarajućeg GRNN modela, razvijeni su i modeli bazirani na višestrukoj linearnoj regresiji (MLR) i regresiji glavnih komponenti (PCR) radi poređenja sa rezultatima GRNN modela. Pokazano je da su ANN modeli za predviđanje emisije čestica adekvatno alternativno rešenje u slučajevima kada države nemaju sve potrebne podatke za procenu emisije korišćenjem nekog od već postojećih linearnih modela (npr. RAINS, GAINS). Kod ANN modela je potrebno i daleko manje ulaznih promenljivih, pa su samim tim oni primenljiviji u slučajevima kada se predviđanja emisije čestica vrše za zemlje u razvoju ili nerazvijene zemlje [164].

6.2.1. Ulazne promenljive i priprema podataka za PM₁₀ model

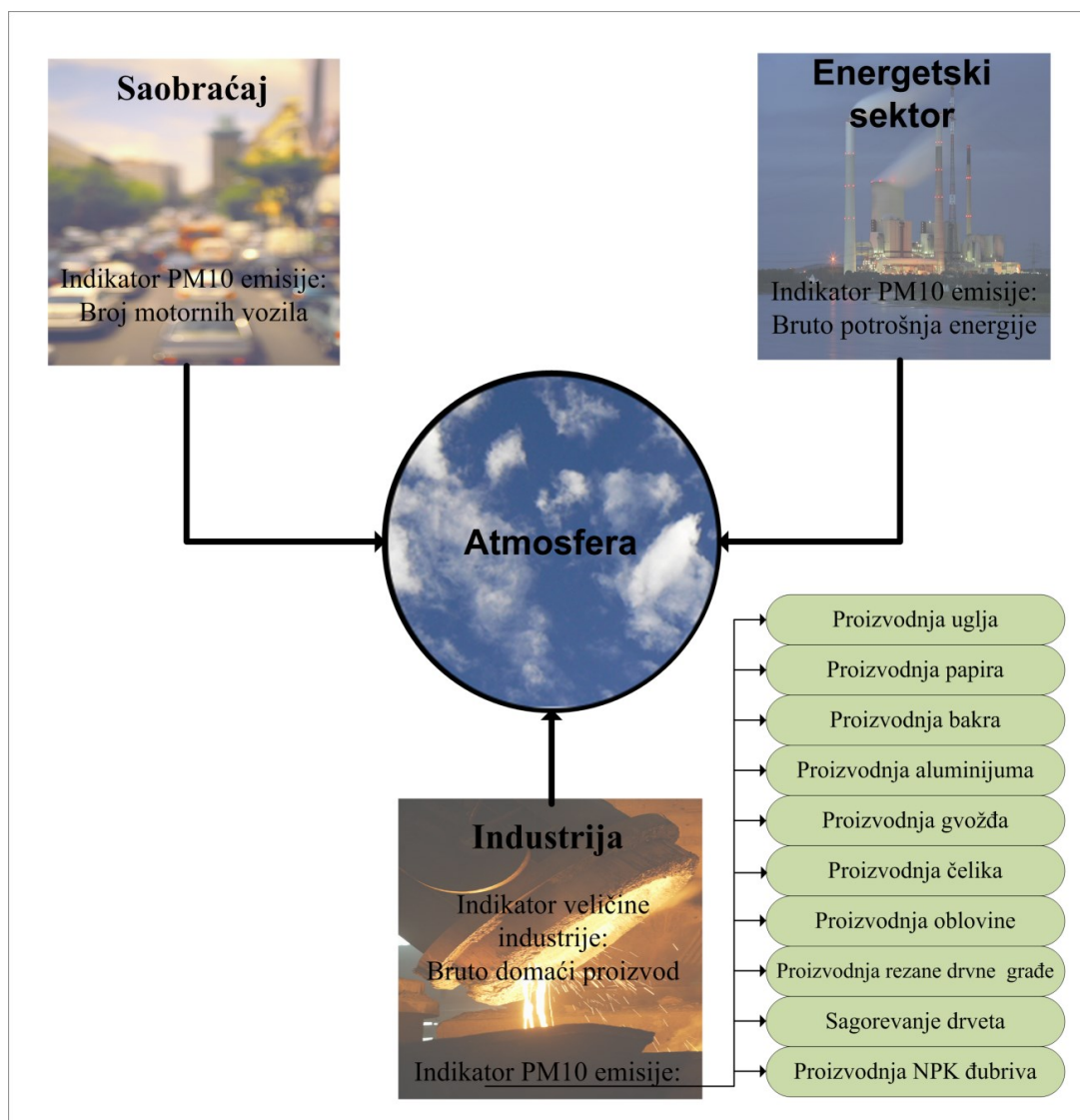
Kao što je već navedeno, široka dostupnost indikatora održivog razvoja čini ih pogodnim za razvoj modela u zaštiti životne sredine. U konkretnom slučaju modelovanja emisije čestica (PM₁₀) na nacionalnom nivou, sledeći indikatori su odabrani kao osnovne ulazne promenljive modela: bruto domaći proizvod (BDP), bruto potrošnja energije (BPE) i broj motornih vozila po stanovniku. BPE i broj motornih vozila su odabrani kao relevantni indikatori energetske sektora i sektora saobraćaja, koji predstavljaju dva od tri glavna sektora emisije PM₁₀ čestica, dok je BDP odabran kao indikator veličine privrede, a samim tim i stepena razvoja industrije, koja je treći glavni izvor antropogene emisije PM₁₀ čestica. Pored navedenih glavnih ulaznih promenljivih, odabrani su i indikatori koji pokazuju aktivnost pojedinih industrijskih grana u kojima takođe dolazi do emisije PM₁₀ čestica. U Tabeli 17. dat je spisak svih ulaznih promenljivih, sa izvorima iz kojih su preuzeti i jedinicama mere posle normalizacije. Normalizacija je vršena po stanovniku, a u slučaju BDP-a

i sa prosečnom vrednošću BDP-a Evropske unije (EU27). Veza između sektora emisije, odabranih ulaznih promenljivih i emisije PM₁₀ čestica, prikazana je na Slici 24.

Tabela 17. Spisak dostupnih ulaznih promenljivih i izvora podataka

Oznaka	Ulazna promenljiva	Jedinica	Izvor podataka
V1	Bruto domaći proizvod (BDP)	-	
V2	Bruto potrošnja energije (BPE)	toe po stan. ¹	
V3	Spaljivanje drveta	m ³ po stan.	
V4	Broj motornih vozila	vozila po stan.	Eurostat [4]
V5	Proizvodnja uglja	toe po stan.	
V6	Proizvodnja papira	t po stan.	
V7	Proizvodnja drvene građe (oblovine)	m ³ po stan.	
V8	Proizvodnja rezane drvene građe	m ³ po stan.	
V9	Proizvodnja bakra	t po stan.	<i>U.S. Geological Survey</i> [165]
V10	Proizvodnja aluminijuma	t po stan.	
V11	Proizvodnja gvožđa	t po stan.	<i>International Iron and Steel Institute</i> [166]
V12	Proizvodnja čelika	t po stan.	
V13	Proizvodnja NPK đubriva	t po stan.	<i>International Fertilizer Industry Association</i> [167]
Izlazna promenljiva			
PM10	Emisija PM ₁₀ čestica	t po stan.	Eurostat [4]

¹Tona ekvivalentne nafte (toe) po stanovniku



Slika 24. Veza odabranih ulaznih promenljivih i PM₁₀ emisije

Kako su na Eurostatu tokom 2011. godine bili dostupni podaci o PM₁₀ emisijama za period 1999-2006. godina [4], podaci od 1999. do 2004. korišćeni su za obučavanje, dok su podaci za 2005. i 2006. godinu korišćeni za proveru modela.

Podaci o PM₁₀ emisijama zemalja Evropske unije (svih osim Malte), kao za i EU27 u celini, korišćeni su za kreiranje PM₁₀ emisionog modela i prikazani su u Tabeli 18. (za godine 1999., 2004. i 2006.).

Tabela 18. Podaci za PM₁₀ emisiju (srednja vrednost za period 1999-2006, vrednosti za 1999., 2004. i 2006.)

Region/Država	Emisija PM ₁₀ čestica (kg po stan.)			
	Srednja vrednost	Godina		
		1999.	2004.	2006.
EU 27	42,44	47,53	40,53	38,48
Austrija	36,88	35,25	37,59	36,23
Belgija	43,89	44,43	42,43	38,64
Bugarska	104,1	98,97	105,7	106,9
Kipar	66,27	75,01	61,70	50,91
Češka	50,08	57,08	49,23	42,69
Danska	55,40	61,70	-	50,18
Estonija	85,61	90,73	86,23	67,61
Finska	58,18	61,37	58,09	55,85
Francuska	42,68	48,52	40,77	37,87
Nemačka	28,88	33,20	27,42	25,72
Grčka	63,11	64,74	-	-
Mađarska	43,42	57,39	38,71	34,38
Irska	66,07	79,17	56,81	-
Italija	33,05	39,81	30,36	26,83
Letonija	26,42	29,60	26,08	28,29
Litvanija	30,15	32,05	29,93	32,52
Luksemburg	23,60	29,02	21,39	14,90
Holandija	29,84	34,39	28,11	26,35
Poljska	51,61	59,13	48,71	49,70
Portugal	59,66	61,47	-	-
Rumunija	33,33	29,76	-	42,89
Slovačka	50,86	64,95	35,67	33,52
Slovenija	38,93	48,21	45,22	35,23
Španija	60,16	64,26	59,19	54,44
Švedska	31,82	36,56	30,16	28,32
Velika Britanija	40,98	48,18	37,83	35,12
Grupa podataka za	Srednja vrednost	Minimalna vrednost	Maksimalna vrednost	
Obučavanje	49,30	21,39	106,8	
Proveru	42,23	14,90	106,9	

6.2.2. Razvoj i optimizacija GRNN modela

Prvi korak u kreiranju modela neuronskih mreža je izdvajanje validacionih podataka iz grupe podataka za obučavanje, koji se koriste tokom obučavanja za sprečavanje "pretreniranosti" mreže. Validacioni podaci su izdvojeni nasumičnim odabirom 24 serije podataka (16 %) od ukupnog broja (154) serija podataka za obučavanje mreže. Kao i u slučaju modelovanja emisije gasova staklene bašte, arhitektura neuronskih mreža sa opštom regresijom (GRNN) odabrana je za kreiranje emisijonog modela PM_{10} čestica. Kako glavni parametri GRNN arhitekture (broj neurona po slojevima) zavise od broja ulaznih i izlaznih promenljivih, kao i od broja serija podataka za obučavanje, vrednosti navedenih parametara su automatski podešene. Svi testirani GRNN modeli imali su 1 neuron u izlaznom sloju, 2 neurona u sloju za sumiranje, 154 neurona u skrivenom sloju, dok je broj neurona u ulaznom sloju menjan od 7 do 13, u zavisnosti od broja ulaznih promenljivih. Jedini parametri arhitekture koji su mogli da se menjaju su:

- funkcija za skaliranje podataka, za koju je u ovom slučaju odabrana linearna funkcija (Jednačina 34.), i
- metod za određivanje rastojanja između podataka, za koji je u ovom slučaju odabran Euklidski metod, koji se inače preporučuje kao najbolji za primenu kod većine ANN mreža.

Najvažniji parametar kod obučavanja GRNN mreže je algoritam za određivanje faktora ujednačavanja, koji određuje performanse GRNN modela. Faktor ujednačavanja može biti određen ili iterativnim ili genetskim algoritmom, pri čemu je u ovom slučaju odabran genetički algoritam, koji se preporučuje kod modelovanja parametara kod kojih ulazne veličine imaju različit uticaj (značajnost) na izlaznu promenljivu, i koji prilikom obučavanja, pored opšteg faktora ujednačavanja, daje i pojedinačne faktore ujednačavanja (ISF) za svaku ulaznu promenljivu. Vrednosti navedenih ISF, koji se kreću od 0 do 3, odgovaraju značajnosti samih ulaznih promenljivih i kao takvi mogu se koristiti kao parametri na osnovu kojih je moguće odrediti ulazne promenljive koje ne utiču značajno na izlaznu promenljivu i potencijalno unose grešku u rezultate modela. Tokom modelovanja emisije PM_{10} čestica

razvijene su dve uzastopne procedure za eliminisanje manje značajnih ulaznih promenljivih na bazi ISF vrednosti:

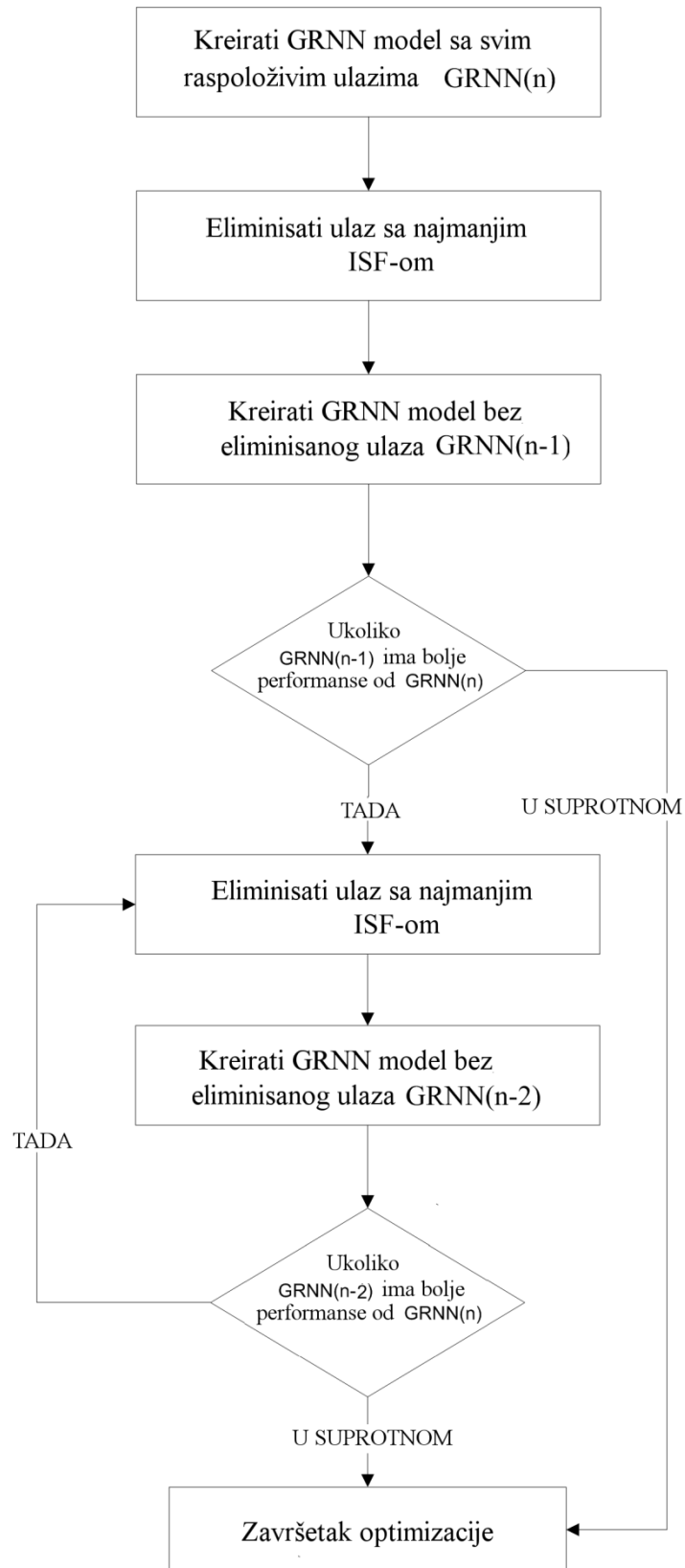
- eliminaciona procedura,
- procedura zasnovana na „glavnim“ ulaznim promenljivim.

Šematski prikaz (algoritam) eliminacione procedure dat je na Slici 25. Iako se ISF vrednosti menjaju svaki put kada se razvije nov GRNN model, bez neke od početnih ulaznih promenljivih, navedena procedura (Slika 25.) se pokazala kao efikasan način za eliminaciju ulaznih promenljivih koje nemaju značajan uticaj na izlaznu promenljivu modela. Važno je naglasiti da je kriterijum za završetak optimizacije bio poređenje poslednjeg razvijenog modela sa prvim modelom koji je razvijen sa svim ulaznim promenljivim (neoptimizovan model). Optimizacija je samim tim vršena sve dok su se eliminisanjem pojedinačnih ulaznih promenljivih dobijali modeli sa boljim performansama u poređenju sa neoptimizovanim modelom.

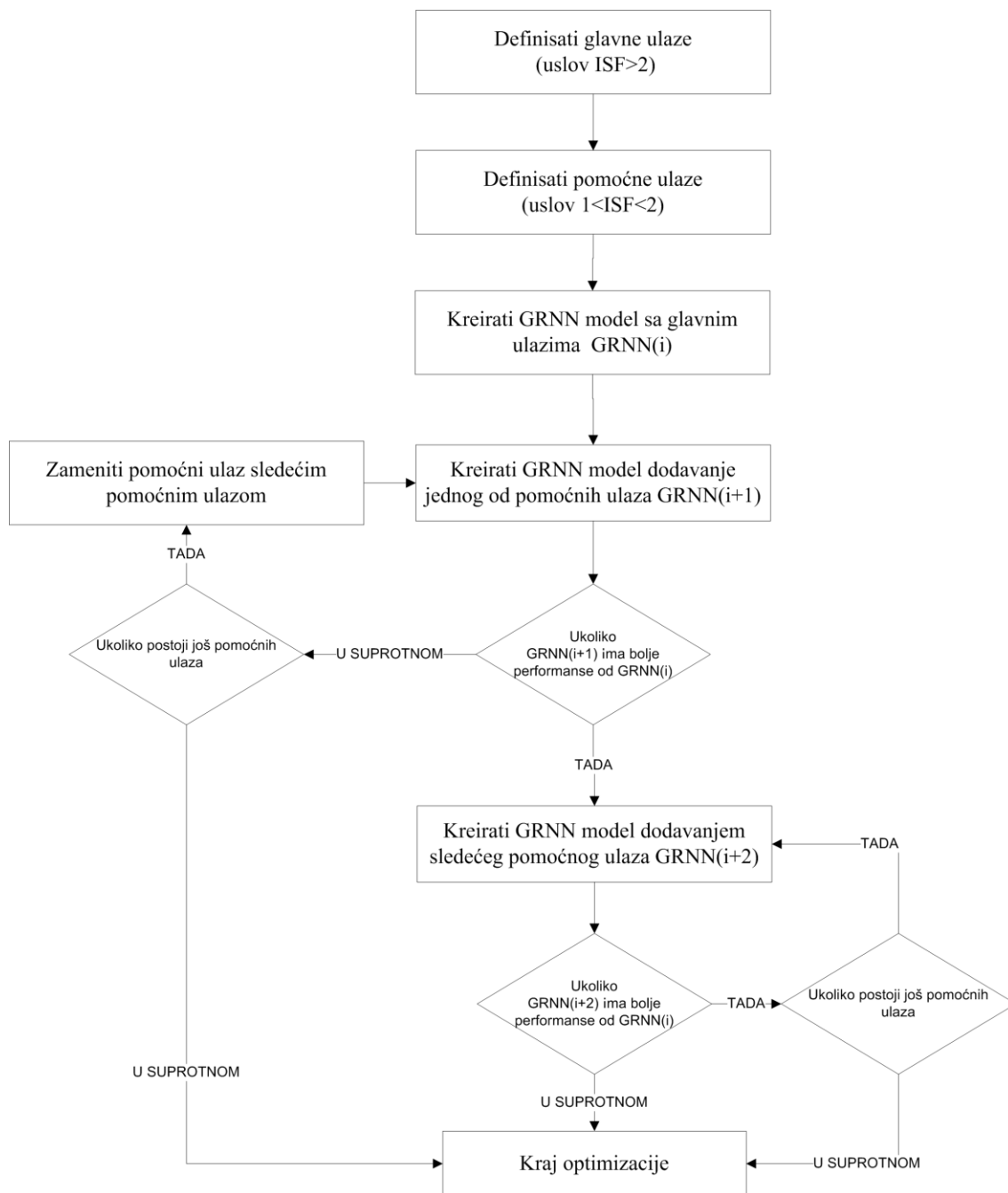
Šematski prikaz (algoritam) procedure zasnovane na „glavnim“ ulaznim promenljivim dat je na Slici 26., vrši se na osnovu poznavanja vrednosti ISF određenih tokom eliminacione procedure optimizacije ulaza. Ulazne promenljive koje su imale vrednost ISF u nekom od GRNN modela, kreiranih tokom eliminacione procedure:

- veću od 2 smatraju se za „glavne“ ulazne promenljive,
- između 1 i 2 smatraju se za „pomoćne“ ulazne promenljive.

Ulazne promenljive koje su tokom eliminacione procedure imali vrednost ISF konstantno ispod 1 se zanemaruju. Na taj način, najpre se kreira GRNN model koji čine samo „glavne“ ulazne promenljive, a zatim i dodatni modeli koji se dobijaju kombinovanjem „glavnih“ i „pomoćnih“ ulaznih promenljivih.



Slika 25. Algoritam eliminacione procedure za odabir ulaznih promenljivih modela (ulaz = ulazna promenljiva)



Slika 26. Algoritam procedure, za odabir ulaznih promenljivih modela, zasnovane na „glavnim“ ulaznim promenljivim (ulaz = ulazna promenljiva)

6.2.3. Rezultati optimizacije GRNN modela

Pet GRNN modela, razvijenih primenom eliminacione procedure optimizacije, zajedno sa dobijenim ISF vrednostima, prikazani su u Tabeli 19. Vrednosti statističkih pokazatelja performansi modela za GRNN modele kreirane eliminacionom procedurom optimizacije date su u Tabeli 20. Kao što se može videti u Tabeli 20., GRNN10 model, koji je kreiran bez tri primarno odabrane ulazne promenljive (proizvodnje uglja, NPK đubriva i oblovine), daje najbolje rezultate. Vrednosti *IA* iznose 0,99 za jednogodišnje i 0,98 za dvogodišnje predviđanje, dok *FA1.25* uzima vrednosti od 100 % do 91 % za jednogodišnje, odnosno dvogodišnje predviđanje emisije PM_{10} čestica. Kako je neoptimizovani GRNN13 model imao slabije vrednosti navedenih statističkih pokazatelja performansi, može se zaključiti da je optimizacija uspela i da je rezultovala razvojem tačnijeg modela, pri čemu se koristi čak 23 % manje ulaznih indikatora i time smanjuje vreme potrebno za pripremu ulaznih podataka, a takođe se smanjuje i nesigurnost modela, kada se vrši predviđanje na bazi ulaznih promenljivih koje su takođe dobijene modelima.

Tabela 19. ISF vrednosti ulaznih promenljivih za razvijene GRNN modele eliminacionom procedurom

Ulaz	GRNN model				
	GRNN13	GRNN12	GRNN11	GRNN10	GRNN9
V1	2,79	2,18	2,32	2,55	1,18
<i>V2</i>	1,61	1,09	0,41	1,66	0,48
<i>V3</i>	1,92	1,64	0,05	1,13	0,34
V4	2,56	2,40	0,42	1,95	0,81
<i>V5</i>	0,01	-	-	-	-
V6	0,33	0,42	2,36	0,07	0,48
<i>V7</i>	0,26	0,32	-	-	-
V8	2,78	2,95	2,93	3,00	2,92
V9	2,44	2,34	2,04	2,88	2,76
V10	1,35	0,41	2,71	1,19	0,75
<i>V11</i>	1,75	1,72	1,07	0,02	-
V12	0,55	0,54	2,15	1,27	1,53
<i>V13</i>	0,24	0,60	0,02	-	-

bold oznaka - "glavne" ulazne promenljive

italic oznaka - "pomoćne" ulazne promenljive

Tabela 20. Statistički pokazatelji performansi GRNN modela kreiranih eliminacionom procedurom

Model	Period	Statistički pokazatelji performansi modela			
		IA	FA1.25 [%]	RMSE [kg po stan.]	MAE [kg po stan.]
GRNN13	2005	0,99	100	3,96	-2,60
	2005-06	0,96	89	7,00	-1,07
GRNN12	2005	0,99	100	3,93	-2,50
	2005-06	0,96	89	6,92	-1,03
GRNN11	2005	0,98	100	4,81	-3,05
	2005-06	0,97	89	6,34	-1,80
GRNN10	2005	0,99	100	3,99	-2,44
	2005-06	0,98	91	5,71	-1,25
GRNN9	2005	0,99	100	4,35	-2,72
	2005-06	0,97	89	6,04	-1,40

Kao što je prikazano u Tabeli 19., „glavne“ ulazne promenljive na osnovu ISF vrednosti su: bruto domaći proizvod (BDP) (V1), broj motornih vozila (V4), kao i proizvodnja papira (V6), rezane drvene građe (V8), bakra (V9), aluminijuma (V10) i čelika (V12). Kao „pomoćne“ ulazne promenljive definisani su bruto potrošnja energije (BPE) (V2), spaljivanje drveta (V3) i proizvodnja gvožđa (V11), dok su preostali ulazni indikatori (proizvodnja uglja (V5), drvene građe – oblovine (V7) i NPK đubriva (V13)) zanemareni. Razvijeni su sledeći modeli:

- GRNN7g - model samo sa „glavnim“ ulaznim promenljivim,
- GRNN8g - model koji pored „glavnih“ ulaznih promenljivih ima i „pomoćnu“ ulaznu promenljivu sa najvećom ISF vrednošću (spaljivanje drveta),
- GRNN9g – model koji pored ulaznih promenljivih koje su korišćene kod GRNN8g modela ima još jedanu „pomoćnu“ ulaznu promenljivu (proizvodnja gvožđa),
- GRNN9ga - model koji pored ulaznih promenljivih koje su korišćene kod GRNN8g modela ima još jedanu „pomoćnu“ ulaznu promenljivu (bruto potrošnja energije).

Vrednosti statističkih pokazatelja performansi modela za GRNN modele razvijene procedurom zasnovanom na „glavnim“ ulaznim promenljivim date su u Tabeli 21.

Tabela 21. Statistički pokazatelji performansi GRNN modela kreiranih procedurom sa „glavnim“ ulaznim indikatorima

Model	Period	Statistički pokazatelji performansi modela			
		IA	FA1.25 [%]	RMSE [kg po stan.]	MAE [kg po stan.]
GRNN7g	2005	0,99	100	4,39	-2,59
	2005-06	0,93	87	8,99	-0,98
GRNN8g	2005	0,99	100	4,32	-2,68
	2005-06	0,97	89	5,98	-1,39
GRNN9g	2005	0,99	100	4,67	-2,78
	2005-06	0,95	87	7,56	-1,39
GRNN9ga	2005	0,99	100	4,59	-2,76
	2005-06	0,95	87	7,64	-1,37

Kao što se može videti, najbolji rezultati su ostvareni sa GRNN8g modelom. Poređenjem GRNN8g (Tabela 21.) sa GRNN10 (Tabela 20.) modelom kreiranim eliminacionom procedurom, može se zaključiti da je GRNN10 model sa najboljim performansama za predviđanje PM₁₀ emisije na nacionalnom nivou.

6.2.4. Poređenje GRNN10 modela sa linearnim modelima

GRNN model sa najboljim rezultatima (GRNN10) u ovom poglavlju je upoređen sa rezultatima dva tipa linearnih modela: MLR modelom i regresijom zasnovanom na glavnim komponentama (PCR). Da bi se ujedno ispitalo i da li eliminaciona optimizaciona procedura daje rezultate i kod linearnih modela, razvijena su po dva različita MLR i PCR modela:

- MLR13 i PCR13 – modeli sa svim raspoloživim ulaznim promenljivim – ekvivalenti neoptimizovanom GRNN modelu,
- MLR10 i PCR10 – sa ulaznim promenljivim kao i kod GRNN10 modela.

Koeficijenti kreiranih linearnih modela prikazani su u Tabelama 22. i 23. Kod razvoja PCR modela vršeno je nasumično izdvajanje validacionih podataka, kao i kod obučavanja GRNN modela. Vrednosti statističkih pokazatelja performansi modela za najbolje linearne modele kod njihove provere prikazane su u Tabeli 24.

Tabela 22. Koeficijenti kreiranih linearnih modela MLR13 i PCR13

Promenljive modela	MLR13		PCR13	
	Koeficijenti	Standardna greška	Koeficijenti	Standardna greška
Konstanta	0,047	0,008	0,056	0,036
V1	-0,023	0,007	-0,020	0,006
V2	0,016	0,004	0,010	0,003
V3	0,032	0,012	0,037	0,011
V4	-0,014	0,018	-0,013	0,019
V5	0,004	0,003	0,003	0,003
V6	0,022	0,007	0,021	0,006
V7	-0,009	0,003	-0,006	0,003
V8	0,001	0,010	-0,012	0,008
V9	0,271	0,143	-0,002	0,123
V10	-0,713	0,160	-0,660	0,133
V11	-0,013	0,003	-0,008	0,003
V12	-0,050	0,011	-0,011	0,003
V13	-0,141	0,034	-0,149	0,028

Tabela 23. Koeficijenti kreiranih linearnih modela MLR10 i PCR10

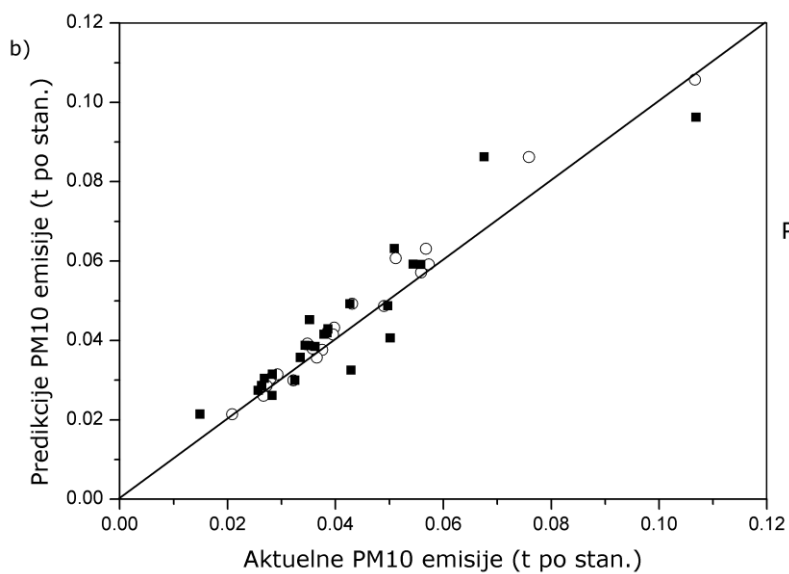
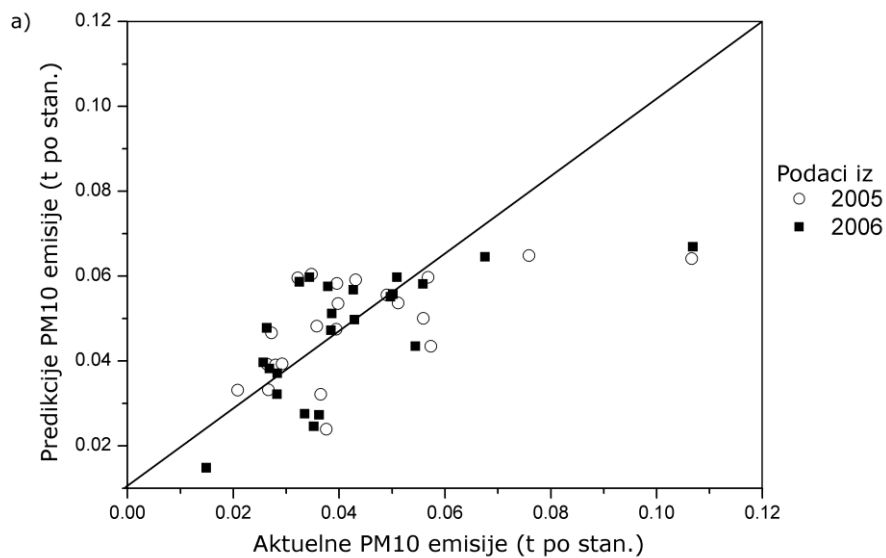
Promenljive modela	MLR10		PCR10	
	Koeficijenti	Standardna greška	Koeficijenti	Standardna greška
Konstanta	0,050	0,006	0,050	0,024
V1	-0,011	0,005	-0,011	0,005
V2	0,007	0,002	0,008	0,003
V3	0,024	0,010	0,020	0,010
V4	-0,013	0,003	-0,019	0,020
V6	0,009	0,006	0,004	0,007
V8	-0,020	0,006	-0,015	0,006
V9	-0,056	0,001	-0,075	0,128
V10	-0,490	0,167	-0,603	0,150
V11	-0,007	0,001	-0,008	0,002
V12	-0,010	0,004	-0,012	0,003

Tabela 24. Statistički pokazatelji performansi najboljih linearnih modela (MLR10 i PCR10)

Model	Period	Statistički pokazatelji performansi modela			
		IA	FA1.25 [%]	RMSE [kg po stan.]	MAE [kg po stan.]
MLR10	2005	0,68	30	16,45	-5,36
	2005-06	0,72	41	15,75	-2,81
PCR10	2005	0,69	39	16,24	-5,37
	2005-06	0,73	39	15,57	-2,64

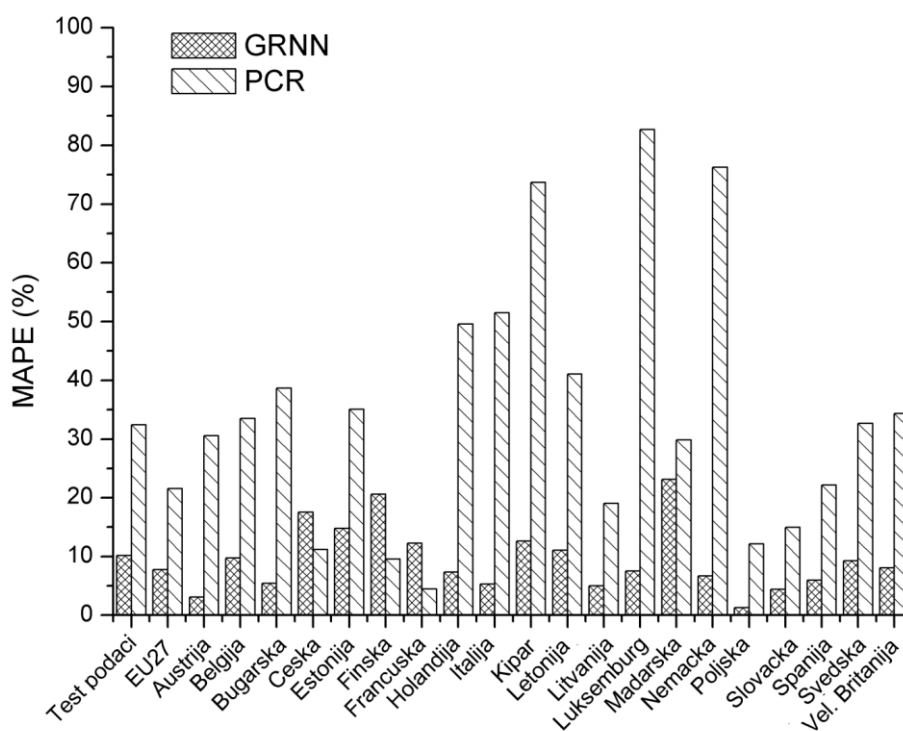
Primena linearnih modela na podatke korišćene za njihov razvoj pokazala je da neoptimizovani modeli (MLR13 i PCR13) imaju bolje performanse u odnosu na modele sa optimizovanim brojem ulaznih promenljivih (MLR10 i PCR10). Međutim, kada su navedeni linearni modeli provereni, i time testirana njihova prava sposobnost predviđanja, rezultati su bili potpuno drugačiji (Tabela 24.). Rezultati optimizovanih modela imali su do 50 % veću vrednost *IA* u poređenju sa neoptimizovanim modelima, pri čemu je najbolje rezultate imao PCR10 model.

Ipak, poređenje performansi najboljeg linearnog modela PCR10 sa najboljim GRNN10 modelom jasno pokazuje da GRNN10 model daje znatno bolja predviđanja emisije PM_{10} čestica. Na Slici 27. prikazano je poređenje GRNN10 i PCR10 predviđanja sa aktuelnim vrednostima emisije PM_{10} čestica za 2005. i 2006. godinu.



Slika 27. Poređenje aktuelnih i modelima predviđenih PM10 emisija za 2005. i 2006. godinu: a) PCR10 b)GRNN10

Srednja relativna greška (MAPE) za predviđanja GRNN10 i PCR10 modela kod provere za pojedinačne EU države prikazana je na Slici 28. (prikazane su samo zemlje za koje su bili dostupni podaci za obe godine korišćene za proveru).



Slika 28. MAPE vrednosti predikcija GRNN10 i PCR10 modela za pojedinačne EU države

Kao što se može videti na Slici 28., GRNN10 model je dao tačnija predviđanja kod provere modela, kao i za većinu država. Jedino u slučaju Češke, Francuske i Finske PCR10 model je imao predviđanja sa greškom manjom od GRNN10 modela.

6.2.5 Analiza grešaka kod GRNN10 modela

Pad tačnosti modela kod dvogodišnjeg, u odnosu na jednogodišnje predviđanje, posledica je promena u samoj emisiji PM₁₀ čestica koja se desila posle perioda obučavanja modela (posle 2004. i tokom 2005-2006. godine). U Tabeli 25. prikazana je promena u emisiji PM₁₀ čestica tokom 1999-2006 i promena u emisiji tokom 2004-2006. Naime, u periodu 1999-2006. koji je obuhvaćen ovim modelom, emisija PM₁₀ čestica na nivou EU27 opala je za 19 %, dok je za pojedine zemlje taj pad iznosio i više od 30 %, pa sve do gotovo 50 % (Tabela 25.). Navedeni pad emisije PM₁₀ čestica posledica je donošenja i primene strožijih EU direktiva vezanih za zaštitu život-

ne sredine (1996/61/EC [168], 1998/69/EC [169], 1999/30/EC [170], 2001/80/EC [26]), kojima je ograničena emisija PM₁₀ čestica.

Tabela 25. Promena količine emitovanih PM10 čestica za analizirane zemlje

Region/Država	Promena u PM10 emisiji [%]	
	Period	
	1999-2006	2004-2006
EU 27	-19,0	-5,05
Austrija	2,79	-3,60
Belgija	-13,0	-8,93
Bugarska	7,99	1,07
Kipar	-32,1	-17,5
Češka	-25,2	-13,3
Danska	-18,9	-
Estonija	-25,5	-21,6
Finska	-8,98	-3,85
Francuska	-22,0	-7,11
Nemačka	-22,6	-6,21
Mađarska	-40,1	-11,2
Italija	-32,6	-11,6
Letonija	-4,41	8,47
Litvanija	1,45	8,63
Luksemburg	-48,6	-30,3
Holandija	-23,4	-6,25
Poljska	-16,0	2,04
Rumunija	44,1	-
Slovačka	-45,8	-6,02
Slovenija	-30,5	-22,1
Španija	-15,3	-8,02
Švedska	-22,5	-6,09
Velika Britanija	-27,1	-7,16

U slučaju jednogodišnjeg predviđanja (za 2005. godinu), GRNN10 model je dao rezultate sa greškom manjom od 25 % (u okviru *FA1.25*) za sve zemlje koje su bile obuhvaćene modelom. U slučaju dvogodišnjeg predviđanja (za 2005-06. godinu), GRNN10 je dao predviđanja iznad *FA1.25* (sa greškom većom od 25 %) u slučaju Luksemburga, Slovenije, Estonije i Rumunije (Slika 28.).

U nekim od navedenih slučajeva greška modela može biti posledica kvaliteta ulaznih i izlaznih podataka koji su korišćeni za kreiranje samog modela. Na primer, GRNN10 model je napravio najveću grešku (43 %) kod predviđanja emisije za Luk-

semburg. Podaci Eurostata pokazuju da je smanjenje PM₁₀ emisije za Luksemburg 2006. u odnosu na 2005. godinu iznosilo 28 %, što je više od 10 puta veće smanjenje nego 2005. u odnosu na 2004. godinu, kada je to smanjenje iznosilo svega 2,6 %. Greška GRNN10 modela za Luksemburg za emisiju u 2005. godini je bila svega 2,6 %, dok je za 2006. greška bila čak 14 puta veća. Takođe treba naglasiti i da je Luksemburg jako specifična država, čija veličina i struktura ekonomije u velikoj meri odstupa od ostalih država obuhvaćenih modelom. Zbog toga je Luksemburg tokom obuke bio predstavljen sa samo 4 % podataka od ukupnog broja koji su korišćeni za obučavanje modela. Iz tog razloga je takođe bila dodatno smanjena sposobnost GRNN modela da dâ tačnija predviđanja.

U slučaju Slovenije i Estonije podaci Eurostata pokazuju takođe neuobičajeno veliko smanjenje u emisiji PM₁₀ čestica tokom 2005. i 2006. godine, koje je iznosilo 22,1 % i 21,6 %, redom. Ovako veliko smanjenje emisije, u poređenju sa ostalim državama, nije moglo da bude očekivano na osnovu podataka korišćenih za obučavanje modela, pa je GRNN10 model dao predviđanja sa velikim odstupanjima.

6.2.6. Analiza značajnosti pojedinih ulaznih promenljivih GRNN10 modela

Kao i u slučaju modelovanja emisije gasova staklene bašte, analiza značajnosti vršena je na osnovu vrednosti pojedinačnih faktora ujednačavanja (*Individual smoothing factor* - ISF), koji su određeni tokom obučavanja genetičkim algoritmom za svaki pojedinačni ulazni parametar GRNN10 modela. Na osnovu ISF vrednosti prikazanih u Tabeli 19. za GRNN10 model, ulazni parametri se mogu podeliti u nekoliko grupa, u odnosu na njihov relativni značaj:

1. grupa najznačajnijih indikatora su BDP i pojedini indikatori koji govore o industrijskoj emisiji PM₁₀ čestica (proizvodnja drvene građe i proizvodnja bakra),
2. druga grupa po važnosti su indikatori koji opisuju proizvodnju energije i saobraćaj, i
3. treću grupu najmanje značajnih čine ostali indikatori industrijske emisije.

6.3. Modelovanje koncentracije PM₁₀ čestica

U ovom poglavlju je ANN metodologija, razvijena tokom modelovanja emisije čestica na nacionalnom nivou, primenjena kod modelovanja izloženosti stanovništva čestičnom zagađenju, koje se kvantifikuje vrednošću PM₁₀ koncentracija u vazduhu. Odabir najznačajnijih ulaznih promenljivih, u odnosu na ulazne promenljive koje su korišćene kod modelovanja emisije PM₁₀ čestica, je vršen korelacionom analizom. Nakon kreiranja odgovarajućeg GRNN modela, razvijeni su i modeli bazirani na višestrukoj linearnoj regresiji (MLR) i regresiji glavnih komponenti (PCR) radi poređenja sa rezultatima GRNN modela. Kao i u slučaju predviđanja emisija PM₁₀ čestica, ANN model se može primeniti kao adekvatan alternativni model kod procene izloženosti stanovništva čestičnom zagađenju u slučajevima kada se predviđanja emisije čestica vrše za zemalje u razvoju ili nerazvijene zemlje, i kada nema dovoljno podataka neophodnih za primenu odgovarajućih linearnih modela [171].

6.3.1. Podaci o PM₁₀ koncentracijama na nacionalnom nivou

Modelovanje koncentracije PM₁₀ čestica na nacionalnom nivou vršeno je sa istim ulaznim promenljivim kao i modelovanje PM₁₀ emisije, jer se pokazalo da su navedene ulazne promenljive dovoljne za njihovo precizno predviđanje, a imajući u vidu da je poznavanje emisije neophodno za adekvatno modelovanje koncentracije zagađujućih materija [172]. Dostupne ulazne promenljive su prikazane u Tabeli 17. Podaci o PM₁₀ koncentracijama u zemljama Evropske unije preuzeti su iz baze podataka Svetske Banke [34], a dobijeni su korišćenjem GMAPS modela. Korišćeni su dostupni podaci od 2000. do 2008. za EU27 i zemlje Evropske unije.

Dostupni podaci (192 serije podataka) podeljeni su na grupu podataka za obučavanje (2000-2006. god.), grupu podataka za proveru (2007-2008. god.) i validacione podatke (23 nasumično odabrane serije podataka iz grupe za obučavanje). Na taj način 67 % podataka je korišćeno za obučavanje, 12 % podataka je primenjeno za validaciju modela, dok je 21 % podataka korišćeno za proveru modela. Statistički podaci o PM₁₀ koncentracijama za svaku od zemalja Evropske unije, kao i za grupu podataka, prikazani su u Tabeli 26.

Tabela 26. Statistička analiza PM₁₀ koncentracija za period 2000-2008. god.

Država/Region	PM10 konc. na nacionalnom nivou [µg/m ³]		
	Minimalna vrednost	Maksimalna vrednost	Srednja vrednost
EU27	20	27	23
Austrija	29	38	34
Belgija	21	28	25
Bugarska	51	73	61
Češka	18	26	22
Danska	16	22	19
Estonija	11	18	15
Finska	15	19	17
Francuska	13	16	15
Nemačka	16	22	19
Grčka	32	48	39
Mađarska	16	24	19
Irska	13	21	17
Italija	23	33	29
Litvanija	17	23	19
Holandija	31	38	34
Poljska	35	40	38
Portugal	21	30	26
Rumunija	12	22	17
Slovačka	13	18	15
Slovenija	28	33	30
Španija	28	40	34
Švedska	10	14	12
Velika Britanija	13	17	15
Grupa podataka za	Minimalna vrednost	Maksimalna vrednost	Srednja vrednost
Obučavanje	11	73	26
Validaciju	12	69	28
Proveru	10	52	21

6.3.2. Razvoj i rezultati optimizacije GRNN modela

Kao što je već opisano u prethodnim poglavljima, većina GRNN parametara zavisi od broja ulaznih/izlaznih promenljivih i od broja serija podataka. Kod modelovanja PM_{10} koncentracije korišćeni su GRNN modeli sa jednim neuronom u izlaznom sloju, dva neurona u sloju za sumiranje, 152 neurona u skrivenom sloju, dok je broj neurona u ulaznom sloju zavisio od broja ulaznih promenljivih koje su korišćene za razvoj modela (broj neurona se kretao od 8 do 13). Ostali parametri, poput funkcije za skaliranje ulaznih podataka i metoda za određivanje faktora ujednačavanja, odabrani su kao i u slučaju modela za PM_{10} emisiju, s obzirom da su dali dobre rezultate. Jedino je promenjena metoda za odabir ulaznih promenljivih, pa su tako u ovom slučaju ulazne promenljive definisane na osnovu korelacione analize (što je bila jedna od metoda za odabir ulaza kod modelovanja gasova staklene bašte). Kao i kod modelovanja gasova staklene bašte i u ovom slučaju su dve različite grupe podataka definisane eliminisanjem manje značajnih ulaznih promenljivih na osnovu rezultata korelacione analize (Tabela 27.):

1. grupa podataka sa međusobno „nekorelisanim“ ulaznim promenljivim koja sadrži samo one promenljive koje su međusobno korelisane sa koeficijentom manjim od 0,8,
2. grupa podataka sa ulaznim promenljivim koje su korelisane sa izlaznom promenljivom sa koeficijentima korelacije većim od 0,1.

Rezultati korelacione analize (Tabela 27.) pokazuju da su proizvodnja oblovine (V7), rezane drvne građe (V8), proizvodnja papira (V6) i spaljivanje drveta (V3) značajno međusobno korelisani. Da bi se dobila grupa podataka sa međusobno „nekorelisanim“ ulaznim promenljivim, indikatori koji se odnose na proizvodnju oblovine i rezane drvne građe su eliminisani. Da bi se razvio GRNN model sa ulazima značajno korelisanim sa PM_{10} koncentracijom bilo je potrebno eliminisati 5 ulaznih promenljivih: proizvodnja uglja (V5), bakra (V9), aluminijuma (V10), gvožđa (V11) i NPK đubriva (V13).

Tabela 27. Rezultati korelacione analize podataka za razvoj modela za predviđanje PM₁₀ koncentracija

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13
V1 ¹	1,00												
V2	0,63	1,00											
V3	-0,03	0,41	1,00										
V4	0,67	0,35	-0,04	1,00									
V5	-0,55	-0,16	0,20	-0,37	1,00								
V6	0,37	0,73	0,60	0,20	-0,18	1,00							
V7	0,14	0,63	0,86	0,00	0,11	0,85	1,00						
V8	0,23	0,64	0,82	0,10	0,06	0,86	0,94	1,00					
V9	0,29	0,61	0,21	0,17	-0,14	0,52	0,42	0,42	1,00				
V10	0,08	0,02	-0,12	0,15	-0,30	0,05	-0,09	-0,05	-0,24	1,00			
V11	0,23	0,57	-0,01	0,12	-0,17	0,40	0,22	0,32	0,61	0,09	1,00		
V12	0,21	0,60	0,05	0,17	-0,18	0,44	0,25	0,34	0,62	0,16	0,78	1,00	
V13	-0,15	0,03	0,00	-0,04	-0,11	-0,01	-0,02	0,00	0,16	-0,22	0,08	0,06	1,00
PM10	-0,29	-0,38	-0,25	-0,13	0,06	-0,22	-0,34	-0,32	-0,02	0,01	-0,06	-0,13	-0,01

¹ ulazne promenljive prikazane bold slovima u koloni 1 su korišćene za razvoj CA-GRNN i CA-PCR modela

Vrednosti pokazatelja performansi za navedene GRNN modele, kreirane na bazi korelacione analize, prikazani su u Tabeli 28., kao i vrednosti statističkih pokazatelja performansi za neoptimizovani GRNN model (NO-GRNN model sa svim raspoloživim ulaznim promenljivim).

Tabela 28. Vrednosti pokazatelja performansi za razvijene GRNN modele za predviđanje PM₁₀ koncentracija

Model	Statistički pokazatelji performansi modela			
	IA	FA1.25 [%]	RMSE [$\mu\text{g}/\text{m}^3$]	MAE [$\mu\text{g}/\text{m}^3$]
IV-GRNN	0,98	95	2,54	2,29
CA-GRNN	0,99	100	2,29	2,05
NO-GRNN	0,98	93	3,09	2,44
CA-PCR	0,72	35	8,88	4,07

Iako su sva tri GRNN modela pokazala dobre performanse kod modelovanja PM₁₀ koncentracije na nacionalnom nivou, najbolja predviđanja su ostvarena sa CA-GRNN modelom ($IA=0,99$ i $FA1.25=100\%$). Kako neoptimizovani GRNN model (NO-GRNN) ima vrednosti $IA=0,98$ i $FA1.25=93\%$, može se zaključiti da je dobijen precizniji model eliminacijom ulaznih promenljivih na bazi korelacione analize. Takođe optimizovani CA-GRNN model ima čak 38 % manje ulaznih promenljivih što značajno umanjuje grešku modela kada se on koristi sa ulaznim promenljivim čije vrednosti su procenjene odgovarajućim modelima, pri čemu se smanjuje i vreme potrebno za prikupljanje i pripremu podataka.

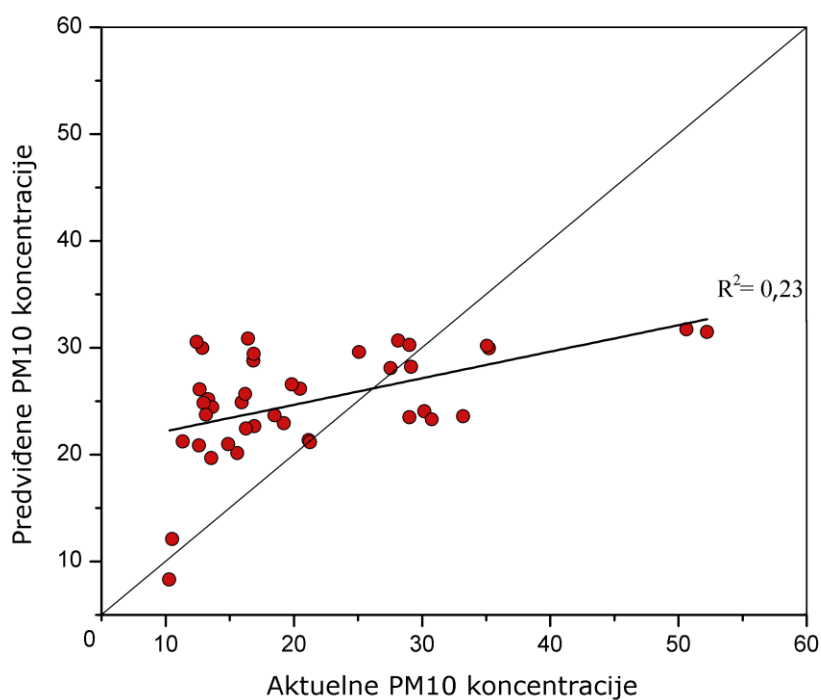
6.3.3. Poređenje GRNN modela za predviđanje PM₁₀ koncentracije sa linearnim modelom

Kako je CA-GRNN model dao najbolje rezultate, radi poređenja kreiran je i odgovarajući linearni PCR model (CA-PCR model), sa istim podacima kao i CA-GRNN model. Tokom kreiranja CA-PCR modela vršena je validacija sa 23 serije podataka (kao i u slučaju GRNN modela). Izdvajanja validacionih podataka i kreiranje PCR modela vršeno je XLSTAT softverskim paketom [173]. Koeficijenti kreiranog CA-PCR modela dati su u Tabeli 29., dok su vrednosti statističkih pokazatelja performansi date u Tabeli 28.

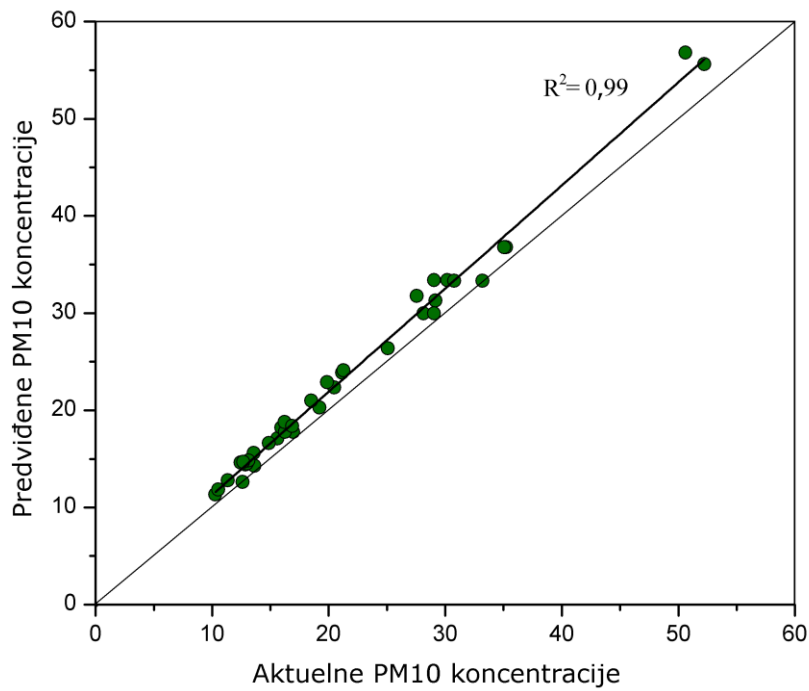
Tabela 29. Koeficijenti razvijenog CA-PCR modela

Promenljive modela	Koeficijenti	Standardna greška
Konstanta	34,907	19,569
V1	-7,217	3,491
V2	-1,240	1,857
V3	4,356	6,519
V4	9,152	11,452
V6	13,762	3,758
V7	-3,753	1,355
V8	0,285	4,375
V11	-3,181	5,135

Na Slikama 29. i 30. prikazano je poređenje rezultata CA-PCR i CA-GRNN modela dobijenih proverom sa aktuelnim vrednostima PM₁₀ koncentracija. Poređenje rezultata CA-GRNN i CA-PCR modela jasno pokazuje da GRNN model ima znatno bolje performanse kod predviđanja PM₁₀ koncentracije. Sva predviđanja CA-GRNN modela imala su grešku u okviru *FA1.25* pokazatelja, dok je CA-PCR imao svega 35 % predviđanja u okviru *FA1.25*.

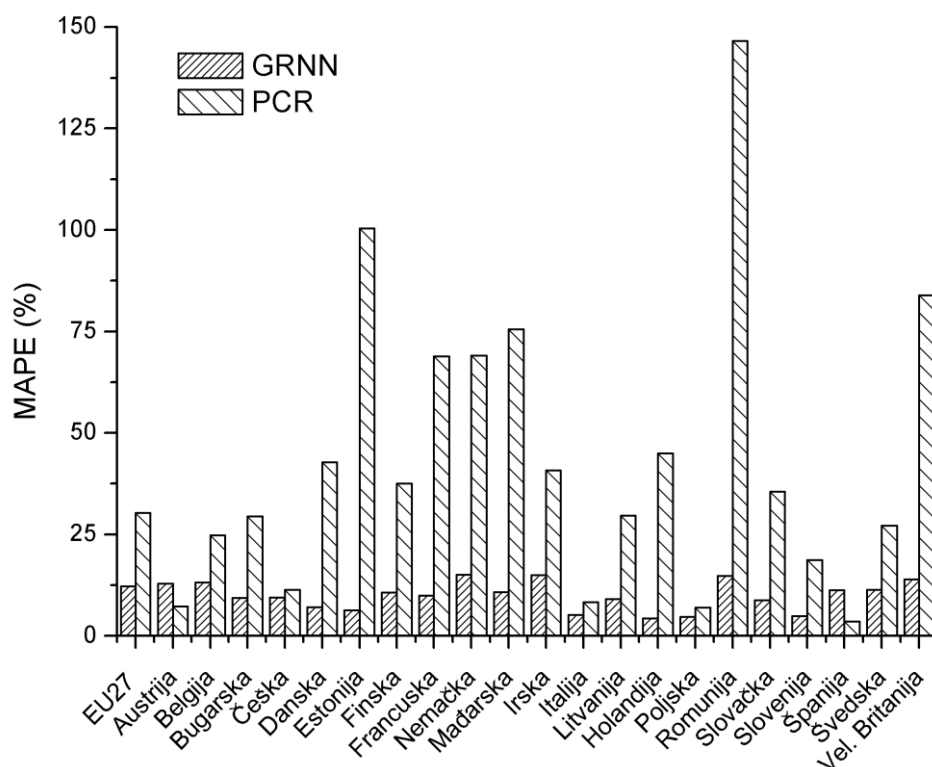


Slika 29. Poređenje rezultata CA-PCR modela sa aktuelnim vrednostima PM₁₀ koncentracija



Slika 30. Poređenje rezultata CA-GRNN modela sa aktuelnim vrednostima PM₁₀ koncentracija

Vrednosti srednje relativne greške (MAPE) za CA-GRNN i CA-PCR model prikazane su na Slici 31., za one evropske zemlje za koje su postojali podaci za obe godine kod provere modela.



Slika 31. MAPE vrednosti predviđanja CA-GRNN i CA-PCR modela za pojedinačne EU države

Kao što se može uočiti na Slici 31., CA-GRNN model je dao znatno bolje rezultate za EU27 (greška manja od 13 %) i većinu analiziranih evropskih država. Samo u slučaju Austrije i Španije CA-PCR model je dao predviđanja sa manjom srednjom relativnom greškom u odnosu na CA-GRNN model, ali i u tim slučajevima greška GRNN modela je bila manja od 13 %.

6.4. Modelovanje količina generisanog komunalnog čvrstog otpada

U ovom poglavlju prikazan je razvoj i evaluacija ANN modela za predviđanje količine generisanog komunalnog čvrstog otpada (KČO) na nacionalnom nivou. ANN modeli su kreirani na bazi dve različite ANN arhitekture, pri čemu su kao ulazne promenljive korišćeni dostupni podaci o ekonomskom razvoju i industrijskoj strukturi 26 evropskih zemalja, kako razvijenih tako i zemalja u razvoju. Modeli koji se mogu koristiti za predviđanja količina KČO na nacionalnom nivou posebno su značajni za zemlje u razvoju, zbog činjenice da se u takvim zemljama upravljanje otpadom ne vrši na zadovoljavajućem nivou, pa samim tim nedostaju podaci o količinama generisanog otpada. Zbog toga je primena razvijenog modela posebno analizirana na primerima zemalja u razvoju, Bugarske i Srbije, pri čemu su u slučaju Srbije za evaluaciju korišćeni procenjeni podaci Agencije za zaštitu životne sredine, jer se količina KČO koja je dovezena na deponije ne meri na celokupnoj teritoriji Republike Srbije [174].

6.4.1. Ulazni indikatori i priprema podataka

Kao što je u prethodnim poglavljima pokazano, ANN modeli na bazi ekonomskih i indikatora održivog razvoja pokazuju veliku preciznost kod modelovanja indikatora kvaliteta životne sredine. Zbog toga je i ANN model za predviđanje količine generisanog čvrstog otpada kreiran na osnovu ulaznih promenljivih koje opisuju ekonomsko stanje i potrošnju (Tabela 30.) u odabranim državama Evropske unije i u Srbiji.

Tabela 30. Ulazne i izlazne promenljive modela sa jedinicama posle normalizacije

Ulazna promenljiva	Jedinica
Bruto domaći proizvod (BDP) po stanovniku	podeljen sa BDP po stan. (EU27)
Domaća potrošnja materijala (DPM)	kg po stan.
Produktivnost resursa (PR)	€/kg
Izlazna promenljiva	Jedinica
Generisani komunalni čvrsti otpad (KČO)	kg po stan.

Indikator domaće potrošnje materijala (DPM) definisan je kao ukupna količina materijala koji se koristi od strane privrede. DPM je jednak direktnoj potrošnji materijala u privredi umanjenoj za izvoz, pa je na taj način DPM mera stvarne potrošnje materijala u privredi. Produktivnost resursa (PR) predstavlja količnik bruto domaćeg proizvoda i potrošnje materijala u privredi.

Kako ANN modeli postižu bolje rezultate sa normalizovanim vrednostima, i u ovom slučaju je BDP normalizovan po stanovniku i sa vrednošću BDP Evropske unije, dok su DPM i generisani KČO normalizovani po stanovniku. Obučavanje i provera modela vršeno je korišćenjem podataka Eurostata [175-178] za 26 evropskih zemalja, EU27 i „stare“ EU15. Podaci za period 2000-2002. god. su korišćeni za obučavanje, dok su podaci od 2003. do 2005. god. korišćeni za proveru razvijenih modela. Primer grupe podataka za obučavanje za 2000. godinu prikazan je u Tabeli 31.

Tabela 31. Podaci iz 2000. godine korišćeni za obučavanje modela

Region/ Država	BDP	DPM [t po stan.]	PR [€/kg]	KČO [kg po stan.]
EU27	1,00	15,480 ^a	1,23 ^c	524
EU15	1,21	20,210 ^a	1,41 ^c	569
Belgija	1,29 ^b	18,856 ^a	1,30 ^c	474
Bugarska	0,09	12,282	0,14	516
Češka	0,31	18,039	0,33	334
Danska	1,70	25,282	1,29	665
Nemačka	1,31	17,728	1,42	643 ^d
Estonija	0,24	21,753	0,20	440 ^d
Irska	1,45	31,368	0,88	603
Grčka	0,66	16,080	0,79	408
Španija	0,82	17,244	0,91	662
Francuska	1,24	14,488	1,64	516
Italija	1,09	16,943	1,23	509
Kipar	0,76	26,333 ^a	0,55 ^c	680
Letonija	0,19	8,075	0,44	270 ^d
Litvanija	0,18	8,223	0,43	363
Mađarska	0,26	10,927	0,47	445
Austrija	1,36	19,006	1,36	581
Poljska	0,26	13,190	0,36	316
Portugal	0,63	18,475	0,65	472
Rumunija	0,09	7,658	0,23	363
Slovenija	0,57	17,024	0,63	513 ^d
Slovačka	0,21	9,635	0,42	254 ^e
Finska	1,34	34,471	0,74	503
Švedska	1,57	13,444	2,24	428
Velika Britanija	1,42	11,906	2,29	578
Norveška	2,13	20,739	1,97	615 ^e
Švajcarska	1,97	12,441	3,04	657

^aprocenjena vrednost [175]

^bprocenjena vrednost [176]

^cprocenjena vrednost [177]

^dprocenjena vrednost [178]

^eodređeno nestandardnom metodologijom [178]

6.4.2. Ulazni podaci za Republiku Srbiju

U slučaju Republike Srbije korišćeni su dostupni podaci Agencije za zaštitu životne sredine (SEPA) [179] i Republičkog zavoda za statistiku [180-182]. Kako Republički zavod za statistiku nema podatke o domaćoj potrošnji materijala, DPM je izračunat na osnovu formule koju su definisali *Weisz* i saradnici 2006. godine [183]: domaća potrošnja materijala je jednaka domaćoj ekstrakciji materijala plus uvoz minus izvoz materijala. Podaci potrebni za proračun domaće potrošnje materijala prikazani su u Tabeli 32. [183].

Tabela 32. Kategorije materijala potrebne za proračun DPM [183]

Glavne kategorije	Podkategorije	Opis
Biomasa	Hrana	Biomasa koja se koristi za prehranu ljudi
	Stočna hrana	Biomasa koja se koristi kao stočna hrana
	Životinje	Uključuje ulovljene divlje životinje i ribu, kao i proizvode životinjskog porekla
	Drvo	Drvena građa i proizvodi od drveta – papir, nameštaj i sl.
	Ostalo	Vlakna i ostali proizvodi od biomase
Fosilna goriva	Ugalj	Svi tipovi uglja
	Nafta	Svi tipovi nafte
	Prirodni gas	Svi tipovi prirodnog gasa
	Ostalo	Proizvodi na bazi fosilnih goriva
Industrijski minerali	Industrijski minerali	Svi tipovi metalnih ruda i proizvodi od metala
	Rude	Svi nemetalni minerali koji se koriste u industriji (izuzev fosilnih goriva)
Konstrukcioni materijali	Konstrukcioni materijali	Svi konstrukcioni materijali

Da bi se eliminisao uticaj kursa dinara prema evru, BDP Republike Srbije preračunat je na osnovu vrednosti cena u 2006. godini, koja je bazna godina u odnosu na period (2006-2009. god.) za koji su prikupljeni podaci (Tabela 33.).

Tabela 33. Ulazni podaci za Republiku Srbiju

Godina	Ulazni indikatori		
	BDP	DPM [tpo stan.]	PR €/kg
2006	0,13	12,216	0,26
2007	0,13	12,321	0,27
2008	0,14	13,397	0,26
2009	0,15	12,624	0,27

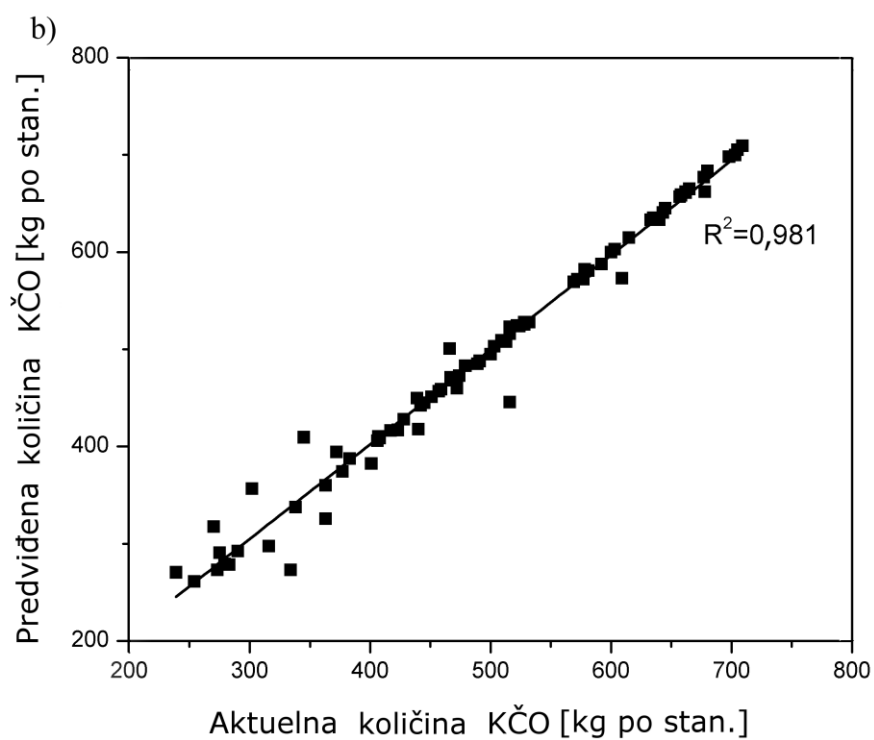
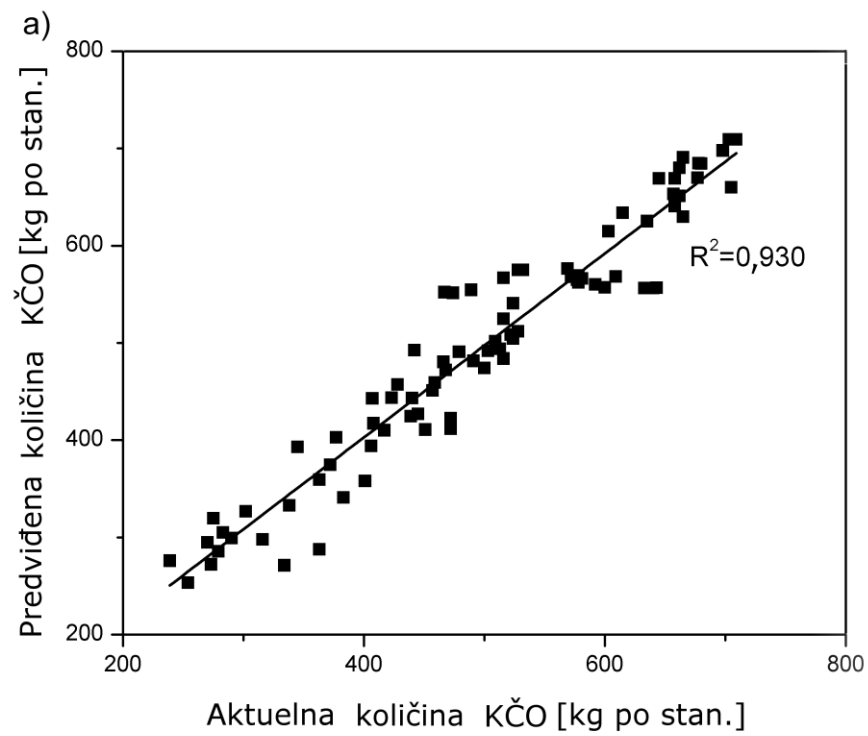
6.4.3. Razvoj ANN modela za predviđanje količine KČO

Za razvoj modela korišćen je softver NeuroShell 2, koji je već dao dobre rezultate kod modelovanja izluživanja solidifikovanog otpada [57]. Pored neuronske mreže sa opštom regresijom (GRNN), koja je u prethodnim poglavljima korišćena za modelovanje emisija i koncentracija gasovitih i čvrstih materija u vazduhu, razvijeni su i modeli na bazi neuronske mreže sa propagacijom greške unazad (*Backpropagation neural network* - BPNN), koja je odabrana kao jedna od najčešće korišćenih arhitektura.

Kod obe arhitekture, broj neurona u ulaznom i izlaznom sloju je definisan brojem ulaznih i izlaznih promenljivih, pa oba modela imaju tri neurona u ulaznom i jedan neuron u izlaznom sloju. Što se tiče skrivenog sloja, GRNN ima onoliko neurona koliko ima serija podataka kod obučavanja (u ovom slučaju 84), dok je broj skrivenih neurona BPNN modela podešen na 10, što je standardna vrednost u slučaju kada se jedan izlaz modeluje sa tri ulaza. Broj neurona u sloju za sumiranje GRNN modela je jednak broju izlaza plus jedan (u ovom slučaju ima 2 neurona).

6.4.4 Analiza rezultata razvijenih ANN modela

Performanse razvijenih GRNN i BPNN modela prikazane su na Slici 32., gde je dato poređenje rezultata modela sa aktuelnim vrednostima KČO. Uočena je i razlika u brzini obučavanja modela, koji je kod GRNN trajao nekoliko minuta, dok je u slučaju BPNN trajao više sati.



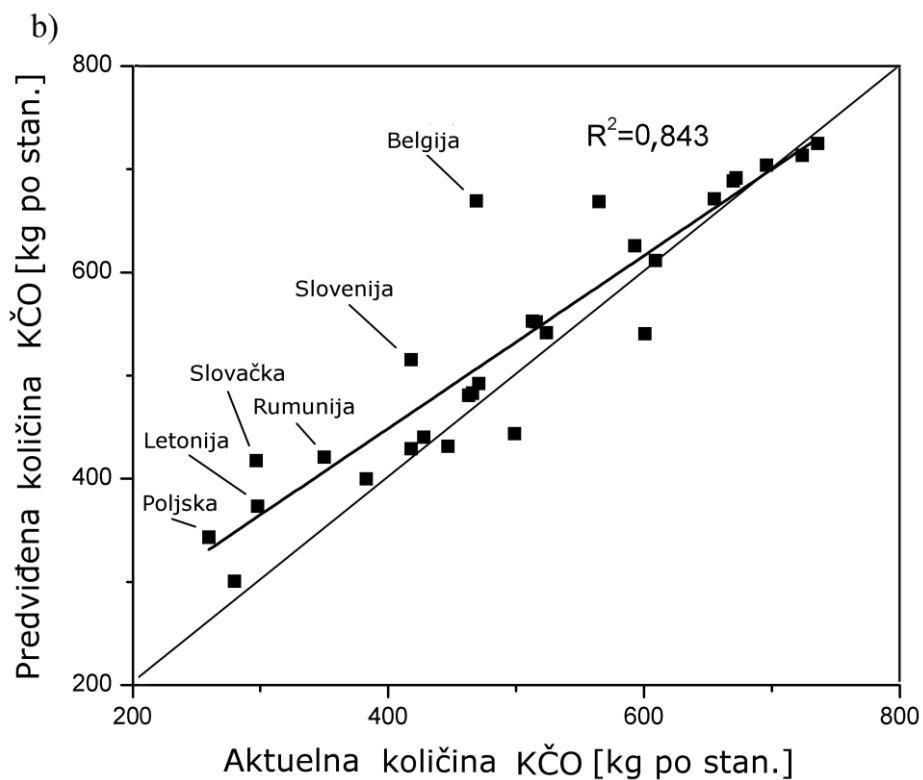
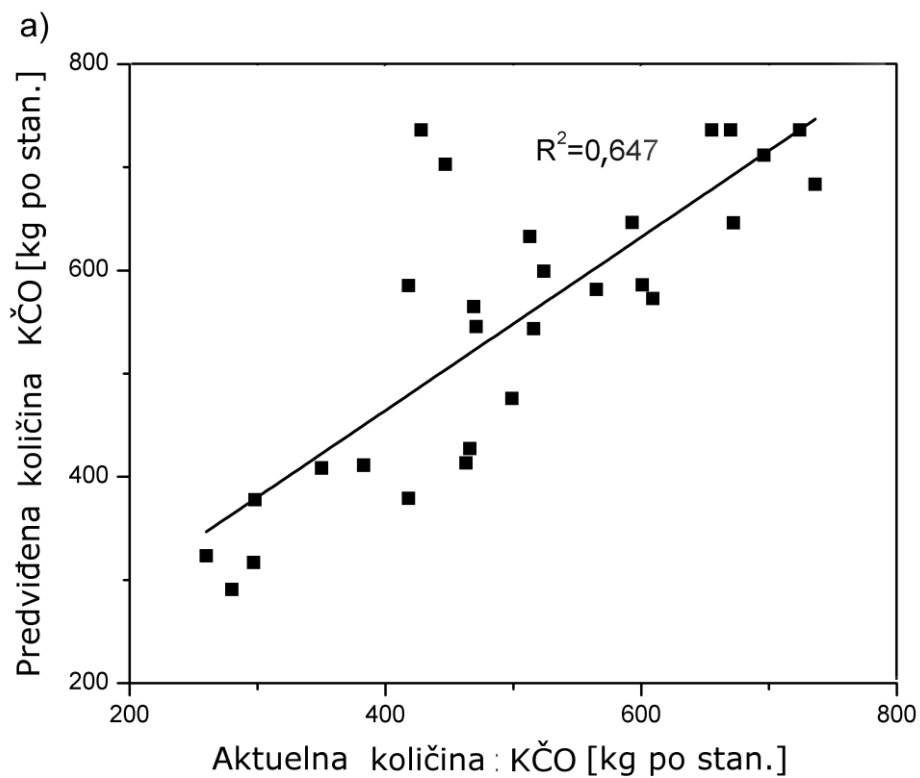
Slika 32. Poređenje aktuelnih vrednosti generisanog KČO sa predviđanjima modela tokom obučavanja: a) BPNN i b) GRNN

Dobijeni rezultati oba modela i korelacioni faktori pokazuju da su mreže zadovoljavajuće obučene. Rezultati GRNN modela sa R^2 od 0,981 su nešto bolji od rezultata BPNN modela ($R^2=0,930$). Ukoliko se posmatraju rezultati za pojedinačne zemlje, vidi se da je primenom GRNN modela 74 % predviđanja bilo sa relativnom greškom manjom od 1 %, dok je u slučaju:

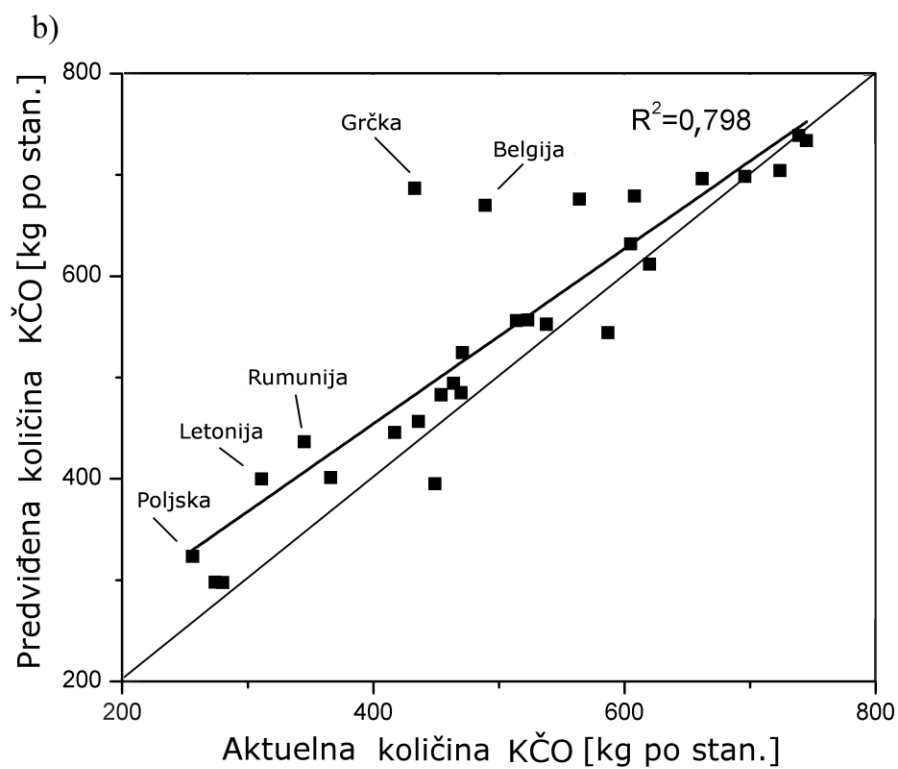
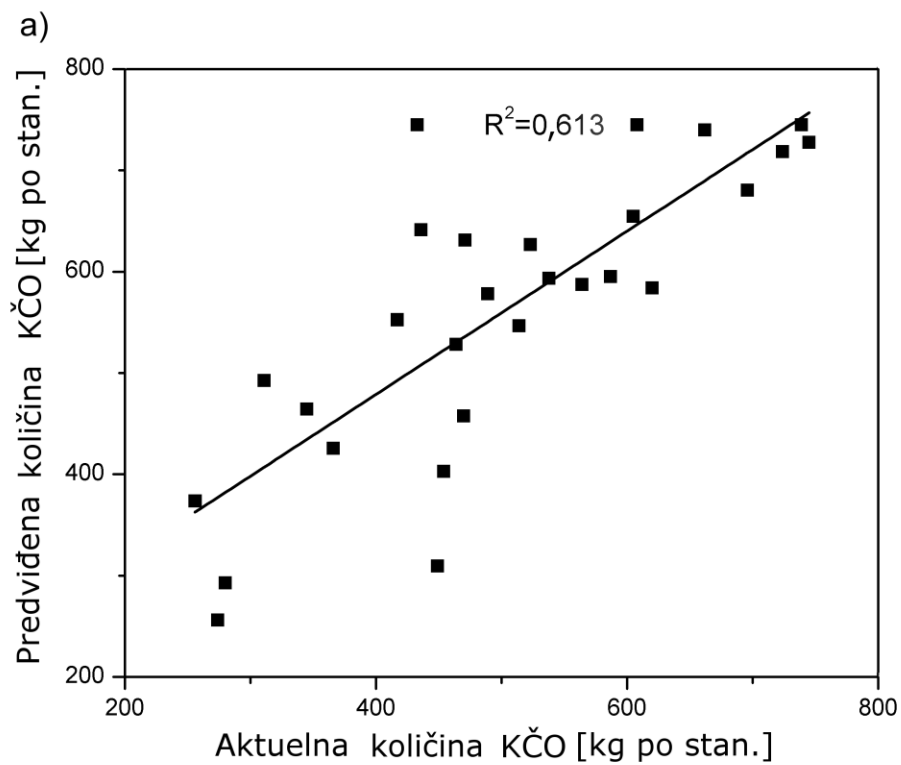
- Danske, Irske i Norveške predviđanje bilo bez greške za sve tri godine,
- Češke, Finske, Švedske, Italije, Mađarske, Francuske, Švajcarske i Evropske unije predviđanje bilo bez greške za dve godine, i
- Velike Britanije, Kipra, Nemačke i Austrije predviđanje bilo bez greške samo za jednu godinu.

Predviđanja BPNN modela sa greškom manjom od 1 % su dobijena samo u slučaju Kipra (za sve tri godine), Estonije i Švajcarske (za dve godine), Rumunije, Slovačke, Češke, Irske, Finske, Švedske, Norveške i Evropske unije (za jednu godinu).

Proverom modela (podaci za period 2003-2005.) ispitane su performanse modela kod predviđanja količine generisanog KČO u posmatranim evropskim zemljama. Rezultati za 2003. i 2004. godinu prikazani su na Slikama 33. i 34., dok su rezultati za 2005. godinu prikazani u Tabeli 34.



Slika 33. Poređenje aktuelnih količina generisanog KČO sa predviđanjima modela tokom provere razvijenih modela (2003 godina): a) BPNN i b) GRNN



Slika 34. Poređenje aktuelnih količina generisanog KČO sa predviđanjima modela tokom provere razvijenih modela (2004 godina): a) BPNN i b) GRNN

Tabela 34. Aktuelne količine i predviđanja modela za 2005. godinu (provera modela)

Region/ Država	BDP	DPM	PR	Aktuelna vrednost	KČO [kg po stan.]			Rel. greška [%]
		[t po stan.]	[€/kg]		BPNN	Rel. greška [%]	GRNN	
EU27	1,00	15,746 ^a	1,30 ^c	517	577	11,79	525	1,52
EU15	1,19	19,483 ^a	1,55 ^c	558	601	7,81	572	2,51
Belgija	1,28	18,262 ^a	1,43 ^c	482	594	23,36	610	26,5
Bugarska	0,12	15,213	0,15	475 ^d	600	26,47	500	5,26
Češka	0,44	18,622	0,39	289	347	20,17	273	5,42
Danska	1,70	27,961	1,22	737	729	0,96	665	9,77
Nemačka	1,21	15,726	1,64	564	624	10,72	516	8,51
Estonija	0,37	22,645	0,29	436	318	26,90	406	6,88
Irska	1,74	32,218	1,04	740	759	2,57	698	5,68
Grčka	0,78 ^b	17,306	0,88	438	652	48,91	662	51,1
Španija	0,93	19,993	0,86	597	759	27,14	645	8,04
Francuska	1,22	13,579	1,84	532	666	25,25	528	0,75
Italija	1,08	14,230	1,49	542	550	1,53	574	5,90
Kipar	0,80	30,047	0,52	739	724	2,00	709	4,06
Letonija	0,25	10,531	0,52	310	461	48,81	444	43,3
Litvanija	0,27	9,886	0,53	376	460	22,28	287	23,6
Mađarska	0,39	16,437	0,39	460	305	33,70	340	26,1
Austrija	1,32	20,629	1,33	620	594	4,24	567	8,47
Poljska	0,28	14,240	0,40	319 ^d	355	11,37	295	7,57
Portugal	0,63	17,180	0,71	446	641	43,66	434	2,60
Rumunija	0,16	15,442	0,16	377 ^d	350	7,05	500	32,6
Slovenija	0,64	18,792	0,68	423	561	32,51	460	8,79
Slovačka	0,32	12,545	0,41	289	289	0,00	430	48,8
Finska	1,33	39,173	0,73	479	547	14,29	459	4,18
Švedska	1,45	18,319	1,83	482	577	19,68	633	31,3
Velika Britanija	1,35	11,652	2,58	585	717	22,53	600	2,56
Norveška	2,34	20,986	2,11	759	732	3,62	635	16,3
Švajcarska	1,79	12,333	3,16	663	486	26,75	658	0,75

^aprocenjena vrednost [175]

^bprocenjena vrednost [176]

^cprocenjena vrednost [177]

^dprocenjena vrednost [178]

Prilikom provere, GRNN model je dao mnogo bolja predviđanja (sa R^2 od 0,798 do 0,843) u poređenju sa BPNN modelom (R^2 od 0,613 do 0,647). BPNN model je samo za 44 % predviđanja imao grešku manju od 10 %. Za sve tri godine BPNN model je imao zadovoljavajuće tačna predviđanja samo u slučaju Kipra, Austrije, Danske, Slovačke, Norveške i Evropske unije. Predviđanja GRNN modela su bila zadovoljavajuće tačna u 69 % slučajeva (rel. greška manja od 10 %), pri čemu posebno treba istaći da tačnost predviđanja nije značajnije opadala tokom provere za različite godine, što je bio slučaj sa BPNN modelom (Tabela 35.).

Tabela 35. Procenat predviđanja sa relativnom greškom manjom od 10 %

Model	Godina		
	2003	2004	2005
BPNN	57%	43%	32%
GRNN	71%	68%	68%

Na osnovu dobijenih rezultata tokom provere modela, može se zaključiti da je GRNN model dao tačnija predviđanja količina generisanog KČO, zbog čega će se analize u narednim poglavljima odnositi samo na GRNN model.

6.4.5 Analiza grešaka GRNN modela

Odstupanja količina KČO dobijenih GRNN modelom u odnosu na aktuelne količine KČO u velikom broju slučajeva mogu se dovesti u vezu sa kvalitetom ulaznih podataka korišćenih za obučavanje. U slučaju Belgije, za koju je GRNN model napravio velike relativne greške za sve tri godine kod provere modela, vrednosti ulaznih promenljivih, osim za BDP i KČO, su dobijene procenom (Tabela 31.), što je mogući uzrok velikih relativnih grešaka tokom provere (Tabela 34., Slike 33. i 34.). U slučaju Letonije i Slovenije procenjeni podatak je sama količina generisanog KČO (Tabela 31.), što takođe može biti uzrok velikih relativnih grešaka prilikom provere.

U slučaju Slovačke, podaci o KČO su sve do 2001. godine prikupljeni u skladu sa Nacionalnim katalogom otpada, a od 2002. podaci su prikupljeni u skladu sa Evropskim katalogom otpada [178]. Kada se ova činjenica uzme u obzir, može se

zaključiti da podaci korišćeni za obučavanje za Slovačku za 2000. i 2001. godinu nisu bili dobijeni identičnom metodom, što je dovelo do velikih grešaka tokom provere modela, od čak 48,8 % za 2005. godinu (Tabela 34.). Takođe, i u Norveškoj je prikupljanje otpada i podataka o količinama KČO bilo po drugačijoj metodologiji u odnosu na većinu EU zemalja sve do 2001. godine [178], što je tokom obučavanja i provere dovelo do značajnijih relativnih odstupanja modelom predviđenih količina KČO za Norvešku u odnosu na aktuelne podatke, npr. 16,3 % za 2005. godinu (Tabela 34.)

U slučaju Poljske i Rumunije, podaci o količinama KČO za test godine dobijeni su procenom, a ne merenjem [178]. Ova činjenica takođe može biti objašnjenje za velika odstupanja predviđenih i zvaničnih vrednosti generisanog KČO za navedene zemlje prilikom provere GRNN modela (Slike 33., 34. i Tabela 34.).

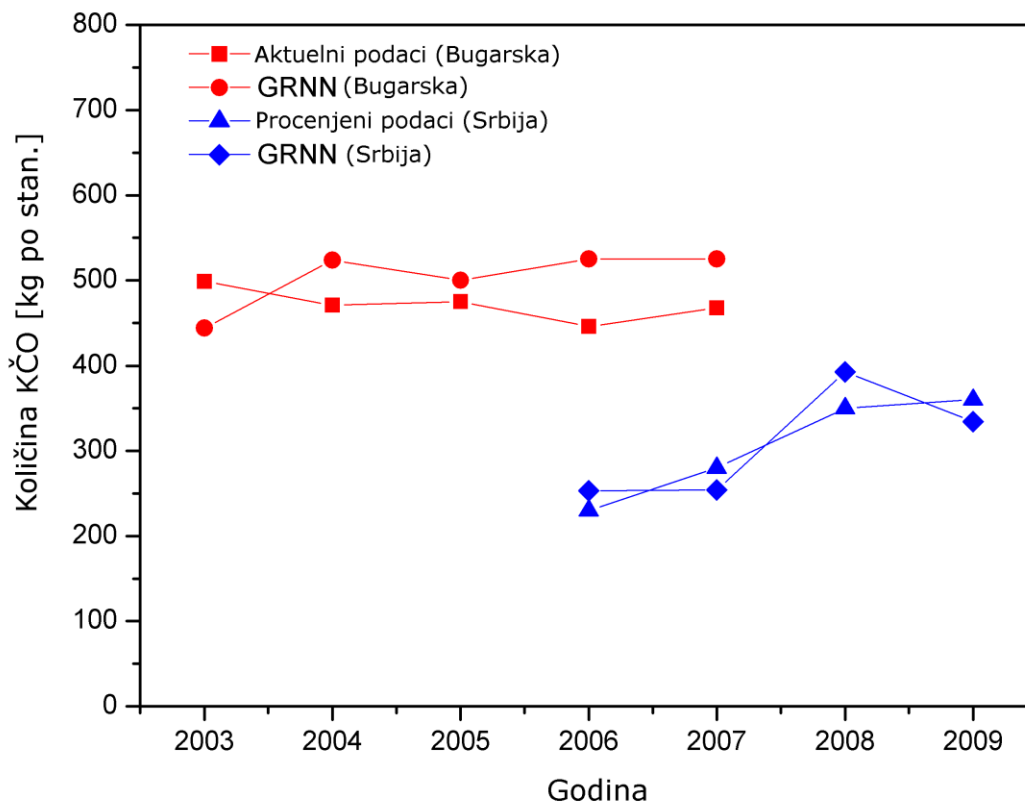
Indikativan je i slučaj Grčke, za koju su BDP podaci na Eurostatu za 2004. i 2005. godinu označeni kao „privremeni“, što je posledica netačnih izveštaja koje je Grčka vlada dostavljala evropskim institucijama. Za navedene godine GRNN model je napravio relativnu grešku od čak 50 %, dok je za 2003. godinu, za koju se BDP Grčke ne dovodi u pitanje, GRNN dao predviđanja količina KČO koja se od aktuelne vrednosti razlikuje za samo 2,9 %.

U slučaju Mađarske, iako je konstantovan napredak u kvalitetu podataka koji su dostavljani počev od 2000. godine, kada je uvedeno merenje količina otpada doveženog na deponije [178], podaci o generisanom KČO ne mogu se smatrati pouzdanim, pa samim tim mogu biti uzrok odstupanja GRNN modelom predviđenih količina generisanog KČO. Takođe, u trenutku razvoja GRNN modela, podaci o KČO za Mađarsku za period 2001-2003. god. nisu bili dostupni ni na sajtu EEA [184].

6.4.6. Primena GRNN modela na Srbiju

Od evropskih zemalja na koje je primenjen GRNN model, Bugarska se može smatrati najbližnjom Srbiji, zbog relativno niskog stepena razvijenosti, geografske lokacije, broja stanovnika i sl. Zbog toga je model dodatno testiran na primeru Bugarske, pre primene na podatke iz Srbije. GRNN model primenjen je na ulazne po-

datke prikupljene za period 2006-2007. god. za Bugarsku, a rezultati su prikazani na Slici 35.



Slika 35. Poređenje aktuelnih vrednosti količina KČO sa predviđanjima GRNN modela za Bugarsku i Srbiju

Predviđanja GRNN modela, primenjenog na podatke iz 2006. i 2007. godine za Bugarsku, bila su slična zvaničnim vrednostima količina KČO, pri čemu je maksimalna relativna greška iznosila 17 % za 2006. godinu. Uzimajući u obzir da je od oktobra 2005. do decembra 2007. deponija u Sofiji bila zatvorena, pri čemu je vršeno privremeno skladištenje otpada u količinama od 2,5 % u 2005. godini do 10 % u 2006. i 2007. godini od ukupnih generisanih količina čvrstog otpada [178], može se zaključiti da je stvarna relativna greška GRNN modela za 2006. godinu zapravo 7%, dok za ostale 2 godine greške zapravo nije ni bilo.

Pošto je na primeru Bugarske pokazano da se GRNN model može uspešno koristiti za predviđanja količine generisanog KČO na zemlje u razvoju, GRNN model je pri-

menjen na pripremljene podatke za Srbiju za period 2006-2009. godina. Treba napomenuti da su podaci o KČO za Srbiju (SEPA [179]), za razliku od većine Evropskih zemalja, dobijeni procenom, a ne merenjem količina otpada, što unosi nesigurnost u analizu rezultata modela.

Procenjene ukupne količine otpada za posmatrani period, dobijene su na osnovu merenja generisanih količina otpada u deset opština u periodu od sedam dana više puta tokom godine. Pri tom, učestalost merenja u tim opštinama nije bila ista svake godine [179]. Rezultati dobijeni GRNN modelom pokazuju da navedena metodologija, koju SEPA koristi za procenu ukupnih količina KČO, daje podatke zadovoljavajuće tačnosti. Dobro slaganje, sa relativnom greškom manjom od 10 %, između procenjenih SEPA vrednosti i predviđanja GRNN modela zabeleženo je za 2006., 2007. i 2009. godinu, dok je relativna greška za 2008. godinu bila 12 % (Tabela 36. i Slika 35.).

Tabela 36. Procenjene i GRNN vrednosti količina generisanog KČO sa relativnim greškama

Godina	SEPA vrednosti [kg po stan.]	GRNN predviđanja [kg po stan.]	Relativna greška [%]
2006	233	253	8,33
2007	280	254	-9,53
2008	350	393	12,22
2009	360	334	-7,22

Kao što se može videti u Tabeli 36., u periodu od 2006. do 2008. godine došlo je do porasta količine generisanog KČO od 50 %, što se ne može objasniti samo porastom potrošnje u Srbiji, već u velikoj meri predstavlja poboljšanje u sistemu prikupljanja otpada, a time i tačnosti podataka o količinama nastalog otpada. Novi Zakon o upravljanju otpadom stupio je na snagu 2009. godine [185] i očekuje se da će njegovim sprovođenjem dodatno biti unapređen kvalitet podataka o generisanim količinama KČO u Srbiji.

6.5. Modelovanje sadržaja rastvorenog kiseonika u površinskim vodama

U ovom poglavlju je opisan razvoj i evaluacija tri ANN modela, baziranih na različitim ANN arhitekturama, za predviđanje sadržaja rastvorenog kiseonika u reci Dunav, kao i njihovo poređenje sa linearnim modelima. Pored odabira najpreciznije ANN arhitekture, cilj je i razvoj modela koji bi, kao ulazne promenljive, koristio samo nespecifične parametre kvaliteta voda. Pored toga što se nespecifični parametri kvaliteta voda lakše, brže i češće prate tokom redovnih ili vanrednih monitoringa kvaliteta voda, prednost ANN model sa nespecifičnim ulaznim parametrima je i znatno šira potencijalna primena, u poređenju sa modelima koji se kreiraju sa specifičnim parametrima kvaliteta voda [186]:

- za dopunu izveštaja o kvalitetu voda - predviđanje vrednosti rastvorenog kiseonika u slučajevima kada rastvoreni kiseonik iz nekog razloga nije bio određen;
- za poboljšanje monitoringa kvaliteta voda - predviđanje rastvorenog kiseonika na monitoring stanicama na kojima se određivanje rastvorenog kiseonika ne vrši;
- za brzo predviđanje rastvorenog kiseonika u akcidentnim i vanrednim situacijama, posebno na manjim rekama na kojima se ne vrši kontinualni monitoring - npr. u situacijama kada dođe do masovnog uginuća riba može se odrediti da li je do pomora došlo usled nedostatka kiseonika izazvanog specifičnim meteorološkim uslovima, ili je do uginuća došlo usled toksičnog dejstva supstanci prisutnih u ispuštenim otpadnim vodama.

6.5.1. Odabir ulaznih promenljivih modela za predviđanje sadržaja rastvorenog kiseonika

Republički hidrometeorološki zavod Srbije (RHMZ) na 588 km, koliki je tok Dunava kroz Srbiju, ima 17 stanica na kojima prati kvalitet vode. Na Slici 36. je prikazan položaj monitoring stanica, uspostavljenih počev od 1965. god. pa do 2000. god., na kojima su uzorci uzimani 10 do 24 puta godišnje (mesečni i polumesečni režim).

Fizičko-hemijske i biološke analize vrši RHMZ, a dostupni izveštaji o kvalitetu vode od 2004. do 2009. godine [146-150] su korišćeni za kreiranje modela.



Slika 36. Lokacije monitoring stanica na Dunavu (Srbija)

Za reprezentativnu monitoring stanicu odabrana je Bezdán stanica, pošto je ona jedna od najstarijih, i za nju postoji najviše podataka u posmatranom periodu.

Izbor odgovarajućih ulaznih promenljivih je jedan od najvažnijih koraka u kreiranju ANN modela. I u ovom slučaju, izbor promenljivih je vršen na osnovu dostupnih literaturnih podataka o uticaju ostalih parametara kvaliteta voda na koncentraciju kiseonika u vodama. Rezultati prethodnih istraživanja pokazuju da je kod ANN modelovanja sadržaja rastvorenog kiseonika najvažniji ulazni parametar pH vrednost [58, 59] i temperatura vode [59]. Pored ovih podataka, za modelovanje rastvorenog kiseonika korišćeni su i mnogi drugi parametri kvaliteta voda:

- $\text{NO}_2\text{-N}$, $\text{NO}_3\text{-N}$, protok, biološka potrošnja kiseonika (BPK) [61];

- mesec u godini, hlorofil-a, $\text{NH}_4\text{-N}$, $\text{NO}_3\text{-N}$ [64];
- ukupni alkalitet, ukupna tvrdoća, ukupne čvrste materije, hemijska potrošnja kiseonika (HPK), $\text{NH}_4\text{-N}$, $\text{NO}_3\text{-N}$, sadržaj hlorida, fosfata, K, Na [65];
- ukupni alkalitet, ukupna tvrdoća, ukupne čvrste materije, hemijska potrošnja kiseonika (HPK), sadržaj $\text{NH}_4\text{-N}$, $\text{NO}_2\text{-N}$, hlorida, fosfata, K, Na [62];
- turbiditet, tvrdoća, alkalitet, sadržaj hlorida, $\text{NH}_4\text{-N}$, $\text{NO}_2\text{-N}$ [59];
- električna provodljivost, protok vode [63];
- električna provodljivost, alkalitet, tvrdoća, sadržaj hlorida, kalcijuma, $\text{NH}_4\text{-N}$, $\text{NO}_2\text{-N}$ [58].

Kako je cilj kreirati ANN model koristeći nespecifične parametre kvaliteta voda, za ulazne promenljive modela izabrani su pH, temperatura, električna provodljivost i protok rečne vode. Navedeni parametri kvaliteta voda su, pored toga što su nespecifični, izabrani i zato što se lako, brzo i kontinualno određuju na većini monitoring stanica.

Kao i kod kreiranja prethodnih ANN modela, dostupni podaci podeljeni su u tri grupe podataka: za obučavanje (podaci od 2004. do 2008. godine, ukupno 62 serije podataka), za validaciju (15 % od podataka za obučavanje) i za proveru (podaci iz 2009. godine, 11 serija - 15 % svih dostupnih podataka).

Statistička analiza formiranih grupa podataka prikazana je u Tabeli 37., dok je grupa podataka za proveru prikazana u Tabeli 38.

Tabela 37. Statistička analiza grupa podataka za obučavanje i proveru modela

God.	Broj serija podataka	Statistički pokazatelji	Promenljive				
			Protok (m ³ s ⁻¹)	pH	Temp. (°C)	Provod. (μS cm ⁻¹)	Rastvoreni kiseonik (mg O ₂ L ⁻¹)
Podaci za obučavanje							
2004	10	minimum	1250	7,7	0,3	301	7,0
2005	21						
2007	20	maksimum	3760	8,7	24,7	504	15,7
2008	11						
	Σ 62	srednja vr.	2176	8,2	12,7	394	10,9
Podaci za proveru							
2009	11	minimum	1600	8,0	3,2	334	7,1
		maksimum	5690	8,7	22,4	518	12,7
		sred. vred.	2671	8,3	13,2	412	10,3

Tabela 38. Grupa podataka, iz 2009. godine, korišćena za proveru modela

Mesec	Protok (m ³ s ⁻¹)	pH	Temp. (°C)	Provodnost (μS cm ⁻¹)	Rastvoreni kiseonik (mg O ₂ L ⁻¹)
Januar	1600	8,2	3,2	518	12,7
Februar	3930	8,1	6,2	448	11,1
Mart	3120	8,3	9,1	469	11,5
April	2890	8,7	15,8	368	11,8
Maj	2160	8,7	20,9	351	11,5
Jun	5690	8,0	20,7	334	7,1
Jul	2860	8,2	22,4	358	8,5
Avgust	2010	8,0	19,7	369	8,5
Septembar	1790	8,3	10,4	443	10,1
Oktoabar	1710	8,2	8,5	399	10,4
Novembar	1620	8,4	8,1	473	10,6
Decembar	-	-	-	-	-

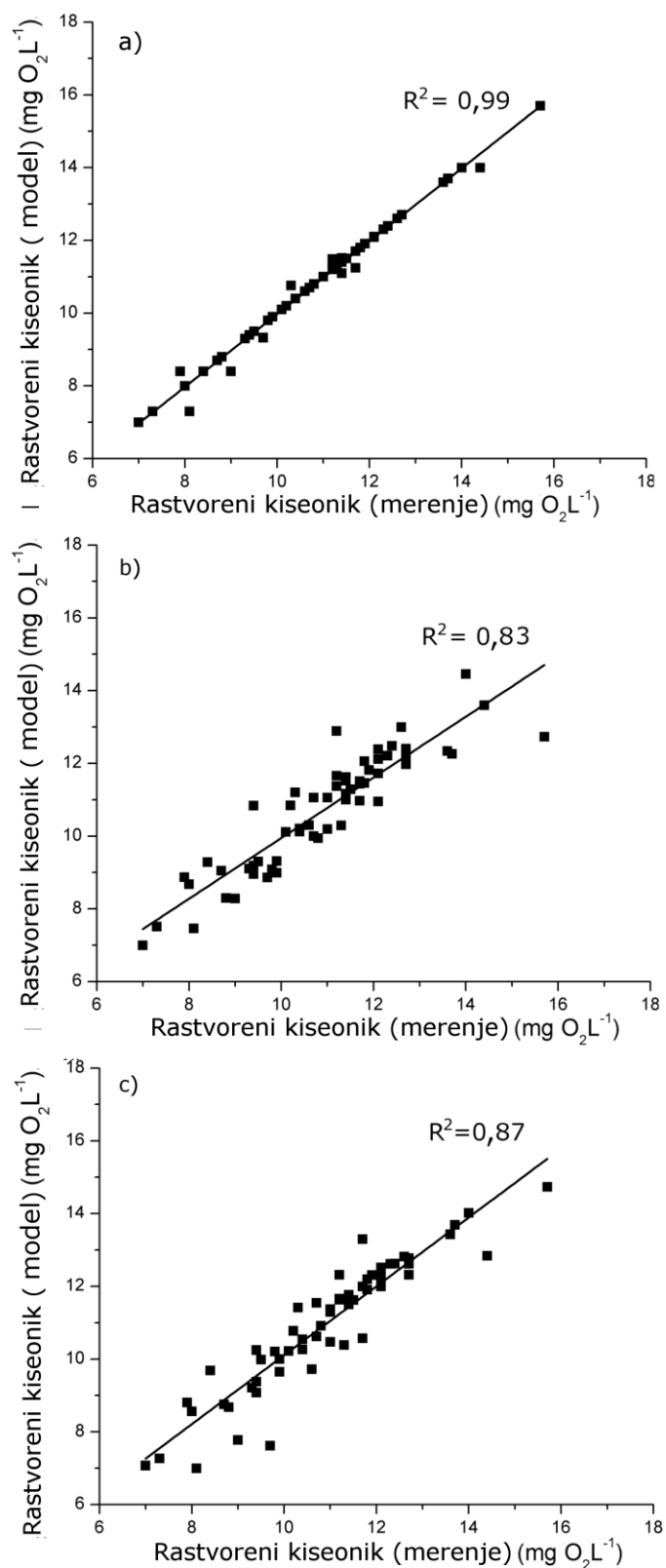
6.5.2. Razvoj i rezultati ANN modela za predviđanje sadržaja rastvorenog kiseonika

Tri različite arhitekture neuronskih mreža su odabrane za kreiranje modela za predviđanje sadržaja rastvorenog kiseonika u vodama:

- a) Neuronska mreža sa opštom regresijom (GRNN) - odabrana je jer se u istraživanjima rađenim u okviru ove doktorske disertacije pokazala kao najbolja,
- b) Neuronska mreža sa propagacijom greške unazad (BPNN) - odabrana je jer se najčešće primenjuje za modelovanje sadržaja rastvorenog kiseonika u vodama [58, 59, 61, 64],
- c) Rekurentna neuronska mreža (RNN) – izabrana je jer je pogodna za učenje sekvenci i podataka koji su vremenske serije [104].

Kao i u prethodnim slučajevima, modeli su kreirani korišćenjem NeuroShell 2 softverskog paketa [161]. Poređenje izmerenog sadržaja rastvorenog kiseonika i rezultata modela tokom obučavanja za sve tri odabrane ANN arhitekture prikazano je na Slici 37. Slaganje izmerenih i modelom predviđenih vrednosti može se smatrati kao dobro. GRNN model je dao odlična predviđanja sadržaja rastvorenog kiseonika ($R^2=0,99$), mnogo bolja od RNN ($R^2=0,87$) i BPNN ($R^2=0,83$) modela. U skladu sa dobijenim rezultatima može se konstatovati da je poredak modela prema performansama GRNN > RNN > BPNN.

Kada se uzmu u obzir i statistički pokazatelji performansi modela (Tabela 39.), može se pretpostaviti da je odličan rezultat GRNN modela zapravo posledica "pre-treniranosti", koja značajno smanjuje sposobnost uopštavanja. Takođe, može se pretpostaviti da će RNN i BPNN imati bolju sposobnost uopštavanja i samim tim i bolje rezultate prilikom provere modela, u odnosu na GRNN model.



Slika 37. Poređenje izmerenog sadržaja rastvorenog kiseonika i vrednosti dobijenih modelima tokom obučavanja: a) GRNN, b) BPNN, c) RNN

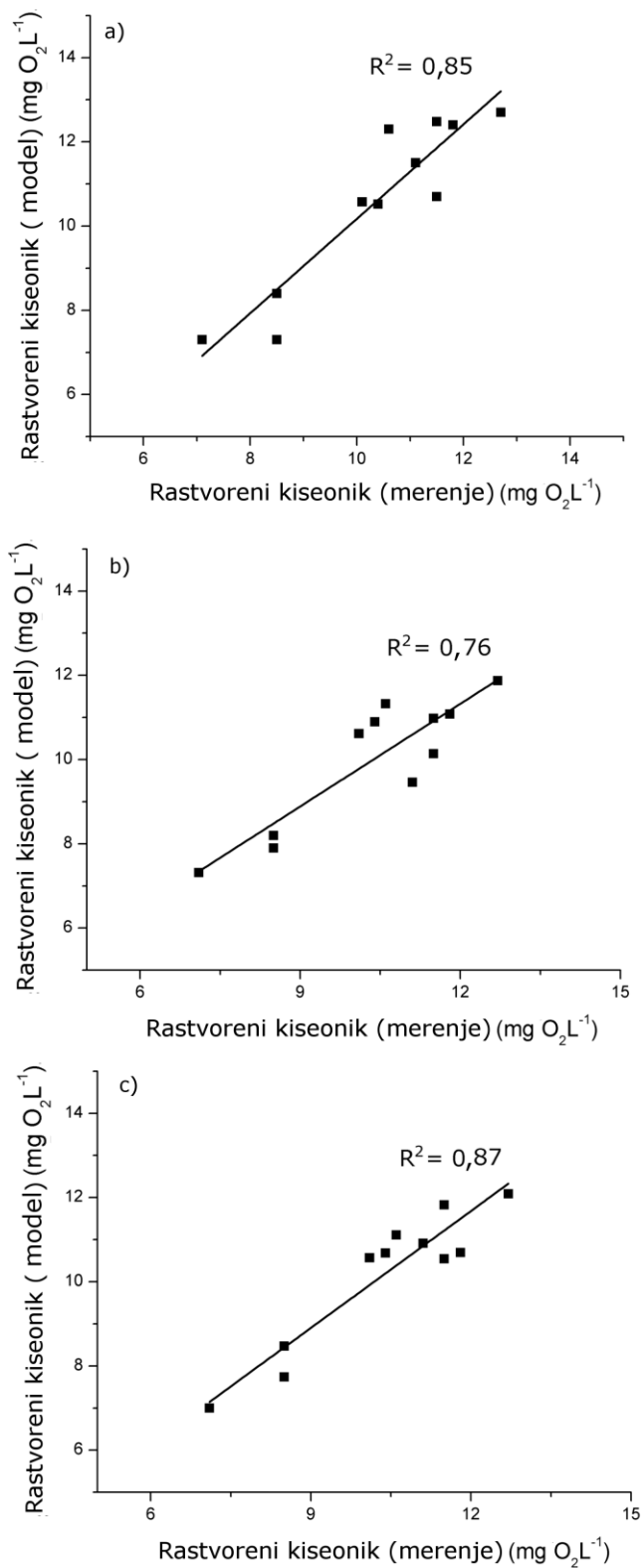
Tabela 39. Vrednosti statističkih pokazatelja performansi kreiranih ANN modela

Model	Statistički pokazatelji performansi			
	IA	FA1.1 (%)	MAE (mg O ₂ L ⁻¹)	RMSE (mg O ₂ L ⁻¹)
Obučavanje				
GRNN	1,00	100	0,07	0,19
BPNN	0,95	90	0,55	0,74
RNN	0,97	87	0,44	0,65
Provera				
GRNN	0,95	82	0,60	0,78
BPNN	0,93	82	0,72	0,83
RNN	0,97	100	0,49	0,59

Poređenje izmerenog sadržaja rastvorenog kiseonika i vrednosti dobijenih modelima tokom provere prikazano je na Slici 38. Kao što se može videti, performanse GRNN modela su značajno opale, procenat rezultata sa greškom manjom od 10 % (FA1.1) opao je sa 100 % na 82 %, dok je srednja apsolutna greška (MAE) porasla sa 0,07 mg O₂ L⁻¹ (obučavanje) na 0,60 mg O₂ L⁻¹ (provera). Na osnovu ovih rezultata, može se doneti i konačan zaključak da je GRNN model "pretreniran". S druge strane, za BPNN model se može konstatovati da nije bio dovoljno treniran. Da bi se poboljšale performanse BPNN modela bilo bi potrebno izvršiti redizajniranje i promeniti neke od osnovnih arhitektonskih parametara same mreže (broj skrivenih slojeva, broj neurona u skrivenim slojevima) i/ili povećati broj ulaznih promenljivih i količinu podataka.

Sto se tiče RNN modela, može se zaključiti da je optimalno treniran, s obzirom da se performanse tokom provere nisu promenile, u poređenju sa obučavanjem. Sva predviđanja za 2009. godinu su bila u okviru FA1.1, dok su vrednosti RMSE i MAE bile slične tokom obučavanja i provere modela (Tabela 39.). Uključujući sve navedeno, može se zaključiti da je realni poredak performansi kreiranih ANN modela RNN > GRNN > BPNN.

Analize u narednim poglavljima biće rađene samo za RNN model, koji je pokazao najbolje performanse kod predviđanja sadržaja rastvorenog kiseonika.



Slika 38. Poređenje izmerenog sadržaja rastvorenog kiseonika i vrednosti dobijenih modelima tokom provere: a) GRNN, b) BPNN, c) RNN

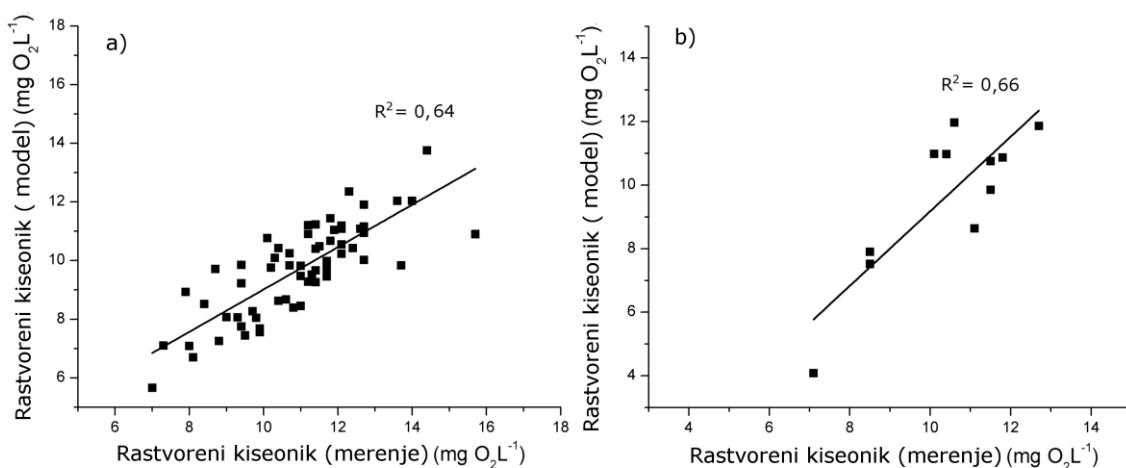
6.5.3. Poređenje RNN modela sa linearnim modelom

Vrednosti sadržaja rastvorenog kiseonika dobijene RNN modelom upoređene su sa predviđanjima modela sa višestrukom regresijom (MLR) koji je razvijen sa istim ulaznim promenljivim i podacima kao i RNN model. MLR model je razvijen korišćenjem SPSS 19 softverskog paketa [162], a koeficijenti modela su prikazani u Tabeli 40.

Tabela 40. Koeficijenti kreiranog MLR modela

Ulazi	Koeficijenti	Standardna greška
Konstanta	-22,87	6,561
Protok	-0,001	0,000
pH	4,560	0,629
Temperatura	-0,169	0,048
Provodnost	-0,001	0,007

Poređenje izmerenog sadržaja rastvorenog kiseonika i rezultata MLR modela tokom obučavanja i provere prikazano je na Slici 39. Vrednosti statističkih parametara za kreirani MLR model prikazani su u Tabeli 41.

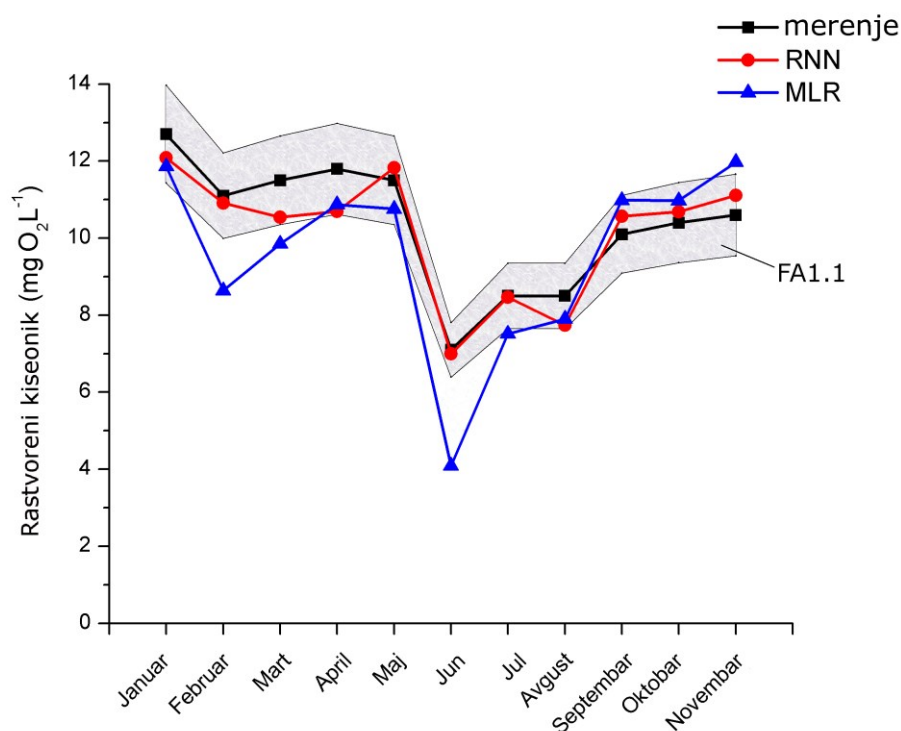


Slika 39. Poređenje izmerenog sadržaja rastvorenog kiseonika i rezultata MLR modela za: a) obučavanje b) proveru

Tabela 41. Vrednosti statističkih pokazatelja performansi MLR modela

Model	Statistički pokazatelji performansi modela			
	IA	FA1.1 (%)	MAE (mg O ₂ L ⁻¹)	RMSE (mg O ₂ L ⁻¹)
obučavanje modela				
MLR	0.75	38	1.35	1.63
provera modela				
MLR	0.85	55	1.28	1.49

Kao što se može uočiti, RNN model je pokazao mnogo bolje performanse od MLR modela kod obučavanja i provere modela. MLR rezultati su imali MAE i RMSE greške čak i do tri puta veće u poređenju sa RNN modelom. Takođe, tek 55 % predviđanja MLR modela za 2009. godinu je bilo u okviru FA1.1, što je gotovo dva puta lošiji rezultat od RNN modela, čije su sve predikcije za navedenu godinu bile u okviru FA1.1 (Slika 40.).



Slika 40. Poređenje izmerenog i modelima (RNN i MLR) predviđenog sadržaja rastvorenog kiseonika za 2009. godinu

7. Zaključak

Osnovni cilj istraživanja opisanih u okviru ove doktorske disertacije je razvoj modela za predviđanje emisija i koncentracija različitih zagađujućih materija čije vrednosti predstavljaju indikatore kvaliteta životne sredine. Primljeni modeli zasnovani su na veštačkim neuronskim mrežama (ANN) i korišćenju dostupnih ekonomskih i industrijskih indikatora i indikatora održivog razvoja, kao ulaznih promenljivih. Poznavanje vrednosti navedenih indikatora kvaliteta životne sredine od suštinske je važnosti u okviru upravljanja životnom sredinom, jer su oni osnova za donošenje mera i strategija vezanih za poboljšanje stanja u životnoj sredini, ali su u isto vreme i mera ostvarenja zadatih ciljeva i sredstvo za uključivanje javnosti u donošenje odluka vezanih za pitanja zaštite životne sredine. Zbog svog značaja, indikatori kvaliteta životne sredine su predmet istraživanja mnogih studija, a sve češće se za predviđanje njihovih vrednosti koristi modelovanje, uz primenu veštačkih neuronskih mreža. U okviru ove doktorske disertacije obuhvaćeni su sledeći indikatori kvaliteta životne sredine, koji se prate na nacionalnom nivou: emisija gasova staklene bašte, emisija PM₁₀ čestica, koncentracija PM₁₀ čestica i količina generisanog komunalnog čvrstog otpada, kao i sadržaj rastvorenog kiseonika u površinskim vodama koji pokazuje kvalitet dela životne sredine na lokalnom nivou.

Kao polazna hipoteza kod razvoja modela za navedene indikatore kvaliteta životne sredine, koji se prate na nacionalnom nivou, pretpostavljena je njihova nelinearna matematička povezanost, kao zavisnih promenljivih, sa ekonomskim indikatorima i indikatorima održivog razvoja, kao nezavisnim promenljivim. Takođe, kod modelovanja sadržaja rastvorenog kiseonika pretpostavljena je nelinearna matematička povezanost specifičnih i nespecifičnih parametara kvaliteta površinskih voda. Na osnovu navedenih pretpostavki, veštačke neuronske mreže su odabrane kao alat za modelovanje, jer su u stanju da vrše predviđanje izlaznih promenljivih, koje su nelinearno povezane sa ulaznim promenljivima modela, preciznije u odnosu na odgovarajuće linearne modele. Tokom razvoja ANN modela, pored njihove provere i ocene različitim statističkim pokazateljima performansi, dobijeni rezultati su u-

poređeni i sa rezultatima linearnih modela baziranih na višestrukoj linearnoj regresiji (MLR) i regresiji glavnih komponenti (PCR).

Za modelovanje emisije gasova staklene bašte korišćena je neuronska mreža sa opštom regresijom (GRNN), pri čemu su svi njeni parametri detaljno optimizovani. Odabir potencijalnih ulaznih promenljivih vršen je iz grupe ekonomskih indikatora i indikatora održivog razvoja. Tokom razvoja modela ulazne promenljive koje imaju najveći uticaj su određene korišćenjem korelacione analize i analize glavnih komponenti (PCA). Model je razvijen korišćenjem podataka za zemlje Evrope iz 2004-2010 i njegove performanse su analizirane primenom nekoliko različitih indikatora performansi modela. Optimizacija modela je pokazala da se najbolji rezultati dobijaju nasumičnim odabirom validacionih podataka, uz njihovo linearno skaliranje sa Euklidskim određivanjem rastojanja između podataka i genetičkim određivanjem faktora ujednačavanja. Poređenje GRNN modela kreiranih sa različitim ulaznim promenljivim (model sa svim dostupnim ulaznim promenljivim, model sa ulaznim promenljivim odabranim korelacionom analizom i model sa ulaznim promenljivim odabranim PCA analizom) je pokazalo da se najbolji rezultati dobijaju primenom modela sa korelaciono odabranim ulaznim promenljivim. Njegovo dalje poređenje sa odgovarajućim linearnim MLR i PCR modelima je pokazalo da GRNN ima znatno bolje performanse kod predviđanja emisije gasova staklene bašte. Kako su tokom korelacione analize eliminisana 3 od 14 primarno odabranih ulaznih indikatora, može se zaključiti da se ovakvim pristupom dobijaju precizniji modeli, uz korišćenje znatno manjeg broja ulaznih promenljivih. Daljom analizom razvijenog GRNN modela, korišćenjem analize značajnosti pojedinih ulaznih indikatora na osnovu vrednosti pojedinačnih faktora ujednačavanja, pokazano je da je najznačajniji ulazni indikator ukupna potrošnja energije. Simulacija primene razvijenog GRNN modela je pokazala da je model sposoban da predvidi vrednosti emisija gasova staklene bašte sa maksimalnom relativnom greškom od svega 8 % i srednjom relativnom greškom od svega 3,6 %. Na osnovu dobijenih rezultata se može zaključiti da se razvijeni GRNN model može uspešno koristiti za predviđanje emisije gasova staklene bašte na nacionalnom nivou (čije vrednosti su potrebne kod izveštavanja u okviru Kjoto protokola i održivog razvoja uopšte). Takođe, či-

njenica da ANN pristup zahteva manje ulaznih podataka koji pri tom nisu striktno definisani, već mogu biti odabrani u okviru dostupnih ekonomskih indikatora i indikatora održivog razvoja za svaku pojedinačnu zemlju, nameće ovaj metod procene emisija kao dobru alternativu u slučajevima kada primena modela zasnovanih na inventarskom pristupu nije moguća zbog nedostatka odgovarajućih podataka.

Kako se ANN metodologija bazirana na primeni GRNN arhitekture u kombinaciji sa ekonomskim indikatorima i indikatorima održivog razvoja pokazala kao jako uspešna kod modelovanja emisije gasova staklene bašte, ista je primenjena i za modelovanje emisije i koncentracije PM₁₀ čestica u atmosferi, na nacionalnom nivou. Kod modelovanja koncentracija PM₁₀ čestica za odabir najznačajnijih ulaznih promenljivih primenjen je metod zasnovan na korelacionoj analizi, koji je razvijen tokom modelovanja gasova staklene bašte, dok su za potrebe modelovanja emisija PM₁₀ čestica definisane dve nove procedure odabira ulaznih promenljivih, obe zasnovane na vrednostima pojedinačnih faktora ujednačavanja određenih genetičkim algoritmom. GRNN modeli sa optimizovanim brojem ulaznih promenljivih su imali bolje performanse od istih modela kreiranih sa svim primarno odabranim ulaznim promenljivim. Oba GRNN modela, za predviđanje emisija i koncentracija PM₁₀ čestica u atmosferi, imaju znatno bolje performanse u poređenju sa odgovarajućim linearnim modelima, MLR i PCR. I u ovom slučaju se pokazalo da primena ANN metodologije značajno smanjuje broj potrebnih ulaznih podataka, u poređenju sa modelima zasnovanim na inventarskom pristupu. Takođe, činjenica da kod ANN modela nije potrebno koristiti emisione faktore, čije određivanje je vezano za terenska istraživanja i koji imaju veoma nesigurne vrednosti, znatno doprinosi smanjenju vremena i troškova potrebnih za prikupljanje neophodnih ulaznih podataka.

Za modelovanje godišnjih količina generisanog komunalnog čvrstog otpada (KČO), pored GRNN arhitekture, primenjena je i neuronska mreža sa propagacijom greške unazad (BPNN). Oba ANN modela razvijena su korišćenjem minimalnog broja ulaznih promenljivih, koje su na kraju svedene na svega tri ekonomska indikatora: bruto domaći proizvod, potrošnja materijala i produktivnost resursa. Obučavanje i

provera ANN modela vršena je podacima za odabrane evropske zemlje. Poređenje GRNN i BPNN modela je pokazalo da GRNN ima znatno kraće vreme obučavanja, kao i da se preciznost predviđanja količina KČO ne menjaju tokom godina za koje je model bio testiran. Primena GRNN KČO modela na zemlje u razvoju, Bugarsku i Srbiju, je pokazala zadovoljavajuću tačnost modela.

U ovoj doktorskoj disertaciji prikazan je i razvoj i primena modela za predviđanje sadržaja rastvorenog kiseonika u reci Dunav na teritoriji Republike Srbije. Pored već testiranih GRNN i BPNN arhitektura, za modelovanje ovog parametra kvaliteta voda korišćena je i rekurentna neuronska mreža (RNN), koja je posebno efikasna u predviđanjima vrednosti promenljivih koje predstavljaju vremenske serije. Kako je cilj modelovanja rastvorenog kiseonika bio razvoj modela koji bi imao široku primenu, uz korišćenje minimalnog broja ulaznih podataka, čije određivanje je na terenu brzo i jeftino, odabrana su četiri nespecifična parametra kvaliteta voda: temperatura, pH vrednost, električna provodljivost i protok rečne vode. Za razvoj i proveru modela korišćeni su podaci dobijeni redovnim monitoringom kvaliteta Dunava na monitoring stanici Bezdán u periodu 2004-2009. godina. Poređenje rezultata navedenih ANN arhitektura pokazalo je da najtačnija predviđanja sadržaja rastvorenog kiseonika daje RNN model. RNN model je imao znatno bolje performanse i od odgovarajućeg linearnog MLR modela. Tokom provere podacima iz 2009. godine, RNN model je dao predviđanja mesečnih vrednosti sadržaja rastvorenog kiseonika sa greškom manjom od 10 %. Kako RNN model koristi samo nespecifične ulazne promenljive, on može biti primenjen za poboljšanje sistema monitoringa kvaliteta voda, bez ulaganja velikih finansijskih sredstava, kao i u vanrednim i akcidentnim situacijama.

Pored već navedene primene ANN metodologije za predviđanje vrednosti indikatora kvaliteta životne sredine, razvijeni modeli mogu biti korišćeni i za simulaciju scenarija emisija ili koncentracija zagađujućih materija. Menjanjem vrednosti ulaznih indikatora, čije promene su uslovljene izmenama zakona ili donošenjem posebnih regulatornih mera, mogu se predvideti buduće vrednosti emisija ili koncentracija zagađujućih materija. Na taj način se omogućava provera efikasnosti pred-

viđenih mera za poboljšanje kvaliteta životne sredine, i vrši izbegavanje donošenja potencijalno skupih mera koje ne bi imale proporcionalno zadovoljavajući efekat tokom primene u praksi.

U okviru budućih istraživanja planirano je ispitivanje mogućnosti primene razvijene ANN metodologije za modelovanja emisije ostalih indikatora kvaliteta životne sredine i održivog razvoja (azotnih oksida, sumpornih oksida, metana, amonijaka i sl.). Pored toga, biće primenjene nove hemometrijske tehnike (npr. Monte Karlo simulacija) za odabir najznačajnijih ulaznih promenljivih..

Literatura

- [1] Report of the World Commission on Environment and Development, United Nations, <http://www.un.org/documents/ga/res/42/ares42-187.htm>, **1987**.
- [2] *Nacionalna strategija održivog razvoja*, Vlada Republike Srbije, Beograd, Srbija, **2008**.
- [3] P. Gabrielsen, P. Bosch, *Environmental Indicators: Typology and Use in Reporting*, European Environment Agency, **2003**.
- [4] Database, Eurostat, http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database, **2013**.
- [5] R.L. Desjardins, M.V.K. Sivakumar, C. de Kimpe, *The contribution of agriculture to the state of climate: Workshop summary and recommendations*, Agricultural and Forest Meteorology **142** (2007) 314-324.
- [6] W. Winiwarter, K. Rypdal, *Assessing the uncertainty associated with national greenhouse gas emission inventories: a case study for Austria*, Atmospheric Environment **35** (2001) 5425–5440.
- [7] EEA, *Greenhouse gas emission trends and projections in Europe 2012*, Publications Office of the European Union, Copenhagen, **2012**.
- [8] United Nations Framework Convention on Climate Change, in: Proceedings of the Convention on Climate Change on the Work of the Second Part of its Fifth session, United Nations, New York, USA., **1992**.
- [9] W. Hediger, *Modeling GHG emissions and carbon sequestration in Swiss agriculture: An integrated economic approach*, International Congress Series **1293** (2006) 86-95.
- [10] V. Dornburg, J. Vandam, A. Faaij, *Estimating GHG emission mitigation supply curves of large-scale biomass use on a country level*, Biomass and Bioenergy **31** (2007) 46-65.
- [11] P. Mancarella, G. Chicco, *Assessment of the greenhouse gas emissions from cogeneration and trigeneration systems. Part II: Analysis techniques and application cases*, Energy **33** (2008) 418-430.
- [12] K. Akimoto, F. Sano, T. Homma, J. Oda, M. Nagashima, M. Kii, *Estimates of GHG emission reduction potential by country, sector, and cost*, Energy Policy **38** (2010) 3384-3393.

- [13] G. Chicco, P. Mancarella, *Assessment of the greenhouse gas emissions from cogeneration and trigeneration systems. Part I: Models and indicators*, Energy **33** (2008) 410-417.
- [14] R. Couth, C. Trois, S. Vaughan-Jones, *Modelling of greenhouse gas emissions from municipal solid waste disposal in Africa*, International Journal of Greenhouse Gas Control **5** (2011) 1443-1453.
- [15] I. Juhos, L. Makra, B. Tóth, *Forecasting of traffic origin NO and NO₂ concentrations by Support Vector Machines and neural networks using Principal Component Analysis*, Simulation Modelling Practice and Theory **16** (2008) 1488-1502.
- [16] A. Rentziou, K. Gkritza, R.R. Souleyrette, *VMT, energy consumption, and GHG emissions forecasting for passenger transportation*, Transportation Research Part A: Policy and Practice **46** (2012) 487-500.
- [17] S. Syri, A. Lehtila, T. Ekholm, I. Savolainen, H. Holttinen, E. Peltola, *Global energy and emissions scenarios for effective climate change mitigation - Deterministic and stochastic scenarios with the TIAM model*, International Journal of Greenhouse Gas Control **2** (2008) 274 – 285.
- [18] G. Villalba, E.D. Gemechu, *Estimating GHG emissions of marine ports - the case of Barcelona*, Energy Policy **39** (2011) 1363-1368.
- [19] GAINS review 2009, GAINS EUROPE,
<http://gains.iiasa.ac.at/index.php/documentation-of-model-methodology/model-reviews/gains-review-2009>, **2012**.
- [20] A. Sözen, Z. Gülseven, E. Arcaklioğlu, *Forecasting based on sectoral energy consumption of GHGs in Turkey and mitigation policies*, Energy Policy **35** (2007) 6491-6505.
- [21] A. Sözen, Z. Gülseven, E. Arcaklioğlu, *Estimation of GHG Emissions in Turkey Using Energy and Economic Indicators*, Energy Sources, Part A: Recovery, Utilization, and Environmental Effects **31** (2009) 1141-1159.
- [22] D. Radojević, V. Pocajt, I. Popović, A. Perić-Grujić, M. Ristić, *Forecasting of greenhouse gas emissions in Serbia using artificial neural networks*, Energy Sources, Part A: Recovery, Utilization, and Environmental Effects **35** (2013) 733-740.
- [23] M.L. Scapellato, C. Canova, A. de Simone, M. Carrieri, P. Maestrelli, L. Simonato, G.B. Bartolucci, *Personal PM₁₀ exposure in asthmatic adults in Padova, Italy: seasonal variability and factors affecting individual concentrations of particulate matter*, International Journal of Hygiene and Environmental Health **212** (2009) 626-636.

- [24] H.J. Jahn, A. Schneider, S. Breitner, R. Eissner, M. Wendisch, A. Kramer, *Particulate matter pollution in the megacities of the Pearl River Delta, China - a systematic literature review and health risk assessment*, International Journal of Hygiene and Environmental Health **214** (2011) 281-295.
- [25] A. Sfetsos, D. Vlachogiannis, *A new methodology development for the regulatory forecasting of PM10. Application in the Greater Athens Area, Greece*, Atmospheric Environment **44** (2010) 3159-3172.
- [26] Directive 2001/81/EC of the European Parliament and of the Council of 23 October 2001 on national emission ceilings for certain atmospheric pollutants, European Commission, Official Journal of the European Communities, **2001**, pp. 22-30.
- [27] Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe, Official Journal of the European Union, **2008**, pp. 1-44.
- [28] Emissions of primary particulates - outlook from LRTAP, European Environment Agency (EEA), http://www.eea.europa.eu/data-and-maps/indicators/emissions-of-primary-particulates-outlook#data_specifications, **2013**.
- [29] P. Kassomenos, S. Karakitsios, C. Papaloukas, *Estimation of daily traffic emissions in a South-European urban agglomeration during a workday. Evaluation of several "what if" scenarios*, The Science of The Total Environment **370** (2006) 480-490.
- [30] I.B. Konovalov, M. Beekmann, F. Meleux, A. Dutot, G. Foret, *Combining deterministic and statistical approaches for PM10 forecasting in Europe*, Atmospheric Environment **43** (2009) 6425-6434.
- [31] *Airborne particles and health: HEI epidemiologic evidence. HEI perspectives*, Health Effects Institute, Cambridge, MA, **2001**.
- [32] A.J. Cohen, H.R. Anderson, B. Ostro, K.D. Pandey, M. Krzyzanowski, N. Künzli, K. Gutschmidt, C.A. Pope III, I. Romieu, J.M. Samet, K.R. Smith, *Urban air pollution*, in: M. Ezzati, A.D. Lopez, A. Rodgers, C.J.L. Murray (Eds.) Comparative quantification of health risks: global and regional burden of disease attributable to selected major risk factors, World Health Organization (WHO), Geneva, Switzerland, **2004**.
- [33] D. Pandey, D. Wheeler, B. Ostro, U. Deichmann, K. Hamilton, K. Bolt, *Ambient Particulate Matter Concentrations in Residential and Pollution Hotspot areas of World Cities: New Estimates based on the Global Model of Ambient Particulates (GMAPS)*, in: The World Bank Development Economics Research Group and the Environment Department Working Paper, The World Bank, Washington DC, **2006**.

- [34] World dataBank, The World Bank Website, <http://databank.worldbank.org/ddp/home.do?Step=1&id=4>, **2012**.
- [35] P. Perez, J. Reyes, *Prediction of maximum of 24-h average of PM₁₀ concentrations 30 h in advance in Santiago, Chile*, Atmospheric Environment **36** (2002) 4555-4561.
- [36] J. Kukkonen, L. Partanen, A. Karppinen, J. Ruuskanen, H. Junninen, M. Kolehmaineb, H. Niska, S. Dorling, T. Chatterton, R. Foxall, G. Cawley, *Extensive evaluation of neural network models for the prediction of NO₂ and PM₁₀ concentrations, compared with a deterministic modelling system and measurements in central Helsinki*, Atmospheric Environment **37** (2003) 4539-4550.
- [37] J. Hooyberghs, C. Mensink, G. Dumont, F. Fierens, O. Brasseur, *A neural network forecast for daily average PM concentrations in Belgium*, Atmospheric Environment **39** (2005) 3279-3289.
- [38] D. Voukantsis, K. Karatzas, J. Kukkonen, T. Rasanen, A. Karppinen, M. Kolehmainen, *Intercomparison of air quality data using principal component analysis, and forecasting of PM₁₀ and PM_{2.5} concentrations using artificial neural networks, in Thessaloniki and Helsinki*, The Science of The Total Environment **409** (2011) 1266-1276.
- [39] A.K. Paschalidou, S. Karakitsios, S. Kleanthous, P.A. Kassomenos, *Forecasting hourly PM(10) concentration in Cyprus through artificial neural networks and multiple regression models: implications to local environmental management*, Environmental Science and Pollution Research International **18** (2011) 316-327.
- [40] P. Schubeler, *Conceptual Framework for Municipal Solid Waste Management in Low-Income Countries*, in: U.M.P.W.P.N. 9, World Bank, **1996**.
- [41] P. Beigl, S. Lebersorger, S. Salhofer, *Modelling municipal solid waste generation: A review*, Waste Management **28** (2008) 200-214.
- [42] S. Mor, K. Ravindra, A. Visscher, R.P. Dahiya, A. Chandra, *Municipal solid waste characterization and its assessment for potential methane generation: A case study*, The Science of The Total Environment **371** (2006) 1-10.
- [43] M. Berkun, E. Aras, S. Nemlioglu, *Disposal of solid waste in Istanbul and along the Black sea coast of Turkey*, Waste Management **25** (2005) 847-855.
- [44] S.S. Chung, C.W.H. Lo, *Local waste management constraints and waste administrators in China*, Waste Management **28** (2008) 272-281.
- [45] N. Gamze Turan, S. Çoruh, A. Akdemir, O. Nuri Ergun, *Municipal solid waste management strategies in Turkey*, Waste Management **29** (2009) 465-469

- [46] R.K. Henry, Z. Yongsheng, D. Jun, *Municipal solid waste management challenges in developing countries – Kenyan case study*, *Waste Management* **29** (2009) 92-100.
- [47] U.N. Ngoc, H. Schnitzer, *Sustainable solutions for solid waste management in Southeast Asian countries*, *Waste Management* **29** (2009) 1982-1995.
- [48] A.V. Shekdar, *Sustainable solid waste management: An integrated approach for Asian countries*, *Waste Management* **29** (2009) 1438-1448.
- [49] K. Hristovski, L. Olson, N. Hild, D. Peterson, S. Burge, *The municipal solid waste system and solid waste characterization at the municipality of Veles, Macedonia*, *Waste Management* **27** (2007) 1680-1689.
- [50] G. Vego, S. Kučar-Dragičević, N. Koprivanac, *Application of multi-criteria decision making on strategic municipal solid waste management in Dalmatia, Croatia*, *Waste Management* **28** (2008) 2192-2201.
- [51] A. Dedijer et al., *ENVIRONMENT in Serbia: an indicator – based review*, Serbian Environmental Protection Agency (SEPA), Belgrade, Serbia, **2007**.
- [52] SEPA, *Report on the environmental situation in the Republic of Serbia in year 2009*, Ministry of Environment and Spatial Planning, Belgrade, Serbia, **2010**.
- [53] R. Yin, Q. Xiang, *An integrative approach to modeling land-use changes: multiple facets of agriculture in the Upper Yangtze basin*, *Sustainability Science* **5** (2010) 9-18.
- [54] J. Phillips, *The advancement of a mathematical model of sustainable development*, *Sustainability Science* **5** (2010) 127-142.
- [55] R. Noori, M.A. Abdoli, A. Farokhnia, M. Abbasi, *Results uncertainty of solid waste generation forecasting by hybrid of wavelet transform-ANFIS and wavelet transform-neural network*, *Expert Systems with Applications* **36** (2009) 9991-9999.
- [56] S. Jahandideh, E.B. Asadabadi, M. Askarian, M.M. Movahedi, S. Hosseini, M. Jahandideh, *The use of artificial neural networks and multiple linear regression to predict rate of medical waste generation*, *Waste Management* **29** (2009) 2874-2879.
- [57] S. Bayar, I. Demir, G.O. Engin, *Modeling leaching behavior of solidified wastes using back-propagation neural networks*, *Ecotoxicology and environmental safety* **72** (2009) 843-850.
- [58] X. Wen, J. Fang, M. Diao, C. Zhang, *Artificial neural network modeling of dissolved oxygen in the Heihe River, Northwestern China*, *Environmental monitoring and assessment* **185** (2013) 4361-4371.

- [59] V. Ranković, J. Radulović, I. Radojević, A. Ostojić, L. Čomić, *Neural network modeling of dissolved oxygen in the Gruža reservoir, Serbia*, *Ecological Modelling* **221** (2010) 1239-1244.
- [60] J. Vieira, A. Fonseca, V.J. Vilar, R.A. Boaventura, C.M. Botelho, *Water quality modelling of Lis River, Portugal*, *Environmental science and pollution research international* **20** (2013) 508-524.
- [61] B. Şengörür, E. Dogan, R. Koklu, A. Samandar, *Dissolved oxygen estimation using artificial neural network for water quality control*, *Fresenius Environmental Bulletin* **15** (2006) 1064-1067.
- [62] N. Basant, S. Gupta, A. Malik, K.P. Singh, *Linear and nonlinear modeling for simultaneous prediction of dissolved oxygen and biochemical oxygen demand of the surface water - A case study*, *Chemometrics and Intelligent Laboratory Systems* **104** (2010) 172-180.
- [63] M. Ay, O. Kisi, *Modeling of Dissolved Oxygen Concentration Using Different Neural Network Techniques in Foundation Creek, El Paso County, Colorado*, *Journal of Environmental Engineering* **138** (2012) 654-662.
- [64] J.-T. Kuo, M.-H. Hsieh, W.-S. Lung, N. She, *Using artificial neural network for reservoir eutrophication prediction*, *Ecological Modelling* **200** (2007) 171-177.
- [65] K.P. Singh, A. Basant, A. Malik, G. Jain, *Artificial neural network modeling of the river water quality—A case study*, *Ecological Modelling* **220** (2009) 888-895.
- [66] EMEP/EEA emission inventory guidebook - 2009, European Environment Agency (EEA), <http://www.eea.europa.eu/publications/emep-eea-emission-inventory-guidebook-2009>, **2009**.
- [67] AP 42, Fifth Edition, Compilation of air pollutant emission factors, Volume I: Stationary point and area sources, U.S. Environmental Protection Agency (US EPA) <http://www.epa.gov/ttnchie1/ap42/>, **1995**.
- [68] Air pollution statistics, Eurostat, http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Air_pollution_statistics, **2012**.
- [69] K. Rypdal, W. Winiwarter, *Uncertainties in greenhouse gas emission inventories - evaluation, comparability and implications*, *Environmental Science & Policy* **4** (2001) 107-116.
- [70] S. Monni, S. Syri, I. Savolainen, *Uncertainties in the Finnish greenhouse gas emission inventory*, *Environmental Science & Policy* **7** (2004) 87-98.

- [71] G. Wang, S. Chen, *A review on parameterization and uncertainty in modeling greenhouse gas emissions from soil*, *Geoderma* **170** (2012) 206-216.
- [72] About IIASA, International Institute for Applied Systems Analysis - IIASA, <http://www.iiasa.ac.at/web/home/research/researchPrograms/Program-Overview.en.html>, **2013**.
- [73] W. Schöpp, M. Amann, J. Cofala, C. Heyes, Z. Klimont, *Integrated assessment of European air pollution emission control strategies*, *Environmental Modelling & Software* **14** (1999) 1-9.
- [74] M. Amann, J. Cofala, C. Heyes, Z. Klimont, R. Mechler, M. Posch, W. Schöpp, *The RAINS model. Documentation of the model approach prepared for the RAINS peer review 2004*, International Institute for Applied Systems Analysis, **2004**.
- [75] RAINS review 2004, GAINS EUROPE, <http://gains.iiasa.ac.at/index.php/documentation-of-model-methodology/model-reviews/rains-review-2004>, **2012**.
- [76] S. Soleille, *The RAINS model, uncertainty and optimisation*, INERIS, Paris, France, **2004**.
- [77] W. Schöpp, Z. Klimont, R. Suutari, J. Cofala, *Uncertainty analysis of emission estimates in the RAINS integrated assessment model*, *Environmental Science & Policy* **8** (2005) 601-613.
- [78] M. Amann, I. Bertok, J. Borken-Kleefeld, J. Cofala, C. Heyes, L. Höglund-Isaksson, Z. Klimont, B. Nguyen, M. Posch, P. Rafaj, R. Sandler, W. Schöpp, F. Wagner, W. Winiwarter, *Cost-effective control of air quality and greenhouse gases in Europe: Modeling and policy applications*, *Environmental Modelling & Software* **26** (2011) 1489-1501.
- [79] Air Pollution in World Cities (PM10 Concentrations), The World Bank, <http://go.worldbank.org/3RDF07T6M0> **2012**.
- [80] S. Sousa, F. Martins, M. Alvimferraz, M. Pereira, *Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations*, *Environmental Modelling & Software* **22** (2007) 97-103.
- [81] S.A. Abdul-Wahab, C.S. Bakheit, S.M. Al-Alawi, *Principal component and multiple regression analysis in modelling of ground-level ozone and factors affecting its concentrations*, *Environmental Modelling & Software* **20** (2005) 1263-1271.
- [82] C.N. Kroll, P. Song, *Impact of multicollinearity on small sample hydrologic regression models*, *Water Resources Research* **49** (2013) 1-14.

- [83] D.J. Dupuis, M.-P. Victoria-Feser, *Robust VIF regression with application to variable selection in large data sets*, *Annals of Applied Statistics* **7** (2013) 319-341.
- [84] S.M. Al-Alawi, S.A. Abdul-Wahab, C.S. Bakheit, *Combining principal component regression and artificial neural networks for more accurate predictions of ground-level ozone*, *Environmental Modelling & Software* **23** (2008) 396-403.
- [85] R. Noori, A. Karbassi, M. Salman Sabahi, *Evaluation of PCA and Gamma test techniques on ANN operation for weekly solid waste prediction*, *Journal of environmental management* **91** (2010) 767-771.
- [86] V. Pocajt, *Razvoj modela atmosfenske disperzije primenom neuronskih mreža*, doktorska disertacija, Tehnološko-metalurški fakultet Univerziteta u Beogradu, Beograd, Srbija, **1999**.
- [87] D.F. Specht, *A General Regression Neural Network*, *IEEE Transactions on Neural Networks* **2** (1991) 568-576.
- [88] D.F. Specht, *Probabilistic Neural Networks*, *Neural Networks* **3** (1990) 109-118.
- [89] T. Kohonen, *Self-organized formation of topologically correct feature maps*, *Biological Cybernetics* **43** (1982) 59-69.
- [90] T. Kohonen, *Self-Organizing Maps*, Third extended edition, Springer, **2001**.
- [91] P.D. Wasserman, *Neural computing theory and practice*, Van Nostrand Reinhold, New York, USA, **1989**.
- [92] S. Gallant, *Neural network learning and expert systems*, MIT Press, Cambridge, Massachusetts, USA, **1993**.
- [93] V. Cherkassky, H. Lari-Najafi, *Data Representation for Diagnostic Neural Networks*, *IEEE Intelligent Systems* **7** (1992) 43-44, 47, 49, 51-53.
- [94] M. Tripathy, *Power transformer differential protection using neural network Principal Component Analysis and Radial Basis Function Neural Network*, *Simulation Modelling Practice and Theory* **18** (2010) 600-611.
- [95] S. Walczak, N. Cerpa, *Heuristic principles for the design of artificial neural networks*, *Information and Software Technology* **41** (1999) 107-117.
- [96] C.E. Balas, M.L. Koç, R. Tür, *Artificial neural networks based on principal component analysis, fuzzy systems and fuzzy neural networks for preliminary design of rubble mound breakwaters*, *Applied Ocean Research* **32** (2010) 425-433.

- [97] F. He, C. Ma, *Modeling greenhouse air humidity by means of artificial neural network and principal component analysis*, Computers and Electronics in Agriculture **71** (2010) S19-S23.
- [98] D.E. Rumelhart, G.E. Hinton, R.J. Williams, *Learning representations by back-propagation errors*, Nature **323** (1986) 533-536.
- [99] T. Bolanča, Š. Cerjan-Stefanović, Š. Ukić, M. Rogošić, M. Luša, *Application of different training methodologies for the development of a back propagation artificial neural network retention model in ion chromatography*, Journal of Chemometrics **22** (2008) 106-113.
- [100] H.B. Demuth, M. Beale, *Neural Network Toolbox for Use with MATLAB, User's Guide.*, The MathWorks, Inc., Natick, MA, USA, **2004**.
- [101] T.P. Vogl, J.K. Mangis, A.K. Zigler, W.T. Zink, D.L. Alkon, *Accelerating the convergence of the backpropagation method.*, Biological Cybernetics **59** (1988) 256-264.
- [102] E.K.P. Chong, S.H. Zak, *An Introduction to Optimization*, Wiley, Singapore, **2004**.
- [103] M.T. Hagan, M.B. Menhaj, *Training feedforward networks with the Marquardt algorithm*, IEEE Transactions on Neural Networks **5** (1994) 989-993.
- [104] Neuroshell 2 Help, Ward systems group
<http://www.wardsystems.com/manuals/neuroshell2/index.html?idxhowuse.htm>, **2008**.
- [105] M.I. Jordan, *Attractor dynamics and parallelism in a connectionist sequential machine*, in: Proceedings of the 8th Annual Conference of the Cognitive Science Society, Hillsdale, NY, **1986**, pp. 531-546.
- [106] H. Schiøler, U. Hartmann, *Mapping neural network derived from the Parzenwindow estimator*, Neural Networks **5** (1992) 903-909.
- [107] D.F. Specht, *The General Regression Neural Network-Rediscovered*, Neural Networks **6** (1993) 1033-1034.
- [108] H. Schiøler, *Response to Letter by D. F. Specht*, Neural Networks **6** (1993) 1034.
- [109] E.A. Nadaraya, *On estimating regression*, Theory of Probability and Its Applications **10** (1964) 186-190.
- [110] G.S. Watson, *Smooth regression analysis*, Sankhya Series: A **26** (1964) 359-372.

- [111] D. Tomandl, A. Schober, *A Modified General Regression Neural Network (MGRNN) with new, efficient training algorithms as a robust 'black box'-tool for data analysis*, Neural Networks **14** (2001) 1023-1034.
- [112] E. Parzen, *On estimation of a probability density function and mode*, The Annals of Mathematical Statistics **33** (1962) 1065-1076.
- [113] M.M.T. Thwin, T.-S. Quah, *Application of neural networks for software quality prediction using object-oriented metrics*, Journal of Systems and Software **76** (2005) 147-156.
- [114] N.A. Zaghoul, M.A.A. Kiefa, *Neural network solution of inverse parameters used in the sensitivity-calibration analyses of the SWMM model simulations*, Advances in Engineering Software **32** (2001) 587-595.
- [115] C. Karul, S. Soyupak, A.F. Çilesiz, N. Akbay, E. Germen, *Case studies on the use of neural networks in eutrophication modelling*, Ecological Modelling **134** (2000) 145-152.
- [116] M.M. Hamed, M.G. Khalafallah, E.A. Hassanien, *Prediction of wastewater treatment plant performance using artificial neural networks*, Environmental Modelling & Software **19** (2004) 919-928.
- [117] B. He, T. Oki, F. Sun, D. Komori, S. Kanae, Y. Wang, H. Kim, D. Yamazaki, *Estimating monthly total nitrogen concentration in streams by using artificial neural network* Journal of environmental management **92** (2011) 172-177.
- [118] S. Dragović, A. Onjia, R. Dragović, G. Bačić, *Implementation of Neural Networks for Classification of Moss and Lichen Samples on the Basis of Gamma-Ray Spectrometric Analysis*, Environmental monitoring and assessment **130** (2007) 245-253.
- [119] E. Akkaya, A. Demir, *Predicting the Heating Value of Municipal Solid Waste-based Materials: An Artificial Neural Network Model*, Energy Sources, Part A: Recovery, Utilization, and Environmental Effects **32** (2010) 1777-1783.
- [120] H. Abu Qdais, K. Bani Hani, N. Shatnawi, *Modeling and optimization of biogas production from a waste digester using artificial neural network and genetic algorithm*, Resources, Conservation and Recycling **54** (2010) 359-363.
- [121] M. Cai, Y. Yin, M. Xie, *Prediction of hourly air pollutant concentrations near urban arterials using artificial neural network approach*, Transportation Research Part D: Transport and Environment **14** (2009) 32-41.
- [122] G. Corani, *Air quality prediction in Milan: feed-forward neural networks, pruned neural networks and lazy learning*, Ecological Modelling **185** (2005) 513-529.

- [123] L. Hrust, Z.B. Klaić, J. Križan, O. Antonić, P. Hercog, *Neural network forecasting of air pollutants hourly concentrations using optimised temporal averages of meteorological variables and pollutant concentrations*, Atmospheric Environment **43** (2009) 5588-5596.
- [124] W.Z. Lu, H.Y. Fan, S.M. Lo, *Application of evolutionary neural network method in predicting pollutant levels in downtown area of Hong Kong*, Neurocomputing **51** (2003) 387-400.
- [125] E.G. Ortiz-García, S. Salcedo-Sanz, Á.M. Pérez-Bellido, J.A. Portilla-Figueras, L. Prieto, *Prediction of hourly O₃ concentrations using support vector regression algorithms*, Atmospheric Environment **44** (2010) 4481-4488.
- [126] C.H. Tsai, L.C. Chang, H.C. Chiang, *Forecasting of ozone episode days by cost-sensitive neural network methods*, The Science of the total environment **407** (2009) 2124-2135.
- [127] F. Li, *Air Quality Prediction in Yinchuan by Using Neural Network*, in: Y. Tan, Y. Shi, K.C. Tan (Eds.) ICSI 2010, Part II, LNCS 6146, Springer-Verlag Berlin, Heidelberg, **2010**, pp. 548-557.
- [128] P. Viotti, G. Liuti, P.D. Genova, *Atmospheric urban pollution: applications of an artificial neural network (ANN) to the city of Perugia*, Ecological Modelling **148** (2002) 27-46.
- [129] S.M. Nagendra, M. Khare, *Artificial neural network based carbon monoxide persistence models for episodic urban air quality management*, Environmental monitoring and assessment **139** (2008) 247-255.
- [130] S.A. Kalogirou, *Artificial intelligence for the modeling and control of combustion processes: a review*, Progress in Energy and Combustion Science **29** (2003) 515-566.
- [131] A. Sözen, M.A. Akçayol, E. Arcaklioğlu, *Forecasting Net Energy Consumption Using Artificial Neural Network*, Energy Sources, Part B: Economics, Planning, and Policy **1** (2006) 147-155.
- [132] A. Sözen, E. Arcaklioglu, *Prediction of net energy consumption based on economic indicators (GNP and GDP) in Turkey*, Energy Policy **35** (2007) 4981-4992.
- [133] A. Mustapha, A.Z. Aris, H. Juahir, M.F. Ramli, N.U. Kura, *River water quality assessment using environmentric techniques: case study of Jakara River Basin*, Environmental Science and Pollution Research International **20** (2013) 5630-5644.
- [134] R.G. Brereton, *Applied chemometrics for scientists*, John Wiley & Sons Ltd, Chippingham, Wiltshire, UK, **2007**.

- [135] A. Perić-Grujić, *Osnovi hemometrije*, Tehnološko-metalurški fakultet Univerziteta u Beogradu, Beograd, **2012**.
- [136] B.R. Cattell, *The Scree Test for the Number of Factors*, *Multivariate Behavioral Research* **1** (1966) 245-276.
- [137] Y. Keho, *The Basics of Linear Principal Components Analysis*, in: P. Sanguansat (Ed.) *Principal Component Analysis*, InTech, Rijeka, Croatia, **2012**, pp. 181-206.
- [138] P.R. Peres-Neto, D.A. Jackson, K.M. Somers, *How many principal components? stopping rules for determining the number of non-trivial axes revisited*, *Computational Statistics & Data Analysis* **49** (2005) 974-997.
- [139] B.C. Geiger, G. Kubin, *Relative Information Loss in the PCA*, in: Proc. IEEE Information Theory Workshop, IEEE, Lausanne, Switzerland, **2012**.
- [140] J.H. Holland, *Adaptation in Natural and Artificial Systems*, MIT Press, Cambridge MA, **1975**.
- [141] L.C. Jain, N.M. Martin, *Fusion of Neural Networks, Fuzzy Systems and Genetic Algorithms: Industrial Applications*, CRC Press LLC, Washington, USA **1998**.
- [142] A. Zalzalá, P. Fleming, *Genetic algorithms in engineering systems*, The Institution of Electrical Engineers, London, UK, **1997**.
- [143] J. Heaton, *Introduction to Neural Networks with Java*, Heaton Research Inc, St. Louis, USA, **2005**.
- [144] C.-S. Ong, J.-J. Huang, G.-H. Tzeng, *Model identification of ARIMA family using genetic algorithms*, *Applied Mathematics and Computation* **164** (2005) 885-912.
- [145] G. Grivas, A. Chaloulakou, *Artificial neural network models for prediction of PM10 hourly concentrations, in the Greater Area of Athens, Greece*, *Atmospheric Environment* **40** (2006) 1216-1229.
- [146] *Hydrological yearbook, vol. 3. Water quality 2004*, Republic Hydrometeorological Service of Serbia (RHMZ) Belgrade, Serbia, **2005**.
- [147] *Hydrological yearbook, vol. 3. Water quality 2005* Republic Hydrometeorological Service of Serbia (RHMZ) **2006**.
- [148] *Hydrological yearbook, vol. 3. Water quality 2007*, Republic Hydrometeorological Service of Serbia (RHMZ) **2008**.
- [149] *Hydrological yearbook, vol. 3. Water quality 2008*, Republic Hydrometeorological Service of Serbia (RHMZ) Belgrade, Serbia, **2009**.

- [150] *Hydrological yearbook, vol. 3. Water quality 2009*, Republic Hydrometeorological Service of Serbia (RHMZ) Belgrade, Serbia, **2010**.
- [151] H.J. Fernando, M.C. Mammarella, G. Grandoni, P. Fedele, R. Di Marco, R. Dimitrova, P. Hyde, *Forecasting PM10 in metropolitan areas: Efficacy of neural networks*, *Environmental Pollution* **163** (2012) 62-67.
- [152] G. Najafi, B. Ghobadian, T. Tavakoli, D.R. Buttsworth, T.F. Yusaf, M. Faizollahnejad, *Performance and exhaust emissions of a gasoline engine with ethanol blended gasoline fuels using artificial neural network*, *Applied Energy* **86** (2009) 630-639.
- [153] Y.-S. Koo, S.-T. Kim, J.-S. Cho, Y.-K. Jang, *Performance evaluation of the updated air quality forecasting system for Seoul predicting PM10*, *Atmospheric Environment* **58** (2012) 56-69.
- [154] M.T. Pay, P. Jiménez-Guerrero, J.M. Baldasano, *Assessing sensitivity regimes of secondary inorganic aerosol formation in Europe with the CALIOPE-EU modeling system*, *Atmospheric Environment* **51** (2012) 146-164.
- [155] V.C. Patel, A. Kumar, *Evaluation of three air dispersion models: ISCST2, ISCLT2, and SCREEN2 for mercury emissions in an urban area*, *Environmental Monitoring and Assessment* **53** (1998) 259-277.
- [156] G. Zhang, B.E. Patuwo, M.Y. Hu, *Forecasting with artificial neural networks: The state of the art*, *International Journal of Forecasting* **14** (1998) 35-62.
- [157] D.Z. Antanasijević, M.Đ. Ristić, A.A. Perić-Grujić, V.V. Pocajt, *Forecasting GHG emissions using an optimized artificial neural network model based on correlation and principal component analysis*, *International Journal of Greenhouse Gas Control* **20** (2014) 244-253.
- [158] Greenhouse gas emission trends and projections in Europe 2012, European Environment Agency (EEA), <http://www.eea.europa.eu/publications/ghg-trends-and-projections-2012>, **2013**.
- [159] United Nations Economic Commission for Europe (UNECE), <http://www.unece.org>, **2013**.
- [160] S. Palani, S.Y. Liong, P. Tkalich, *An ANN application for water quality forecasting*, *Marine Pollution Bulletin* **56** (2008) 1586-1597.
- [161] Neuroshell 2, Ward systems group, Inc., **1993**.
- [162] IBM SPSS 19 Statistics for Windows, IBM Corp., **2010**.
- [163] H.F. Kaiser, *Index of factorial simplicity*, *Psychometrika* **39** (1974) 31-36.

- [164] D.Z. Antanasijević, V.V. Pocajt, D.S. Povrenović, M.D. Ristić, A.A. Perić-Grujić, *PM(10) emission forecasting using artificial neural networks and genetic algorithm input variable optimization*, *The Science of The Total Environment* **443** (2013) 511-519.
- [165] U.S. Geological Survey Minerals Yearbook-1999-2006, Volume III, U.S. Geological Survey (USGS), <http://minerals.usgs.gov/minerals/pubs/country/europe.html>, **2011**.
- [166] Steel Statistical Yearbook 1999-2006, International Iron and Steel Institute (IISI), <http://www.worldsteel.org/statistics/statistics-archive.html>, **2011**.
- [167] IFADATA, International Fertilizer Industry Association (IFA), <http://www.fertilizer.org/ifa/ifadata/search>, **2011**.
- [168] Council Directive 96/61/EC of 24 September 1996 concerning integrated pollution prevention and control, *Official Journal of the European Communities*, **1996**, pp. 26-40.
- [169] Directive 98/69/EC of the European Parliament and of the Council of 13 October 1998 relating to measures to be taken against air pollution by emissions from motor vehicles and amending Council Directive 70/220/EEC, *Official Journal of the European Communities*, **1998**, pp. 1-65.
- [170] Council Directive 1999/30/EC of 22 April 1999 relating to limit values for sulphur dioxide, nitrogen dioxide and oxides of nitrogen, particulate matter and lead in ambient air, *Official Journal of the European Communities*, **1999**, pp. 41-60.
- [171] D.Z. Antanasijević, M.Đ. Ristić, A.A. Perić-Grujić, V.V. Pocajt, *Forecasting human exposure to PM10 at the national level using an artificial neural network approach*, *Journal of Chemometrics* **27** (2013) 170-177.
- [172] R. Dennis, T. Fox, M. Fuentes, A. Gilliland, S. Hanna, C. Hogrefe, J. Irwin, S.T. Rao, R. Scheffe, K. Schere, D. Steyn, A. Venkatram, *A Framework for Evaluating Regional-Scale Numerical Photochemical Modeling Systems*, *Environmental Fluid Mechanics* **10** (2010) 471-489.
- [173] XLSTAT, Addinsoft, <http://www.xlstat.com/en>, **2012**.
- [174] D. Antanasijević, V. Pocajt, I. Popović, N. Redžić, M. Ristić, *The forecasting of municipal waste generation using artificial neural networks and sustainability indicators*, *Sustainability Science* **8** (2013) 37-46.
- [175] Domestic material consumption, Eurostat, <http://epp.eurostat.ec.europa.eu/tgm/table.do?tab=table&init=1&plugin=1&language=en&pcode=tsdpc230>, **2010**.

- [176] Gross domestic product at market prices, Eurostat, <http://epp.eurostat.ec.europa.eu/tgm/table.do?tab=table&init=1&language=en&pcode=tec00001&plugin=0>, **2010**.
- [177] Resource productivity, Eurostat, <http://epp.eurostat.ec.europa.eu/tgm/table.do?tab=table&init=1&plugin=1&language=en&pcode=tsdpc100>, **2010**.
- [178] Municipal waste generated, Eurostat, <http://epp.eurostat.ec.europa.eu/tgm/table.do?tab=table&init=1&plugin=1&language=en&pcode=tsdpc210>, **2010**.
- [179] *Report on the environmental situation in the Republic of Serbia in year 2009*, Serbian Environmental Protection Agency (SEPA) Belgrade, **2010**.
- [180] Statistical yearbook of Serbia 2008, Statistical Office of the Republic of Serbia <http://webrzs.stat.gov.rs/axd/en/index.php>, **2008**.
- [181] Statistical yearbook of Serbia 2009, Statistical Office of the Republic of Serbia <http://webrzs.stat.gov.rs/axd/en/index.php>, **2009**.
- [182] Statistical yearbook of Serbia 2010, Statistical Office of the Republic of Serbia <http://webrzs.stat.gov.rs/axd/en/index.php>, **2010**.
- [183] H. Weisz, F. Krausmann, C. Amann, N. Eisenmenger, K.-H. Erb, K. Hubacek, M. Fischer-Kowalski, *The physical economy of the European Union: Cross-country comparison and determinants of material consumption*, *Ecological Economics* **58** (2006) 676-698.
- [184] *The European environment – state and outlook 2010*, European Environment Agency (EEA), www.eea.europa.eu **2011**.
- [185] *Zakon o upravljanju otpadom*, Službeni glasnik RS, **2010**.
- [186] D. Antanasijević, V. Pocajt, D. Povrenović, A. Peric-Grujić, M. Ristić, *Modelling of dissolved oxygen content using artificial neural networks: Danube River, North Serbia, case study*, *Environmental Science and Pollution Research* **20** (2013) 9006–9013.

BIOGRAFIJA AUTORA

Davor Z. Antanasijević, istraživač saradnik, rođen je 26.03.1986. godine u Pirotu, gde je 2005. godine završio gimnaziju na prirodno-matematičkom smeru.

Iste godine upisao je Tehnološko-metalurški fakultet u Beogradu, smer Inženjerstvo zaštite životne sredine, koji je završio 2009. godine sa prosečnom ocenom 8,89. Završni rad „*Određivanje izotopa olova masenom spektrometrijom sa indukovano spregnutom plazmom*“ odbranio je sa ocenom 10. Nagrađen je za visok prosek u trećoj i završnoj godini studija nagradom fonda „Panta S. Tutundžić“.

Školske 2009/2010. godine upisao je doktorske studije na Tehnološko-metalurškom fakultetu Univerziteta u Beogradu, na studijskom programu Inženjerstvo zaštite životne sredine. U okviru doktorskih studija položio je sve ispite predviđene studijskim programom, sa prosečnom ocenom 9,58. Završni ispit pod nazivom „*Razvoj modela neuronskih mreža za modelovanje indikatora održivog razvoja*“ je odbranio u septembru 2011. godine sa ocenom 10.

Od novembra 2009. godine angažovan je na Projektima osnovnih istraživanja Ministarstva nauke Republike Srbije (142002 i 172007). Kao član Organizacionog odbora učestvovao je u organizaciji „*Prve konferencije mladih hemičara Srbije*“ koja je održana 19. i 20. oktobra 2012. godine na Tehnološko-metalurškom fakultetu u Beogradu. Takođe je bio angažovan na izradi projekta „*Lokalni registar izvora zagađivanja životne sredine na teritoriji Beograda*“ koji su od 2010 do 2013. godine realizovali Sekretarijat za zaštitu životne sredine Grada Beograda i Tehnološko-metalurški fakultet Univerziteta u Beogradu.

Koautor je 11 radova objavljenih u časopisima međunarodnog i nacionalnog značaja i 16 naučnih saopštenja u zbornicima radova sa međunarodnih i nacionalnih skupova.

Izjava o autorstvu

Potpisani-a Davor Z. Antansijević

broj indeksa 4009/2009

Izjavljujem

da je doktorska disertacija pod naslovom

Modelovanje indikatora kvaliteta životne sredine primenom veštačkih neuronskih mreža

- rezultat sopstvenog istraživačkog rada,
- da predložena disertacija u celini ni u delovima nije bila predložena za dobijanje bilo koje diplome prema studijskim programima drugih visokoškolskih ustanova,
- da su rezultati korektno navedeni i
- da nisam kršio/la autorska prava i koristio intelektualnu svojinu drugih lica.

Potpis doktoranda

U Beogradu, 14.02.2014.



Izjava o istovetnosti štampane i elektronske verzije doktorskog rada

Ime i prezime autora Davor Z. Antanasijević

Broj indeksa 4009/2009

Studijski program Inženjerstvo zaštite životne sredine

Naslov rada Modelovanje indikatora kvaliteta životne sredine primenom veštačkih neuronskih mreža

Mentor Dr Viktor Pocajt, vanredni profesor

Potpisani/a Davor Z. Antanasijević

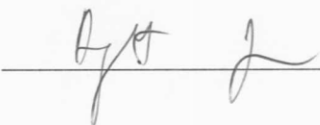
Izjavljujem da je štampana verzija mog doktorskog rada istovetna elektronskoj verziji koju sam predao/la za objavljivanje na portalu Digitalnog repozitorijuma Univerziteta u Beogradu.

Dozvoljavam da se objave moji lični podaci vezani za dobijanje akademskog zvanja doktora nauka, kao što su ime i prezime, godina i mesto rođenja i datum odbrane rada.

Ovi lični podaci mogu se objaviti na mrežnim stranicama digitalne biblioteke, u elektronskom katalogu i u publikacijama Univerziteta u Beogradu.

Potpis doktoranda

U Beogradu, 14.02.2014.



A handwritten signature in black ink, appearing to read 'Davor Z. Antanasijević', is written over a horizontal line.

Izjava o korišćenju

Ovlašćujem Univerzitetsku biblioteku „Svetozar Marković“ da u Digitalni repozitorijum Univerziteta u Beogradu unese moju doktorsku disertaciju pod naslovom:

Modelovanje indikatora kvaliteta životne sredine primenom veštačkih neuronskih mreža

koja je moje autorsko delo.

Disertaciju sa svim priložima predao/la sam u elektronskom formatu pogodnom za trajno arhiviranje.

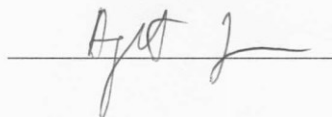
Moju doktorsku disertaciju pohranjenu u Digitalni repozitorijum Univerziteta u Beogradu mogu da koriste svi koji poštuju odredbe sadržane u odabranom tipu licence Kreativne zajednice (*Creative Commons*) za koju sam se odlučio/la.

1. Autorstvo
2. Autorstvo - nekomercijalno
3. **Autorstvo – nekomercijalno – bez prerade**
4. Autorstvo – nekomercijalno – deliti pod istim uslovima
5. Autorstvo – bez prerade
6. Autorstvo – deliti pod istim uslovima

(Molimo da zaokružite samo jednu od šest ponuđenih licenci, kratak opis licenci dat je na poleđini lista).

U Beogradu, 14.02.2014.

Potpis doktoranda



A handwritten signature in black ink, appearing to be 'A. J.', is written over a horizontal line.

1. Autorstvo - Dozvoljavate umnožavanje, distribuciju i javno saopštavanje dela, i prerade, ako se navede ime autora na način određen od strane autora ili davaoca licence, čak i u komercijalne svrhe. Ovo je najslobodnija od svih licenci.

2. Autorstvo – nekomercijalno. Dozvoljavate umnožavanje, distribuciju i javno saopštavanje dela, i prerade, ako se navede ime autora na način određen od strane autora ili davaoca licence. Ova licenca ne dozvoljava komercijalnu upotrebu dela.

3. Autorstvo - nekomercijalno – bez prerade. Dozvoljavate umnožavanje, distribuciju i javno saopštavanje dela, bez promena, preoblikovanja ili upotrebe dela u svom delu, ako se navede ime autora na način određen od strane autora ili davaoca licence. Ova licenca ne dozvoljava komercijalnu upotrebu dela. U odnosu na sve ostale licence, ovom licencom se ograničava najveći obim prava korišćenja dela.

4. Autorstvo - nekomercijalno – deliti pod istim uslovima. Dozvoljavate umnožavanje, distribuciju i javno saopštavanje dela, i prerade, ako se navede ime autora na način određen od strane autora ili davaoca licence i ako se prerada distribuira pod istom ili sličnom licencom. Ova licenca ne dozvoljava komercijalnu upotrebu dela i prerada.

5. Autorstvo – bez prerade. Dozvoljavate umnožavanje, distribuciju i javno saopštavanje dela, bez promena, preoblikovanja ili upotrebe dela u svom delu, ako se navede ime autora na način određen od strane autora ili davaoca licence. Ova licenca dozvoljava komercijalnu upotrebu dela.

6. Autorstvo - deliti pod istim uslovima. Dozvoljavate umnožavanje, distribuciju i javno saopštavanje dela, i prerade, ako se navede ime autora na način određen od strane autora ili davaoca licence i ako se prerada distribuira pod istom ili sličnom licencom. Ova licenca dozvoljava komercijalnu upotrebu dela i prerada. Slična je softverskim licencama, odnosno licencama otvorenog koda.