

UNIVERZITET U BEOGRADU
TEHNOLOŠKO-METALURŠKI FAKULTET

Lidija J. Stamenković

**PREDVIĐANJE EMISIJE GASOVITIH
ZAGAĐUJUĆIH MATERIJA NA
NACIONALNOM NIVOU PRIMENOM
MODELA ZASNOVANIH NA VEŠTAČKIM
NEURONSKIM MREŽAMA**

doktorska disertacija

Beograd, 2016.

UNIVERSITY OF BELGRADE
FACULTY OF TECHNOLOGY AND METALLURGY

Lidija J. Stamenković

**PREDICTION OF EMISSION OF
GASEOUS POLLUTANTS ON NATIONAL
LEVEL USING ARTIFICIAL NEURAL
NETWORKS MODELS**

Doctoral Dissertation

Belgrade, 2016

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: _____

ZAHVALNICA

Ova doktorska disertacija je urađena na Katedri za inženjerstvo zaštite životne sredine Tehnološko-metalurškog fakulteta Univerziteta u Beogradu, u periodu od 2012. do 2016. godine.

Na predloženoj temi, savetima i ukazanoj pomoći pri izradi ove doktorske disertacije, izražavam iskrenu zahvalnost mom mentoru, prof. dr Viktoru Pocajtu. Profesoru Pocajtu se zahvaljujem što mi je omogućio da se bavim ovom aktuelnom tematikom, bila mi je čast i zadovoljstvo da radim sa njim i učim od njega.

Posebnu zahvalnost dugujem dr Davoru Antanasijeviću, naučnom saradniku IC TMF-a, koji je bio uključen u sve segmente izrade ove doktorske disertacije. Njegova stručnost, znanje i nesebični saveti su značajno doprineli kvalitetu ove teze.

Takođe bih želela da se zahvalim prof. dr Mirjani Ristić i prof. dr Aleksandri Perić Grujić na saradnji, stručnoj pomoći, kao i na sadržajnim sugestijama i dobronamernim primedbama tokom naučno-istraživačkog rada u okviru ove doktorske disertacije.

Na kraju bih želela da izrazim ogromnu zahvalnost svojoj porodici, čije su mi strpljenje, stalna moralna i materijalna podrška i razumevanje umnogome olakšali uspešan završetak doktorskih studija.

Predviđanje emisije gasovitih zagađujućih materija na nacionalnom nivou primenom modela zasnovanih na veštačkim neuronskim mrežama

REZIME

Radi realizacije koncepta održivog razvoja u narednim decenijama, kao jedan od značajnijih segmenata jeste očuvanje i kontrola kvaliteta vazduha. U tom smislu su na globalnom nivou osnovane brojne organizacije, podržane određenim međunarodno zakonodavno-pravnim mehanizmima. Obaveza država članica pomenutih organizacija i potpisnica konvencija je podnošenje izveštaja o trenutnim i budućim emisijama određenih zagađujućih materija, definisanih kroz indikatore koji se odnose na vazduh i klimatske promene. Radi realizacije preuzetih obaveza potrebno je primeniti odgovarajuće modele koji će na što precizniji, jednostavniji i ekonomičniji način proceniti emisiju određenih polutanata u vazduh.

Postojeći modeli koji se koriste za proračun trenutnih i budućih emisija zagađujućih materija su zasnovani na inventarskom pristupu i podrazumevaju poznavanje i primenu velikog broja ulaznih parametara. Da bi proračun emisije prema postojećim modelima bio što precizniji, zahteva se poznavanje više stotina specifičnih parametara za određenu državu i svaki od izvora emisije, koji zavise od primenjene tehnologije, vrste goriva, kao i drugih informacija. Određivanje ulaznih parametara za postojeće modele je izuzetno složeno i zahteva korišćenje brojnih resursa da bi se utvrdila njihova vrednost.

Poslednjih godina, sa razvojem računarske tehnologije, veštačke neuronske mreže (ANN - *Artificial Neural Networks*) su vrlo često korišćene za modelovanje u različitim oblastima. Predstavljaju sofisticirane tehnike modelovanja koje su u mogućnosti da modeluju veoma kompleksne i nelinearne funkcije. U okviru ove disertacije osnovni ciljevi su bili razvoj ANN modela za predviđanje nacionalnih emisija sledećih gasovitih zagađujućih materija: amonijaka, nemetanskih isparljivih organskih jedinjenja, metana, azotovih oksida i gasova staklene bašte.

Za razvoj ANN modela za predviđenje emisije amonijaka korišćena je višeslojna perceptron arhitektura (MLP - *Multilayer Perceptron*) – troslojna mreža. MLP model je najpre

optimizovan primenom proba i greška procedure kojom je određen optimalan broj skrivenih neurona, aktivaciona funkcija i *backpropagation* algoritam obučavanja. Analiza glavnih komponenti (PCA - *Principal Component Analysis*) je primenjena na originalnim ulaznim podacima radi redukcije korelacije između ulaznih promenljivih. Dobijeni rezultati ANN modela kreiranih sa transformisanim ulazima, tj. glavnim komponentama (PCA - MLP) su pokazali da ima mnogo bolje performanse u odnosu na ANN model kreiran sa originalnim ulaznim promenljivama (MLP). U fazi validacije modela, kreirani MLP i PCA - MLP modeli su poređeni sa regresionim modelom razvijenim sa glavnim komponentama, kao ulaznim parametrima (PCR - *Principal Component Regression*). Rezultati poređenja ova tri modela su pokazali da PCA - MLP model daje najbolje rezultate predviđanja sa relativnom greškom ispod 20% za SAD i većinu Evropskih država koje su bile uključene u razvoj modela.

ANN model za predviđanje emisije nemetanskih isparljivih organskih jedinjenja (NMVOC - *Non-Methane Volatile Organic Compounds*) je razvijen za Kinu za period 2005 – 2011. godine. Predviđena vrednost emisije NMVOC za Kinu kreiranim ANN modelom je poređena sa vrednostima dobijenu inventarskim pristupom. ANN model za Kinu je razvijen korišćenjem dostupnih podataka za devet država Evropskih unije, koje su po stepenu emisije NMVOC po glavi stanovnika najbližije Kini, za period 2004 - 2012. godine. Za odabir najznačajnijih ulaznih promenljivih za model, kako za Kinu tako i za EU države, primenjena je *forward* strategija. Nakon odabira najznačajnijih ulaznih promenljivih, od inicijalno odabranih, finalni ANN model je kreiran sa samo pet ulaznih promenljivih. Finalni ANN model je pokazao slične performanse u predviđanju, kako za EU države čiji su podaci korišćeni za razvoj modela, tako i za Kinu čiji su podaci prezentovani modelu tek u fazi validacije, tj. provere stvarnih mogućnosti razvijenog modela sa novim podacima. Dobijene su vrednosti srednje apsolutne procentualne greške (*MAPE - Mean Absolute Percentage Error*) za EU države 8% i Kinu 5%. Kreirani ANN model je pokazao slaganje trenda promene NMVOC emisije u Kini sa trendom promene emisije dobijenim inventarskim pristupom.

Za razvoj ANN modela za predviđanje nacionalne emisije metana korišćene su dve arhitekture neuronskih mreža; troslojna *backpropagation* neuronska mreža (BPNN -

Backpropagation Neural network) i generalna regresiona neuronska mreža (GRNN - *General Regression Neural Network*). Za poređenje performansi razvijenog ANN modela, uporedo je razvijen i model zasnovan na višestrukoj linearnoj regresiji (MLR - *Multiple Linear Regression*). Modeli su razvijani i testirani podacima za dvadeset EU država za period 2000 - 2008. godine. Dobijeni rezultati su pokazali da ANN model daje veoma dobre rezultate predviđanja emisije metana, značajno bolje u odnosu na kreirani MLR model.

ANN model za predviđanje emisije azotovih oksida (NO_x) je razvijen korišćenjem dostupnih podataka (2001 - 2008. godine) za 17 EU država, SAD, Kinu, Indiju, Japan i Rusiju. Za razvoj modela je korišćena GRNN arhitektura i inicijalni ANN model je razvijen primenom 11 potencijalnih lako dostupnih ekonomskih i indikatora održivog razvoja kao ulaznih promenljivih. Za odabir najznačajnijih i nekorelisanih ulaznih promenljivih primenjene su korelaciona i faktor povećanja varijanse (VIF - *Variance Inflation Factor*) analiza. Najbolje rezultate je pokazao model kod koga je redukcija broja ulaznih promenljivih izvršena na osnovu korelacije između ulaznih promenljivih. Rezultati kreiranog modela su značajno bolji u odnosu na GRNN model koji je bio kreiran sa svim ulaznim promenljivama.

Jedan od ciljeva u okviru ovog rada je bio i da se razvije ANN model za predviđanje emisija gasova staklene bašte za Republiku Srbiju. Zbog nedostatka podataka za Republiku Srbiju korišćeni su dostupni podaci za Republiku Bugarsku, koja je po osnovnim karakteristikama najsljednija Srbiji. Kao ulazne promenljive za razvoj modela su korišćena samo dva indikatora održivog razvoja; bruto domaći proizvod i godišnja proizvodnja energije. Za razvoj modela je korišćena rekurentna arhitektura neuronske mreže (RNN - *Recurrent Neural Network*). Proverom mogućnosti predviđanja kreiranog RNN modela na validacionim podacima za Bugarsku, dobijena je vrednost greške od 16%. Razvijeni model je zatim primenjen na Srbiju i pokazao je odstupanje od samo 3% između proračunate i modelom predviđene vrednosti za 1998. godinu.

Na osnovu dobijenih rezultata, pristupom zasnovanim na neuronskim mrežama i široko dostupnim ekonomskim, industrijskim i indikatorima održivog razvoja, kao ulaznim promenljivim, koji je primenjen u okviru ove doktorske disertacije, može se zaključiti da se

primenjena metodologija može primeniti za predviđanje emisije zagađujućih materija u vazduh na nacionalnom nivou.

Ključne reči: održivi razvoj, veštačke neuronske mreže, amonijak, nemetanska isparljiva organska jedinjenja, metan, azotni oksidi, gasovi staklene bašte.

Naučna oblast: Tehničko-tehnološke nauke

Uža naučna oblast: Inženjerstvo zaštite životne sredine

UDK broj: 628.4.08 : 004.032.26

Prediction of emission of gaseous pollutants on national level using artificial neural networks models

SUMMARY

In order to implement concept of sustainable development in the coming decades, one of significant segments is to prevent further degradation of air quality influenced by emission of pollutants. Regarding this, numerous organizations have been founded on global level, supported by certain international legislative-legal mechanisms. Obligation of member countries and convention signees is submission of reports on current and future emissions of specific pollutants defined through indicators regarding air and climate changes. In order to realize commitments, it is necessary to apply suitable models which will, in the simplest, most precise and most economical way estimate the emission of certain pollutants into the air.

Existing models that are used for estimation of current and future emission of pollutants are based on inventory approach and imply knowledge and implementation of numerous input parameters. In order to estimate emission more precisely according to existing models, it is necessary to have the knowledge of hundreds of specific parameters for certain country as well as every emission source, that depend on applied technology, type of fuel and other information. Setting input parameters for existing models is extremely complicated and requires use of numerous resources so as to determine their value.

In recent years, along with development of computer technology, artificial neural networks (ANN) have often been used for modeling in different fields. They present sophisticated modeling techniques that are able to model very complex and nonlinear functions. Basic goals within this dissertation was development of ANN model for prediction of national emissions of following gaseous pollutants: ammonia, non-methane volatile organic compounds, methane, nitrogen oxides and greenhouse gases.

Multilayer perceptron architecture (MLP) - three-layer network was used for the development of ANN model for estimation of ammonia emission. MLP model was firstly

optimized by application of trials and errors of the procedure by which the optimum number of hidden neurons, activation function and back-propagation training algorithm is determined. Principal Components Analysis (PCA) is applied on original input data in order to reduce correlation between input variables. Obtained results of ANN model, created with transformed inputs, or, principal components (PCA - MLP) have shown that it has much better performance in comparison with ANN model, created with original input variables (MLP). In the phase of model validation, created MLP and PCA - MLP models are compared with regression model developed with principal components as input parameters (PCR). Comparison results of these three models have shown that PCA - MLP model provides best estimation results with relative error under 20% for USA and most European countries that are involved in model development.

ANN model for estimation of non-methane volatile organic compounds (NMVOC) has been developed for China for the period from 2005 - 2011. Estimated emission value of NMVOC for China with created ANN model has been compared to values obtained in inventory approach. ANN model for China has been developed using available data for nine European countries which are, by the emission NMVOC level per capita, most similar to China for the period from 2004 - 2012. *Forward* strategy was applied for selection of most significant input variables, for China as well as for EU countries. After the selection of most significant input variables, from initially selected, final ANN model was created with only five input variables. Final ANN model has shown similar performances in estimation for EU countries, whose data were used for model development, as well as for China, whose data were presented to the model in the validation phase, that is, during audit of real possibilities of developed model on new data. MAPE values have been obtained for EU countries 8% and China 5%. Created ANN model has shown matching of the trend of change in NMVOC emission in China to the trend of change in emission obtained by inventory approach.

Two architectures of neural networks; three-layered backpropagation neuronal network (BPNN) and general regressive neuronal network (GRNN) were used for the development of ANN model for estimation of national emission of methane. A model based on multiple linear regression (MLR) was developed at the same time for comparison of performances of

ANN model. Models are developed and tested according to data for twenty EU countries for the period from 2000 - 2008. Obtained results have shown that ANN model provides very good results in estimation of methane emission, significantly better comparing to created MLR model.

ANN model for estimation of nitrogen oxides emission (NO_x) has been developed by using available data (2001 - 2008.) for 17 EU countries, USA, China, India, Japan and Russia. GRNN architecture was used for the development of the model and initial ANN model was developed by using all 11 potential widely available economic indicators of sustainable development as input variables. For selection of most significant and non-correlated input variables, correlation and variance inflation factor (VIF) analysis were applied. Best results have been shown by the model in which the reduction of number of input variables was undertaken based on correlation between input variables. Results of created model were significantly better compared to GRNN model, created with all input variables.

One of the goals regarding this study was to develop ANN model for estimation of greenhouse gases emission for Republic of Serbia. Due to lack of data for Republic of Serbia, available data for Republic of Bulgaria were used, which is, according to basic characteristics, most similar to Serbia. Input variables, used for model development, were only two indicators of sustainable development; gross domestic product and annual energy production. Recurrent neural network architecture (RNN) was used for model development. By checking estimation possibilities of created RNN model on validation data for Bulgaria, error value of 16% was obtained. Developed model was then applied onto Serbia and has shown deviation of only 3% between calculated value and value estimated by the model for the year 1998.

According to obtained results, by approach based on neural networks and widely available economic, industrial and indicators of sustainable development as input variables, which was applied within this doctoral dissertation, it can be concluded that applied methodology can be applied for estimation of national emission of pollutants into the air.

Key words: sustainable development, artificial neural networks, ammonia, non-methane volatile organic compounds, methane, nitrogen oxides, greenhouse gases.

Scientific area: Technical-technological sciences

Scientific sub-area: Environmental Engineering

UDC number: 628.4.08 : 004.032.26

Sadržaj

Lista simbola i skraćenica.....	I
Spisak slika.....	VI
Spisak tabela.....	VIII
1. Uvod.....	1
Teorijski deo	
2. Održivi razvoj i očuvanje kvaliteta vazduha.....	4
2.1. Definicija i ciljevi održivog razvoja.....	4
2.1.1. Indikatori kvaliteta životne sredine.....	6
2.2. Prikaz odabranih indikatora iz kategorije kvaliteta vazduha.....	8
2.2.1. Emisija amonijaka (NH ₃).....	9
2.2.2. Emisija nemetanskih isparljivih organskih jedinjenja (NMVOCs).....	10
2.2.3. Emisija azotnih oksida (NO _x).....	11
2.2.4. Emisija metana (CH ₄).....	14
2.2.5. Međunarodni zakonodavno-pravni mehanizmi za smanjenje zagađenja vazduha.....	16
2.3. Postojeći modeli za procenu emisija.....	18
2.3.1. EDGAR v4 model.....	19
2.3.2. EMEP/CORINAIR model.....	20
3. Veštačke neuronske mreže.....	24
3.1. Biološke i veštačke neuronske mreže.....	24
3.2. Osnovni principi funkcionisanja veštačkih neuronskih mreža.....	26
3.3. Proces razvoja modela zasnovanih na veštačkim neuronskim mrežama.....	30
3.3.1. Odabir ulaznih promenljivih za razvoj modela.....	32
3.3.2. Odabir arhitekture mreže.....	35
3.3.3. Obučavanje i validacija neuronskih mreža.....	36
3.4. Primena neuronskih mreža za modelovanje emisija zagađujućih materija.....	37
Eksperimentalni deo	
4. Materijali i metode.....	39

4.1.	Prikupljanje ulaznih i izlaznih podataka	39
4.2.	Arhitekture mreža primenjene za razvoj modela	40
4.2.1.	Višeslojni perceptron (MLP)	40
4.2.2.	Generalne regresione neuronske mreže (GRNN)	42
4.2.3.	Rekurentne neuronske mreže (RNN)	44
4.3.	Odabrane metode primenjene za selekciju ulaznih promenljivih	45
4.3.1.	Analiza glavnih komponenti (PCA)	45
4.3.2.	Korelaciona i Faktor povećanja varijanse (VIF) analiza	49
4.4.	Regresioni modeli za komparaciju rezultata ANN modela	52
4.5.	Statistički indikatori performansi modela	54

Rezultati i diskusija

5.	Predviđanje emisije amonijaka u SAD i EU	57
5.1.	Ulazni i izlazni podaci	57
5.2.	Rezultati optimizacije MLP modela	60
5.3.	Rezultati MLP modela kreiranih nakon primene analize glavnih komponenti	64
5.4.	Rezultati validacije MLP modela	66
5.5.	Rezultati analize osetljivosti (značajnosti) ulaznih promenljivih	70
6.	Predviđanje emisije nemetanskih isparljivih organskih jedinjenja u EU i Kini	72
6.1.	Prikupljanje i selekcija podataka	72
6.2.	Razvoj i rezultati razvijenog MLP modela	77
7.	Predviđanje emisije metana u EU	84
7.1.	Ulazne i izlazne promenljive i priprema podataka za razvoj modela	84
7.2.	Razvoj i rezultati modela	87
8.	Predviđanje emisije azotnih oksida u EU, SAD, Kini, Indiji, Japanu i Rusiji	97
8.1.	Ulazni i izlazni podaci	97
8.2.	Razvoj i optimizacija GRNN modela	100
8.3.	Rezultati optimizacije GRNN modela	103
8.4.	Rezultati analize osetljivosti (značajnosti) ulaznih promenljivih	109
9.	Predviđanje emisije gasova staklene bašte u Srbiji	110
9.1.	Ulazne i izlazne promenljive i podaci	110

9.2. Razvoj ANN modela i rezultati	112
10. Zaključak	115
Literatura	119
Biografija autora	
Prilozi	

LISTA SIMBOLA I SKRAĆENICA

- AD - Stopa aktivnosti u određenom sektoru izvoru emisije
- A_k - *Hessian* matrica drugog izvoda funkcije greške
- AL - Površina poljoprivrednog zemljišta (*Agricultural Land*)
- ANN - Veštačke neuronske mreže (*Artificial Neural Networks*)
- BDP - Bruto domaći proizvod
- B_i - Broj izlaznih promenljivih
- BPNN - Neuronska mreža sa propagacijom greške unazad (*Backpropagation Neural Network*)
- B_{tr} - Broj serija podataka u setu podataka za obučavanje
- B_u - Broj ulaznih promenljivih
- CF - Faktor doprinosa (*Contribution Factor*)
- CLRTAP - Konvencija o prekograničnom zagađenju vazduha na velikim udaljenostima (*Convention on Long Range Transboundary Air Pollution*)
- CNF - Potrošnja azotnih đubriva (*Consumption of Nitrogen Fertilizers*)
- CS - Proizvodnja iz hemijske industrije (*Chemical Shipments*)
- CTPP - Ukupna potrošnja naftnih derivata (*Consumption Total Petroleum Products*)
- do - Željeni izlaz
- d_r - Indeks performansi modela (*The index of model performance*)
- E - Greška mreže
- ECH - Potrošnja električne energije u domaćinstvima (*Electricity Consumption by Households*)
- EDGAR - *Emission Database for Global Atmospheric Research*
- EEA - Evropska agencija za zaštitu životne sredine (*European Environment Agency*)
- EF - Emisioni faktor za svaki sektor i tehnologiju
- EGRS - Električna energija dobijena iz obnovljivih izvora (*Electricity Generated from Renewable Sources*)

EM - Godišnja emisija određene zagađujuće materije

EMEP - *European Monitoring and Evaluation Programme*

CORINAIR - *CORe Inventory of AIR emissions*

EOP - Primenjena mera za smanjenje emisije za svaki sektor izvora emisije

EPA - Agencija za zaštitu životne sredine Sjedinjenih Država (*United States Environmental Protection Agency*)

EPC - Potrošnja električne energije (*Electric Power Consumption*)

EPCS - Proizvodnja električne energije iz uglja (*Electricity Production from Coal Sources*)

EPRS - Električna energije proizvedena iz obnovljivih izvora (*Electricity Production from Renewable Sources*)

EU - Evropska unija (*European Union*)

EU27 - Evropska unija sa 27 članica

FAI.25 - Procenat približno tačnih rezultata koje daje model u odnosu na stvarne vrednosti

FECT - Ukupna potrošnja energije u transportu (*Final Energy Consumption in Transport*)

FFEC - Potrošnja energije iz fosilnih goriva (*Fossil Fuel Energy Consumption*)

GAINS - *Greenhouse Gas and Air Pollution Interactions and Synergies*

GDP - Bruto domaći proizvod (*Gross Domestic Product*)

GHGs - Gasovi staklene bašte (*Greenhouse gases*)

GIEC - Bruto unutrašnja potrošnja energije (*Gross Inland Energy Consumption*)

g_k - Trenutni gradijent greške modela

GPE - Godišnja potrošnja energije

GRNN - Generalna regresiona neuronska mreža (*General Regression Neural Network*)

GWP - Potencijal globalnog zagrevanja (*Global Warming Potential*)

IA - Indeks slaganja (*Index of Agreement*)

IIASA - *International Institute for Applied Systems Analysis*

IPCC - Međuvladin panel o klimatskim promenama (*The Intergovernmental Panel on Climate Change*)

ISF - Pojedinačni faktor ujednačavanja (*Individual Smoothing Factor*)

IVS - *Input Variable Selection*

JRC - Zajednički istraživački centar Evropske unije (*Joint Research Centre*)

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

LR - površinu zemljišta korišćenog za uzgajanje pirinča (*Land use – Rice*)

MAE - Srednja apsolutna greška (*the Mean Absolute Error*)

MAPE - Srednja apsolutna procentualna greška (*Mean Absolute Percentage Error*)

MDGs - Milenijumski ciljevi razvoja (*Millenium Development Goals*)

MLP - Višeslojna perceptron mreža (*Multi - layer perceptron*)

MLR - Višestruka linearna regresija (*Multiple Linear Regression*)

MSE - Srednja kvadratna greška (*Mean Squared Error*)

MVP - Proizvodnja vozila (*Motor Vehicle Production*)

MWG - Generisanje komunalnog otpada (*Municipal Waste Generation*)

NC - Broj grla stoke (*Number of Cattle*)

NMVOC - Nemetanska isparljiva organska jedinjenja (*Non-Methane volatile Organic Compounds*)

no - Dobijena vrednost izlazne promenljive

NSE - *The Nash-Sutcliffe Efficiency*

NV - Broj proizvedenih vozila (*Number of Vehicles*)

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

OICA - Međunarodna organizacija proizvođača motornih vozila (fr. *Organisation Internationale des Constructeurs d'Automobiles*, engl. *The International Organization of Motor Vehicle Manufacturers*)

PBL - Holandska agencija za zaštitu životne sredine (*Planbureau voor de Leefomgeving*)

PC - Glavne komponente (*Principal Components*)

PCA - Analiza glavnih komponenti (*Principal Component Analysis*)

PCR - *Principal Component Regression*

PMI - *Partial Mutual Information*

PPG - Primarna proizvodnja prirodnog gasa (*Primary Production of Gas*)

r - Pirsonov koeficijent korelacije

RDFC - Potrošnja dizel goriva u drumskom sektoru (*Road sector diesel fuel consumption*)

RED - Efikasnost smanjenja emisije u odnosu na primenjenu meru

RIVM - Holandski nacionalni institut za javno zdravlje i životnu sredinu (*Rijksinstituut voor Volksgezondheid en Milieu*)

R_j^2 - Koeficijent determinacije

RMSE - Greška korena srednjeg kvadrata (*Root Mean Squared Error*)

RNN - Rekurentna neuronska mreža (*Recurrent Neural Network*)

RSEC - Potrošnja energije u drumskom sektoru (*Road Sector Energy Consumption*)

RSR - RMSE – *observation standard deviation ratio*

RVS - Recursive Variable Selection

SAD - Sjedinjene Američke Države

SCR - Selektivna katalitička redukcija (*Selective Catalytic Reduction*)

SDG - Ciljevi održivog razvoja (*Sustainable Development Goals*)

SNAP - *Selected Nomenclature for Air Pollution*

TECH - Vrsta primenjene tehnologije za svaki sektor izvora emisije

TNO - Holandska organizacija za primenjena naučna istraživanja (*Nederlandse Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek*)

toe - Tona ekvivalenta nafte

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

UNFCCC - Okvirna konvencija Ujedinjenih nacija o promeni klime (*United Nations Framework Convention on Climate Change*)

USDA - Nacionalni poljoprivredni statistički servis Sjedinjenih Država (*U.S. Department of Agriculture*)

USEIA - Nacionalni statistički servis iz oblasti energetike Sjedinjenih Država (*U.S. Energy Information Administration*)

VIF - *Variance Inflation Factor*

WB - Svetska Banka (*World Bank*)

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

WDL - Odlaganje komunalnog otpada na ili ispod površine zemlje (*Waste Deposit onto or into Land*)

w_i - Težinski koeficijenti

x_i - Ulazna promenljiva

y - Izlazni signal iz neurona

Y - Zavisna - izlazna promenljiva

y_k - Konjugovani gradijent greške

α_k - Stopa obučavanja (*learning rate*)

β_k - Impuls (*Momentum*)

λ - Svojtvena vrednost glavne komponente

ρ - Koeficijent Spirmanove korelacije

SPISAK SLIKA

Slika 1. Udeo pojedinih sektora u emisiji NO _x u EU	13
Slika 2. Udeo pojedinih sektora u emisiji NO _x u SAD	13
Slika 3. Biološki neuron.....	24
Slika 4. Šematski prikaz veštačkog neurona	26
Slika 5. Višeslojna neuronska mreža	27
Slika 6. Šematski prikaz procesa razvoja neuronske mreže.....	31
Slika 7. IVS tehnike za odabir ulaznih promenljivih	34
Slika 8. Arhitekture neuronskih mreža.....	36
Slika 9. Katelov kriterijum za određivanje značajnog broja glavnih komponenti [96]	48
Slika 10. Optimizacija MLP mreže – određivanje optimalnog broja skrivenih neurona.....	61
Slika 11. Optimizacija MLP modela – određivanje optimalne aktivacione funkcije	63
Slika 12. Optimizacija MLP modela – određivanje pogodnog BP algoritma.....	64
Slika 13. Poređenje aktuelnih emisija amonijaka i PCA – MLP – V modelom predviđenih vrednosti	66
Slika 14. Rezultati validacije modela. Dijagram aktuelna – predviđena emisija amonijaka za a) MLP – V; b) PCA – MLP – V; c) PCR.....	68
Slika 15. MAPE vrednosti predviđanja PCA- MLP – V i PCR modela za pojedinačne države	70
Slika 16. Doprinos različitih sektora emisiji NMVOCs u EU državama.....	73
Slika 17. Poređenje aktuelne i predviđene vrednosti emisije NMVOC razvijenih ANN modela: a-e testiranje sa EU podacima iz 2012. godine i f-h primena na podacima za Kinu 2005 – 2010.	79
Slika 18. Šematski prikaz NMVOC MLP modela	80
Slika 19. Poređenje aktuelne i modelom predviđene NMVOC emisije za EU države	81
Slika 20. Poređenje a) aktuelne i modelom predviđene NMVOC emisije za Kinu i b) doprinos pojedinih sektora emisiji NMVOC u Kini [207].....	82
Slika 21. Šematski prikaz ANN arhitektura: a) BPNN i b) GRNN	88
Slika 22. Poređenje aktuelnih i predviđenih vrednosti emisije metana za trening podatke a) BNNN, b) GRNN, c) MLR	90

Slika 23. Poređenje aktuelnih i predviđenih vrednosti emisije metana za validacioni set podataka: a) BPNN, b) GRNN, c) MLR	93
Slika 24. GRNN arhitektura.....	100
Slika 25. Poređenje aktuelnih podataka i GRNN1 rezultata (validacioni set podataka)....	104
Slika 26. Rezultati GRNN3d modela za validacioni set podataka: a) poređenje aktuelne i predviđene NO _x emisije, b) <i>FAI.25</i> dijagram.....	106
Slika 27. Relativne greške za a) evropske države b) države van Evrope.....	108
Slika 28. Individualni faktori ujednačavanja (<i>ISF</i>) za ulazne promenljive u GRNN3d modelu	109
Slika 29. Poređenje Srbije i Bugarske: a) položaj, broj stanovnika i površina, b) vrednost BDP-a i GPE-a [220].....	111
Slika 30. Rekurentna arhitektura (RNN) [220]	113
Slika 31. Predviđanje GHG emisije za Srbiju [220]	114

SPISAK TABELA

Tabela 1. Protokoli LRTAP Konvencije	17
Tabela 2. Jedanaest glavnih sektora u CORINAIR metodologiji [65]	21
Tabela 3. Lista ulaznih promenljivih za razvoj ANN modela i osnovni statistički pokazatelji	58
Tabela 4. Pirsonov koeficijent korelacije između ulaznih promenljivih i odgovarajući faktor povećanja varijanse (VIF)	59
Tabela 5. Statistički pokazatelji podataka emisije amonijaka za period od 2000. do 2010. godine	60
Tabela 6. Testirane aktivacione funkcije u procesu optimizacije MLP mreže.....	62
Tabela 7. Rezime primenjene PCA: komunalitet, KMO vrednost i Bartletov test	65
Tabela 8. Faktorske težine, strukturna matrica, svojstvene vrednosti, varijansa po komponenti, kao i ukupna varijansa objašnjena izdvojenim glavnim komponentama	65
Tabela 9. Vrednosti statističkih pokazatelja performansi modela (faza validacije modela).....	69
Tabela 10. Statistička analiza potencijalnih ulaznih promenljivih za odabrane EU države	75
Tabela 11. Statistička analiza emisije NMVOC (kg po st.) za izabrane EU države za period 2004 – 2012	76
Tabela 12. Podaci za Kinu	76
Tabela 13. R^2 vrednosti za kreirane ANN modele za svaku ulaznu promenljivu pojedinačno	78
Tabela 14. Lista ulaznih i izlaznih promenljivih	86
Tabela 15. Podaci za obučavanje za 2006. godinu	87
Tabela 16. Vrednosti indikatora performansi kreiranih modela.....	89
Tabela 17. Validacioni set podataka (vrednosti za 2007. i 2008. godinu).....	92
Tabela 18. Aktuelne i GRNN modelom predviđene vrednosti emisije metana sa relativnim greškama za 2007. i 2008. godinu	94
Tabela 19. Statistička analiza 11 inicijalnih ulaznih promenljivih.....	98
Tabela 20. Statistička analiza NOx emisije na nacionalnom nivou	99
Tabela 21. Detaljne informacije kreiranih GRNN modela	101
Tabela 22. Rezultati korelacione analize	102

Tabela 23. VIF analiza.....	102
Tabela 24. Vrednosti indikatora performansi kreiranih GRNN modela	103
Tabela 25. Ulazni i izlazni podaci za razvoj ANN modela	112

1. Uvod

Zaštita životne sredine, klimatske promene i zagađivanje vazduha su postali značajni problemi sa kojima se suočava čovek u današnjem vremenu. Zagađenje vazduha zbog potencijalnog prekograničnog transporta predstavlja globalni problem za čije je rešavanje potrebna međunarodna saradnja svih zemalja, a posebno onih koje predstavljaju najveće emitere zagađujućih materija. U tom smislu je postalo neophodno praćenje emisije svih onih zagađujućih materija koje deluju štetno, kako na zdravlje čoveka, tako i na životnu sredinu. Kao deo globalnih napora za redukciju emisije različitih polutanata u vazduh, Konvencija o prekograničnom zagađenju vazduha na velikim udaljenostima i Okvirna konvencija Ujedinjenih nacija o promeni klime, kao i njihovi protokoli, podrazumevaju obavezu članica, država potpisnica ovih konvencija, podnošenje godišnjih izveštaja trenutnih i budućih emisija određenih polutanata. Da bi države ispunile svoje preuzete obaveze, neophodna je primena odgovarajućih metoda i modela za procenu sadašnjih i simulaciju budućih emisija zagađujućih materija.

Prema postojećim modelima i metodologijama, procena emisije svih zagađujućih materija u vazduh je zasnovana na inventarskom pristupu. Kod ovog pristupa emisija zagađujućih materija se procenjuje na osnovu dva osnovna parametra: vrednosti emisionih faktora i stopa aktivnosti za svaki izvor emisije, kao i nekih drugih parametara koji zavise od primenjenog modela. To podrazumeva detaljno poznavanje više stotina specifičnih vrednosti ova dva parametra za određeni proces i primenjenu tehnologiju, za čije je određivanje potrebno određeno vreme, ljudski resursi i izvesna finansijska ulaganja. Brojne države, a pogotovu one u razvoju, usled nedostatka podataka o ovim parametrima, koriste vrednosti emisionih faktora koje su određene od strane međunarodnih institucija, a koji nisu specifični za tu državu i proces, što za posledicu ima nesigurnost podataka procenjene emisije.

Kao jedna od mogućih alternativa postojećim modelima za predviđanje emisije zagađujućih materija u vazduh u ovom radu se predlaže pristup modelovanja zasnovan na veštačkim neuronskim mrežama. Očigledna veza između ekonomskog i privrednog razvoja i produkcije

zagađujućih materija, otvara mogućnost primene lako dostupnih ekonomskih, industrijskih, poljoprivrednih i ostalih indikatora održivog razvoja kao ulaznih promenljivih za razvoj modela zasnovanih na neuronskim mrežama. Osnovni cilj istraživanja opisanih u okviru ove doktorske disertacije je razvoj modela zasnovanog na veštačkim neuronskim mrežama i ekonomskim i indikatorima održivog razvoja kao ulaznih promenljivih, za predviđanje nacionalnih emisija gasovitih zagađujućih materija u vazduh. Predmet disertacije je razvoj modela za predviđanje emisije sledećih gasovitih zagađujućih materija: amonijaka, nemetanskih isparljivih organskih jedinjenja, metana, azotnih oksida i gasova staklene bašte. Emisije navedenih zagađujućih materija predstavljaju indikatore kvaliteta životne sredine koji se odnose na vazduh i klimatske promene. Pri razvoju modela za predviđanje nacionalnih emisija navedenih zagađujućih materija, pošlo se od pretpostavke da su ulazne – nezavisne promenljive (ekonomski, industrijski, parametri održivog razvoja itd.) i izlazne – zavisne promenljive (emisije gasovitih zagađujućih materija) nelinearno povezane.

Na osnovu polaznih hipoteza, definisanog predmeta i naučnih ciljeva istraživanja, doktorska disertacija je prikazana kroz deset poglavlja.

U prvom poglavlju koje predstavlja uvod u istraživanje, dat je kratak pregled problema koji se tretira u okviru ove disertacije, zatim, kratak pregled postojećih modela koji se koriste za proračun emisija, kratak opis pristupa koji se predlaže u ovom radu, definisani su predmet i ciljevi istraživanja.

Sledeća dva poglavlja, Održivi razvoj i očuvanje kvaliteta vazduha i Veštačke neuronske mreže, čine Teorijski deo disertacije. U teorijskom delu disertacije, najpre je data definicija održivog razvoja, značaj i postavljeni ciljevi, kao i definisani indikatori kvaliteta životne sredine, sa posebnim osvrtom na indikatore koji se odnose na vazduh i klimatske promene, a čija je emisija predmet istraživanja u okviru ove doktorske disertacije. Nakon toga je dat pregled međunarodno zakonodavno-pravnih mehanizama koji su uspostavljeni radi smanjenja emisija. Zatim je dat prikaz nekih od postojećih modela koji su zasnovani na inventarskom pristupu i koji su načešće korišćeni modeli za procenu emisije. U poglavlju koje se odnosi na teorijsku osnovu veštačkih neuronskih mreža, opisani su osnovni principi

funkcionisanja neuronskih mreža, proces razvoja modela zasnovanih na neuronskim mrežama, pregled tehnika za odabir ulaznih promenljivih, arhitekture neuronskih mreža, proces obučavanja kao i literaturni pregled primene neuronskih mreža za modelovanje emisija zagađujućih materija u atmosferu.

U poglavlju Eksperimentalni deo, opisani su materijali i metode korišćeni pri razvoju modela. U okviru ovog poglavlja postoje pet potpoglavlja: Prikupljanje ulaznih i izlaznih podataka, Arhitekture mreža primenjene za razvoj modela, Odabrane metode primenjene za selekciju ulaznih promenljivih, Regresioni modeli za komparaciju rezultata ANN modela i Statistički indikatori performansi modela koji su korišćeni za analizu predikcionih mogućnosti razvijenih modela. Navedena potpoglavlja detaljnije opisuju izvore podataka, odabranih arhitektura, kao i metoda koje su korišćene u ovoj doktorskoj disertaciji za razvoj i optimizaciju ANN modela, kao i detalje o regresionim modelima koji su korišćeni za komparaciju postignutih rezultata predviđanja.

Rezultati i diskusija su predstavljeni kroz pet poglavlja: Predviđanje emisije aminijaka u EU i SAD, Predviđanje emisije nemetanskih isparljivih organskih jedinjenja u Kini, Predviđanje emisije metana u EU, Predviđanje emisije azotovih oksida i Predviđanje emisije gasova staklene bašte u Srbiji. Svako od navedenih potpoglavlja sadrži detaljno opisan proces razvoja modela zasnovanih na neuronskim mrežama, koji se sastoji od odabira potencijalnih ulaznih promenljivih i njihove pripreme za razvoj modela, optimizacije parametara mreže i razvoj inicijalnog ANN modela, primena neke od metoda za selekciju ulaznih promenljivih radi odabira značajnih i nekorelisanih ulaza za razvoj finalnog modela, analize značajnosti pojedinačnih ulaznih promenljivih i analiza performansi kreiranih modela.

U desetom poglavlju je dat kratak prikaz sprovedenih istraživanja u okviru doktorske disertacije, sa zaključcima koji su proistekli na osnovu dobijenih rezultata, kao i pravci budućih istraživanja.

Teorijski deo

2. Održivi razvoj i očuvanje kvaliteta vazduha

2.1. Definicija i ciljevi održivog razvoja

Ubrzani industrijski i privredni razvoj poslednjih decenija za posledicu ima degradaciju životne sredine. U tom smislu se, kako na globalnom, tako i na regionalnom i nacionalnom nivou, čine brojni napori radi smanjenja tih posledica, a kroz usvajanje i primenu zakonskih mera koje uspostavljaju ravnotežu između ekonomskog i industrijskog razvoja i zagađenja životne sredine. I upravo pojam održivog razvoja koji je uveden 1987. godine od strane Svetske komisije za životnu sredinu i razvoj (WCED - *World Commission on Environment and Development*) podrazumeva usklađivanje između ekonomskog i socijalnog razvoja i pritisaka kojima je izložena životna sredina i društvo. Naime, u izveštaju pod nazivom "Naša zajednička budućnost" navodi se da bi usvajanje principa održivog razvoja doprinelo zadovoljavanju potreba sadašnjih generacija, bez ugrožavanja mogućnosti budućim generacijama da zadovolje svoje potrebe [1]. Radi realizacije koncepta održivog razvoja, razvojne politike na globalnom nivou proteklih godina bile su usmeravane u pravcu postizanja osam usvojenih milenijumskih ciljeva razvoja (MDGs - *Millenium Development Goals*). Milenijumski ciljevi razvoja usvojeni su od strane 189 zemalja članica Ujedinjenih nacija (UN - *United Nations*) 2000-te godine sa ciljem ostvarivanja do 2015-te godine, pri čemu je 1990-ta godina bila osnov za procenu napretka ka ostvarivanju ciljeva [2]. Nakon isteka perioda za ostvarivanje MDGs, UN su 2015. razvile novu agendu baziranu na principima održivog razvoja, a koja će u narednim decenijama usmeravati razvojne politike na globalnom nivou [3]. Nova agenda, koju je usvojilo 193 članica UN, sadrži 17 ciljeva održivog razvoja (SDG - *Sustainable Development Goals*) i 169 "meta" koje bi trebalo ostvariti do 2030. godine [4]:

1. Siromaštvo - okončati siromaštvo svuda i u svim oblicima;
2. Glad i bezbednost hrane - okončati glad, postići bezbednost hrane i poboljšanu ishranu i promovisati održivu poljoprivredu;
3. Zdravlje - obezbediti zdrav život i promovisati blagostanje za sve u svim dobima;

4. Obrazovanje - obezbediti inkluzivno i kvalitetno obrazovanje i promovisati mogućnosti doživotnog učenja;
5. Rodna ravnopravnost - postići rodnu ravnopravnost i osnaživati žene i devojčice;
6. Voda i kanalizacija - obezbediti pristup pijaćoj vodi i obezbediti sanitarne uslove za sve;
7. Energija - obezbediti pristup dostupnoj, pouzdanoj, održivoj i modernoj energiji za sve;
8. Ekonomski razvoj - promovisati inkluzivan i održiv ekonomski rast, zaposlenost i dostojan rad za sve;
9. Infrastruktura, industrijalizacija - izgraditi prilagodljivu infrastrukturu, promovisati održivu industrijalizaciju i podržavati inovativnost;
10. Nejednakost - smanjiti nejednakost između i unutar država;
11. Gradovi - učiniti gradove i ljudska naselja inkluzivnim, bezbednim, prilagodljivim i održivim;
12. Održiva potrošnja i proizvodnja - obezbediti održive oblike potrošnje i proizvodnje;
13. Klimatske promene - preduzeti hitnu akciju u borbi protiv klimatskih promena i njenih posledica;
14. Okeani - sačuvati i održivo koristiti okeane, mora i morske resurse;
15. Biodiverzitet, šume, dezertifikacija - zaštititi, obnoviti i promovisati održivo upravljanje šumama, sprečiti dezertifikaciju zemljišta i zaustaviti gubitak biodiverziteta;
16. Mir i pravda - promovisati miroljubiva i inkluzivna društva;
17. Partnerstva - revitalizovati globalno partnerstvo za održivi razvoj.

Kao rezultat globalne saradnje velikog broja stručnjaka iz UN organizacija, akademske zajednice, civilnog društva, biznisa i velikog broja nacionalnih statističkih organizacija, dogovoreno je da je praćenje uspostavljanja koncepta održivog razvoja, pa samim tim i postizanje milenijumskih ciljeva, moguće uz primenu seta indikatora održivog razvoja. U tom smislu je 2015. godine od strane UN objavljen izveštaj pod nazivom ” *Indicators and a Monitoring Framework for the Sustainable Development Goals*” koji predlaže 100 globalnih

indikatora i daje predloge za razvoj komplementarnih nacionalnih indikatora [5]. U skladu sa Izveštajem, svaka država treba da uspostavi određen broj i opseg komplementarnih nacionalnih indikatora koji su u skladu sa potrebama i kapacitetima za prikupljanje i analizu podataka. Indikatori predstavljaju okosnicu za praćenje napretka ka MDG na lokalnom, nacionalnom, regionalnom i globalnom nivou i značajno sredstvo da se pomogne državama da razviju strategiju razvoja koja će biti usklađena sa zadatim ciljevima održivog razvoja.

2.1.1. Indikatori kvaliteta životne sredine

S obzirom na to da zagađenje životne sredine ima tzv. prekograničnu strukturu, tj. predstavlja globalni problem, to se javlja potreba da vlade država, privreda i javnost imaju pouzdane i usaglašene podatke o stanju životne sredine. I upravo razvijeni indikatori životne sredine na nacionalnom nivou, a koji su usklađeni sa međunarodno prihvaćenim indikatorima radi poređenja, pružaju informacije o stanju ili opisuju pojave u oblasti životne sredine.

Pored nacionalnih indikatora, mnoge međunarodne organizacije u okviru svojih programa za postizanje ciljeva održivog razvoja, razvile su idikatore životne sredine. Ekonomska komisija za Evropu (UNECE - *United Nations Economic Commission for Europe*) koja broji 56 država članica, među kojima je i Srbija, za praćenje stanja životne sredine razvila je 10 glavnih indikatora [6]:

1. Zagađenje vazduha i oštećenje ozonskog omotača;
2. Klimatske promene;
3. Vode;
4. Biodiverzitet;
5. Zemljište;
6. Poljoprivreda;
7. Energija;
8. Saobraćaj;
9. Otpad;
10. Finansiranje zaštite životne sredine.

Organizacija za ekonomsku saradnju i razvoj (OECD - *Organisation for Economic Co-operation and Development*) je za bolje razumevanje, deljenje i poređenje podataka o stanju i kvalitetu životne sredine na međunarodnom nivou, razvila četiri glavna indikatora životne sredine [7]:

1. Vazduh i klimatske promene;
2. Šumski resursi;
3. Komunalni otpad;
4. Vode.

Pored gore navedenih i mnoge druge međunarodne organizacije poput Ujedinjenih nacija (UN - *United Nations*), Svetske banke (WB - *The World Bank*), Evropske agencije za zaštitu životne sredine (EEA - *The European Environment Agency*), koriste indikatore zaštite životne sredine da odgovore na ključna pitanja koja se tiču politike upravljanja i zaštite životne sredine, projektovanja okvira budućih politika i ciljeva, monitoringa i evaluacije strategije održivog razvoja.

U okviru Nacionalne strategije održivog razvoja republike Srbije, predviđeno je godišnje izveštavanje o napretku u pravcu ostvarivanja ciljeva strategije, a koje se zasniva na indikatorima Milenijumskih ciljeva razvoja. U tom smislu je, radi efikasnijeg izveštavanja o stanju životne sredine i usaglašavanja sa međunarodnom metodologijom, u Srbiji razvijena Nacionalna lista indikatora zaštite životne sredine koja je podeljena prema kategorijama u tematske celine [8]:

- Vazduh i klimatske promene;
- Vode;
- Zemljište;
- Otpad;
- Buka;
- Nejonizujuće zračenje;
- Šumarstvo, lov i ribolov;
- Održivo korišćenje prirodnih resursa;

- Privredni i društveni potencijali i aktivnosti od značaja za životnu sredinu;
- Međunarodna i nacionalna zakonska regulativa, kao i mere (strategije, planovi, programi, sporazumi), izveštaji i ostala dokumenta i aktivnosti iz oblasti zaštite životne sredine;
- Subjekti sistema zaštite životne sredine.

Indikatori su razvijeni primenom sistema Pokretački faktori – Pritisци - Stanje - Uticaji - Reakcije društva (*D - Driving Forces, P - Pressures, S - State, I - Impact, R - Response - DPSIR framework*) koji je razvijen od strane EEA [9].

Kao značajan segment u okviru različitih lista indikatora kvaliteta životne sredine, jeste praćenje stanja kvaliteta vazduha. Indikatori koji se odnose na vazduh i klimatske promene a koji su tema ove doktorske disertacije, su opisani u sledećem poglavlju.

2.2. Prikaz odabranih indikatora iz kategorije kvaliteta vazduha

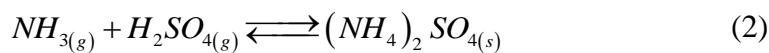
Osnovni cilj indikatora koji prate stanje u atmosferi je da daje meru postojećih i očekivanih pritisaka na životnu sredinu u pogledu emisije štetnih materija u vazduh, kao i napredak u odnosu na zadati cilj, ukoliko postoji. Značaj indikatora nije samo u proceni pritisaka na životnu sredinu od zagađenja vazduha, već i u proceni i identifikovanju stacionarnih i mobilnih izvora emisije koji su uslovljeni određenim ekonomskim aktivnostima. U tom smislu se, na godišnjem nivou, podnose izveštaji o emisijama zagađujućih materija (emisija iz pojedinih izvora i ukupna emisija) koje imaju negativan efekat kako na zdravlje ljudi tako i na ekosisteme. U zagađujuće materije čija se emisija prati kroz indikator koji se odnosi na vazduh i klimatske promene spadaju: sumpor-dioksid (SO₂), oksidi azota (NO_x), nemetanska isparljiva organska jedinjenja (NMVOCs - *non-methane volatile organic compounds*), amonijak (NH₃), ugljen-monoksid (CO), suspendovane čestice (PM₁₀, PM_{2.5}), dugotrajne organske zagađujuće materije (POPs - *Persistent Organic Pollutants*), teški metali (uglavnom Pb, Cd, Hg, Ni i As), gasovi sa efektom staklene bašte (CO₂, N₂O, CH₄, F- gasovi HFCs, PFCs i SF₆) kao i neke specifične zagađujuće supstance.

U nastavku rada su dati opisi indikatora iz kategorije vazduha čije su emisije na nacionalnom nivou predmet modelovanja u ovoj doktorskoj disertaciji.

2.2.1. Emisija amonijaka (NH₃)

Amonijak je bezbojan gas, karakterističnog nadražujućeg mirisa koji pri višim koncentracijama ima negativan efekat na ljudsko zdravlje. Spada u najrasprostranjenije gasove baznog karaktera u atmosferi, sa širokim spektrom uticaja, uključujući fotohemijsko zagađenje, smanjenu vidljivost, uticaj na biodiverzitet kao i oštećenje ozonskog omotača [10,11].

Amonijak gas ima relativno kratak životni vek u atmosferi, od nekoliko sati do nekoliko dana (10-100 km), dok se amonijum jon, kao aerosol, može transporovati na mnogo većim udaljenostima (100-1000 km) i životni vek mu je 1-15 dana [12]. Amonijak gas se iz atmosfere uklanja na osnovu aviniteta prema vodi i njegovog baznog karaktera, što se ogleda u formiranju i neutralizaciji nitratnih i sulfatnih aerosola u zagađenoj atmosferi. Amonijak reaguje sa ovim kiselim aerosolima uz formiranje čestica amonijumovih soli, kao što je prikazano u reakcijama (1) i (2) [13]:



Amonijak dakle, spada u zagađujuće materije koje su prekursori za formiranje "sekundarnih čestica" putem kompleksnih reakcija, koje su značajni deo ukupnog zagađenja suspendovanim česticama. I upravo soli sulfata ili nitrata predstavljaju značajan deo mase finih suspendovanih čestica PM2.5.

Rezultati nedavnih istraživanja sprovedenih u Sjedinjenim Američkim Državama (SAD) kod dugotrajne izloženosti PM2.5 česticama su pokazali da porast koncentracije finih čestica za 10 µg/m³ za posledicu ima za 6% povećanje svih vrsta zdravstvenih rizika, 9% povećan rizik od kardiovaskularnih problema i problema sa disajnim organima i 14% povećan rizik od raka pluća [14]. U skladu sa novim naučnim saznanjima koja se odnose na zagađenje vazduha

suspendovanim česticama, države razvijaju ili redefinišu svoje zakonske regulative u skladu sa postavljenim standardima, koje se odnose i na granične vrednosti respirabilnih čestica. Radi sprovođenja zadatih planova, unapređeni disperzioni modeli za predviđanje koncentracija atmosferskih zagađivača su korisna podrška u odlučivanju u sistemima upravljanja kvalitetom vazduha. Ovi modeli koriste meteorološke podatke i podatke o emisijama određenih zagađujućih materija (zasnovanih na inventaru emisija), kao ulazne parametre [15].

S obzirom na to da amonijum-jon predstavlja značajan sastojak suspendovanih čestica, veoma je bitno razvijanje što preciznijih modela za predviđanje emisije amonijaka čiji se podaci koriste kao ulazni parametar u modelima za predviđanje koncentracije pomenutih zagađujućih materija.

Amonijak se u atmosferu emituje iz različitih izvora: poljoprivrede, drumskog saobraćaja, industrijskih procesa, itd. i rezultati nedavnih istraživanja ukazuju na ubranu proizvodnju reaktivnog azota na globalnom nivou, nastalog kao posledica ljudske aktivnosti [16]. Pokazano je da su jedan od najznačajnijih izvora emisije amonijaka u atmosferu aktivnosti vezane za poljoprivredu [17]. Procene pokazuju da se 83% azota korišćenog u obliku đubriva oslobađa u životnu sredinu [18], pri čemu se u SAD od ukupnog azota iz đubriva, 30% oslobađa u vidu emisije amonijaka [19]. Poljoprivredni sektor je takođe najveći izvor emisije amonijaka i u Evropi. Preko 90% od ukupne emisije NH_3 potiče iz ovog sektora [20].

2.2.2. Emisija nemetanskih isparljivih organskih jedinjenja (NMVOCs)

Praćenje emisije nemetanskih isparljivih organskih jedinjenja (NMVOCs), postalo je veoma značajan segment u okviru zakonodavnih mera od strane raznih regulatornih tela, u cilju redukcije emisije ovih zagađujućih materija budući da imaju veoma značajnu ulogu u hemiji atmosfere, od lokalnog do globalnog nivoa.

NMVOCs čine grupu od nekoliko stotina organskih jedinjenja (razni alifatični i aromatični ugljovodonici i njihova suspsituisana jedinjenja) koja se razlikuju po svom hemijskom sastavu, ali pokazuju slično ponašanje u atmosferi. NMVOCs su veoma značajni polutanti sa

dva aspekta: prvo, zbog njihovog štetnog uticaja na ljudsko zdravlje, jer su se pokazali kao veoma toksični, mutageni i kancerogeni; drugo, oni predstavljaju jedan od dominantnih prekursora iz grupe onih koji su uključeni u nastajanje troposferskog ozona. Jedan broj NMVOCs učestvuje u složenim fotohemijskim reakcijama stvarajući fotohemijski smog, organske aerosole, organske kiseline u vazduhu i oni su takođe veoma štetni za ljudsko zdravlje [21]. Stoga, SAD i zemlje EU imaju predložene liste određenih NMVOCs, koji deluju kao prekursori za formiranje fotohemijskog smoga, koji su označeni kao polutanti čija je kontrola prioritetna [22].

Izvori emisija NMVOCs u atmosferu su raznovrsni i obuhvataju: prirodne, kao što je vegetacija, i antropogene izvore u koje spadaju industrija, saobraćaj i poljoprivreda [23]. Ukupna emisija NMVOCs na nacionalnom nivou i odnos između pojedinih izvora emisije varira od zemlje do zemlje i zavisi pre svega od njihovog stepena razvijenosti [24].

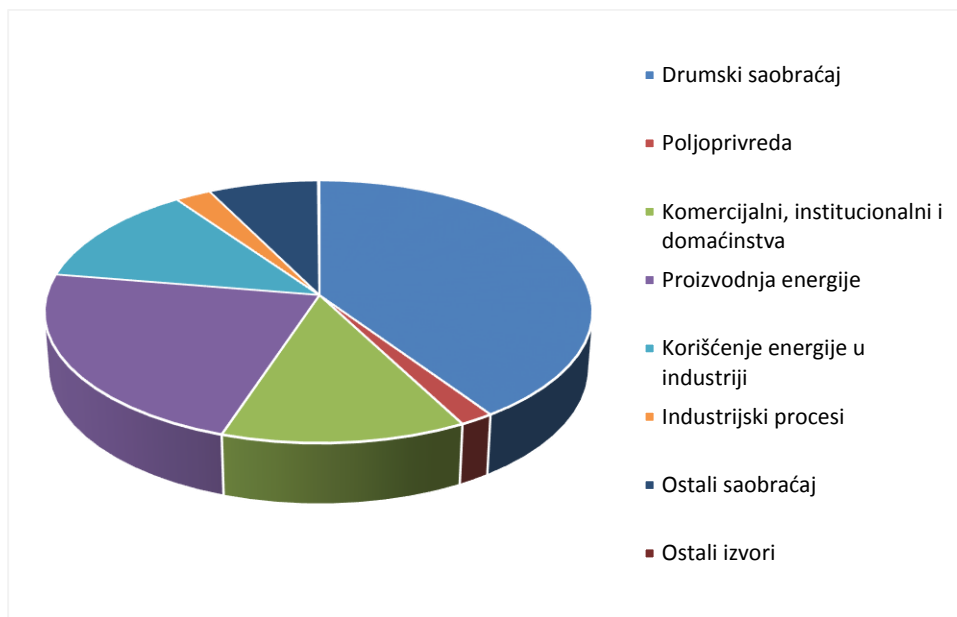
Na globalnom nivou, najveći emiteri NMVOCs su upravo industrijski najrazvijenije države, među kojima Kina zauzima prvo mesto. Glavni izvori emisije u Kini su procesi sagorevanja goriva, upotreba rastvarača, petrohemijska industrija i drumski saobraćaj [25]. Kako je emisija ovih zagađujućih materija u Kini velika, to je njeno praćenje značajno u smislu pomoći da se prati formiranje troposferskog ozona u istočnoj Aziji [26]. Takođe, u nekoliko istraživanja je pokazano da su emisije NMVOCs jedan od najvažnijih prekursora za formiranje površinskog ozona i organskih aerosola, čija je koncentracija značajno porasla poslednjih godina, kako u urbanim tako i u ruralnim regionima Kine [27,28].

2.2.3. Emisija azotnih oksida (NO_x)

Dve veoma značajne gasovite zagađujuće materije atmosfere su azotova jedinjenja i to bezbojni azot-oksidi (NO) i crveno-braon obojeni azot-dioksidi (NO_2), koji se zbirno označavaju kao azotovi oksidi (NO_x). Generalno, najveći antropogeni izvori emisija NO_x su procesi sagorevanja fosilnih goriva, kako iz stacionarnih izvora, tako i iz sektora saobraćaja. Pri tome, smeša azotnih oksida nastaje oksidacijom atmosferskog azota i azota iz goriva prilikom procesa sagorevanja goriva na visokim temperaturama (iznad $1500\text{ }^\circ\text{C}$), kroz seriju kompleksnih hemijskih reakcija.

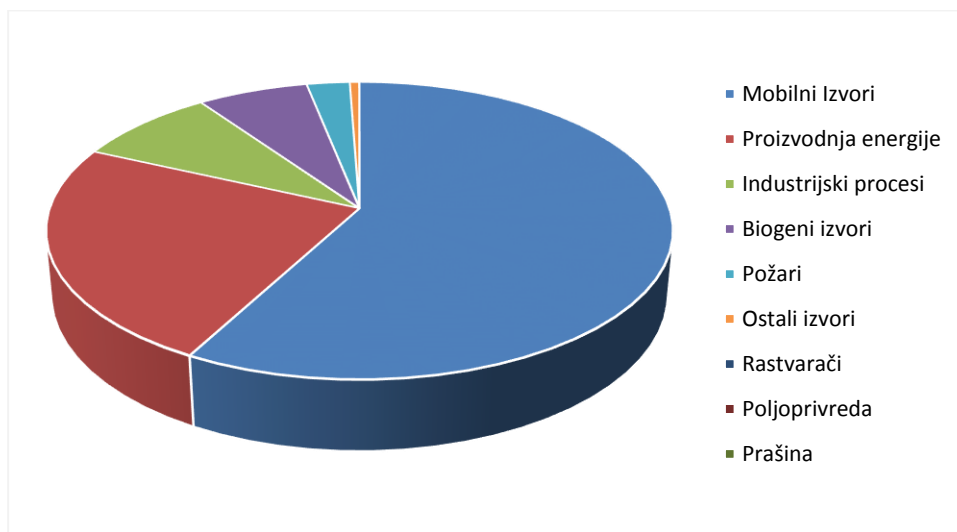
Emisija NO_x ima višestruke efekte na životnu sredinu, uključujući stvaranje kiselih kiša (putem formiranja azotne kiseline), finih suspendovanih čestica $\text{PM}_{2.5}$ (putem formiranja sekundarnih čestica, kao što je amonijum-nitrat), nastajanje fotohemijskog smoga (putem fotohemijskih reakcija sa organskim materijama) i povećanje koncentracije troposferskog ozona (putem reakcija sa ugljovodicima i ugljen-monoksidom), što za posledicu ima štetan uticaj na zdravlje ljudi, eutrofikaciju vodenih ekosistema, kao i na poljoprivredu [29–32]. Sa aspekta uticaja na kvalitet vazduha i ljudsko zdravlje NO_2 je mnogo značajniji, jer je nivo toksičnosti NO_2 pet puta veći u odnosu na NO čime direktno utiče na ljudsko zdravlje, što se ogleda u povećanju broja obolelih od plućnih bolesti i smanjenoj otpornosti od respiratornih infekcija [33].

Oksidi azota se u atmosferu emituju iz hiljade antropogenih izvora, kao i značajnog broja prirodnih izvora. Doprinos pojedinih sektora emisiji NO_x se razlikuje među državama. U Evropi je dominantni sektor NO_x emisije, sa četrdeset procenata od ukupne emisije, drumski saobraćaj (Slika 1.) [34].



Slika 1. Udeo pojedinih sektora u emisiji NO_x u EU

U Sjedinjenim Američkim Državama, takođe, dominantni izvor emisije azotovih oksida su procesi sagorevanja u mobilnim izvorima (saobraćaj) (Slika 2.), kao i proizvodnja električne energije, čiji je doprinos oko 25% od ukupne antropogene emisije azotovih oksida, budući da se oko 90% od ukupne potrošnje uglja koristi upravo za proizvodnju električne energije [35].



Slika 2. Udeo pojedinih sektora u emisiji NO_x u SAD

Najveći emiter azotovih oksida u Aziji je Kina, sa udelom 41-47% od ukupne emisije u Aziji i udelom od oko 10.1% u ukupnoj emisiji u svetu [36,37]. Dominantni sektor, izvor emisije NO_x u Kini, sa udelom od oko 70%, je proces sagorevanja uglja, koji trenutno predstavlja osnovni energent za proizvodnju energije [38]. Procenjuje se da će se takav trend upotrebe uglja nastaviti i u budućnosti.

Poslednjih godina uvedene su određene kontrolne mere i strategije koje zahtevaju primenu naprednih tehnologija radi redukcije NO_x emisija. Selektivna katalitička redukcija (SCR - *Selective Catalytic Reduction*) je jedna od tehnologija koja se primenjuje radi smanjenja NO_x emisija u procesima sagorevanja fosilnih goriva, kako u mobilnim tako i u stacionarnim izvorima. Selektivnom katalitičkom redukcijom može se redukovati emisija NO_x za 70-90% [39].

2.2.4. Emisija metana (CH₄)

U okviru indikatora životne sredine koji se odnose na klimatske promene, predviđeno je i praćenje emisije gasova sa efektom staklene bašte (GHGs - *Greenhouse Gases*). Poslednjih decenija, emisija GHG i njihov uticaj na životnu sredinu predmet je brojnih istraživanja i primene međunarodnih sporazuma radi redukcije emisije. Pokazalo se da povišene koncentracije ovih gasova imaju višestruko negativno dejstvo. Globalno zagrevanje, promena klime iz koje proističu negativni uticaji na ekosisteme, poljoprivredu i druge društveno-ekonomske aktivnosti, su samo neke od posledica povećane emisije GHG u atmosferu.

Doprinos pojedinih gasova staklene bašte globalnom zagrevanju zavisi od njihovog potencijala da apsorbuju zračenje sa zemlje i njihovog životnog veka u vazduhu. Od svih GHGs, najveći pritisak na životnu sredinu koji vodi ka promeni klime, potiče od CO₂, CH₄ i N₂O. Da bi se procenio uticaj svakog od GHGs na globalno zagrevanje, uveden je pojam potencijala globalnog zagrevanja (GWP - *Global Warming Potential*) izražen preko CO₂ ekvivalenta. GWP predstavlja odnos zagrevanja koje vrši određeni gas u odnosu na zagrevanje od slične mase ugljen-dioksida (za koji je usvojeno da je GWP jednak 1). GWP se obračunava u određenom intervalu vremena, najčešće za period od 100 godina [6].

Pored ugljen-dioksida (CO₂), metan (CH₄) je sledeći po značaju gas koji doprinosi efektu staklene bašte sa nekih 20% u odnosu na ostale gasove. Životni vek metana u atmosferi je mnogo kraći u odnosu na ugljen-dioksid, međutim njegov potencijal za apsorpciju zračenja je veći u odnosu na ugljen-dioksid. Smatra se da je uticaj metana na klimatske promene oko 25 puta veći u odnosu na ugljen-dioksid u periodu od 100 godina [40]. Pored toga, metan spada u gasove prekursore za formiranje sekundarne zagađujuće materije, ozona, koji takođe, ima vrlo negativne efekte na životnu sredinu. I pored toga, iako je uticaj metana, kako na globalno zagrevanje, tako i na stvaranje ozona, značajan, poslednjih 250 godina njegova koncentracija u atmosferi se povećala za 152% [41].

Na globalnom nivou, više od 60% od ukupne emisije metana je posledica antropogenih aktivnosti [42]. U najznačajnije izvore emisije spadaju:

- Transport i eksploatacija prirodnog gasa;
- Deponije;
- Eksploatacija uglja;
- Tretman otpadnih voda;
- Uzgajanje pirinča;
- Petrohemijska industrija;
- Procesi sagorevanja (mobilni i stacionarni izvori);
- Industrijski procesi;
- Fermentacija stajskog đubriva;
- Procesi varenja preživara, itd.

Kao jedan od značajnijih antropogenih izvora emisije metana su i deponije. Procenjuje se da emisija sa deponija iznosi oko 30% od ukupne antropogene emisije metana u Evropi, u SAD iznosi 34%, i 10% od ukupne antropogene emisije u svetu [43]. U skladu sa tim, donose se određene zakonodavne mere kako bi se smanjila produkcija metana iz tog sektora. Jedna od tih mera je EU direktiva o deponijama otpada [44] koja se odnosi na države EU. Predloženom direktivom se zahteva od država članica redukovanje deponovanih količina biorazgradivog komunalnog otpada od 65% do 2020. u odnosu na nivo iz 1995. godine.

2.2.5. Međunarodni zakonodavno-pravni mehanizmi za smanjenje zagađenja vazduha

Kao što je već istaknuto, emisija zagađujućih materija u vazduh i izloženost stanovništva tom zagađenju može imati značajne negativne efekte na zdravlje ljudi u vidu akutnih i hroničnih bolesti pa sve do smrti. Organizacija za ekonomsku saradnju i razvoj (OECD - *Organisation for Economic Co-operation and Development*) procenjuje da će se kvalitet vazduha, ukoliko se ne implementiraju nove razvojne politike koje bi kontrolisale zagađenje atmosfere, i dalje globalno pogoršavati. OECD predviđa da će do 2050. godine, zagađenje vazduha biti najveći uzrok smrti povezanih sa uslovima u životnoj sredini [45].

Prekogračni, ili globalni karakter zagađenja vazduha uslovalo je potrebu za međunarodnim zakonodavnim mehanizmima, kojima bi se zajednički uredila ova oblast životne sredine. U tom smislu, zaštita vazduha uslovala je dinamičan razvoj zakonodavstva u mnogim zemljama, koje je usko povezano sa Evropskim i svetskim konvencijama. Međunarodne obaveze preuzete iz tih konvencija, za zemlje potpisnice, propisuju redovno izveštavanje o stanju zagađivanja vazduha i emisiji određenih zagađujućih materija. U nastavku rada će biti predstavljeni neki od najznačajnijih međunarodnih ugovora iz oblasti koja se odnosi na kontrolu zagađivanja vazduha.

Radi poboljšanja kvaliteta vazduha od lokalnog do regionalnog nivoa, 1979. godine u organizaciji Evropske ekonomske komisije za Evropu (UNECE - *United Nations Economic Commission for Europe*), u Ženevi je usvojena Konvencija o prekograničnom zagađenju vazduha na velikim udaljenostima (CLRTAP - *Convention on Long Range Transboundary Air Pollution*). Od usvajanja Konvencije do danas, ona je proširena sa osam protokola koji identifikuju specifične mere koje je trebalo preduzeti od strane država potpisnica radi smanjenja emisije zagađujućih materija. Konvenciju su potpisale 51 država, članica UNECE [46]. Protokoli ove konvencije su prikazani u Tabeli 1. [47,48].

Tabela 1. Protokoli LRTAP Konvencije

Protokol LRTAP Konvencije	Godina stupanja na snagu
Protokol o zajedničkom praćenju i proceni prenosa zagađujućih materija na velikim udaljenostima u Evropi (EMEP) (1984.)	1988.
Protokol o 30% smanjenju emisije oksida sumpora i njihov prekogranični prenos (1985.)	1987.
Protokol o kontroli emisija azotovih oksida i njihovom prekograničnom prenosu (1988.)	1991.
Protokol o kontroli emisija isparljivih organskih jedinjenja (VOC) i njihovom prekograničnom prenosu (1991.)	1997.
Protokol o daljem smanjenju emisija sumpora (1994.)	1998.
Protokol o teškim metalima i njegova 2012. izmenjena verzija (1998.)	2003.
Protokol o perzistentnim organskim polutantima (POPs) i njegova 2009. izmenjena verzija (1998.)	2003
Protokol o suzbijanju acidifikacije, eutrofikacije i prizemnog ozona i njegova 2012. izmenjena verzija (1999.)	2005.

Prema EMEP protokolu (EMEP - *European Monitoring and Evaluation Programme*), koji daje naučnu podršku LRTAP konvenciji, propisuje se redovno podnošenje podataka o emisijama sledećih zagađujućih materija: SO₂, NO_x, CO₂, CH₄, NMVOC, CO, NH₃, teških metala i POPs. U skladu sa tim, za navedene zagađujuće materije, svaka od potpisnica konvencije, među kojima je i Srbija, dostavlja izveštaj o ukupnoj nacionalnoj emisiji.

Okvirna konvencija Ujedinjenih nacija o promeni klime (UNFCCC - *United Nations Framework Convention on Climate Change*) je jedan od najznačajnijih globalnih napora usmerenih ka rešavanju najvećeg izazova današnjice, promene klime i globalnog zagrevanja. Osnovni cilj ove konvencije je da se međudržavnim naporima i zakonskim instrumentima usklađuju nastojanja svih članica, kako bi se sprečio "opasan" uticaj antropogenog delovanja na klimatski sistem. Konvencija je stupila na snagu 1994. godine.

Međunarodni sporazum kojim je proširena UNFCCC konvencija je Kjoto protokol, koji je usvojen 1997. godine. U prvom obavezujućem periodu Kjoto protokola, od 2008. do 2012. godine, industrijski razvijene države su bile u obavezi da smanje emisije GHG za 5% u odnosu na vrednosti iz 1990. godine. Drugi obavezujući period ovog protokola počeo je od

2013. godine i traje do 2020. godine. U drugom obavezujućem periodu, države potpisnice su se obavezale da smanje emisiju GHG za najmanje 18% u odnosu na nivo iz 1990. godine. Okvirna konvencija Ujedinjenih nacija trenutno ima 195 članica, dok Kjoto protokol broji 192 članice, od kojih su 83 potpisnice protokola [49]. Srbija je članica UNFCCC od 2001. godine i spada u red onih država koje su ratifikovale Kjoto protokol. U okviru konvencije, države članice su dužne da podnesu godišnje izveštaje o emisiji GHG (od bazne 1990. godine pa do danas), kao i izveštaje o svojim razvojnim politikama usmerenim ka redukciji emisije GHG.

Za uspešno sprovođenje preuzetih međunarodnih obaveza i ostvarivanje zadatih ciljeva u pogledu redukcije koncentracije zagađujućih materija u atmosferu, kako za industrijski razvijene, tako i države u razvoju, od velikog je značaja postojanje što preciznijih modela za procenu i predviđanje nacionalnih emisija određenih zagađujućih materija. Dosadašnji modeli, o kojima će biti više reči u poglavlju 2.3, kojima se procenjuje i predviđa emisija, zasnovani su na inventarima emisija i razvoju relevantnih scenarija za procenu budućih emisija. Razvoj računarskih nauka pruža mogućnost razvijanja novih, savremenih modela. U ovom radu, to se pre svega odnosi na razvijanje modela za predviđanje emisije primenom veštačkih neuronskih mreža, tehniku modelovanja koja je pokazala dobre performanse kako u drugim oblastima, tako i u modelovanju emisija [50–52].

2.3. Postojeći modeli za procenu emisija

Direktna merenja emisije zagađujućih materija u vazduh karakteriše i najveća tačnost [53]. Međutim, taj pristup je finansijski zahtevan, pogotovu za države u razvoju, budući da zahteva odgovarajuću mernu opremu. Postojeći modeli za procenu nacionalnih, regionalnih i globalnih emisija zagađujućih materija u atmosferu, zasnovani su na inventarima emisija. Inventari emisija predstavljaju i osnovne izvore podataka koje koriste disperzioni modeli atmosferskih zagađujućih materija, koji su podrška u odlučivanju u sistemima upravljanja kvalitetom vazduha. Generalno, procena emisije u inventarima, zasnovana je na emisionim faktorima za određenu zagađujuću materiju i primenjenu tehnologiju izvora emisije. Da bi procenjenjena emisija bila što pouzdanija, potrebno je odrediti emisione faktore za svaki

emisioni proces, odvojeno na nacionalnom nivou. Međutim, kako je to proces koji iziskuje određene troškove i vreme, to se radi proračuna emisije najčešće koriste postojeći emisioni faktori iz literature, što za posledicu može imati nedovoljno tačne podatke o emisiji. Emisioni faktori određeni od strane US EPA su najčešće korišćeni [54].

2.3.1. EDGAR v4 model

Globalni EDGAR model (EDGAR - *Emission Database for Global Atmospheric Research*) (1992-2004.) prvobitno je razvijen od strane Holandskog nacionalnog instituta za javno zdravlje i životnu sredinu (RIVM - *Rijksinstituut voor Volksgezondheid en Milieu*) i Holandske organizacije za primenjena naučna istraživanja (TNO - *Nederlandse Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek*). Nakon 2004. godine, razvijene su nove dopunjene verzije EDGAR modela, kao rezultat rada Holandske agencije za zaštitu životne sredine (PBL - *Planbureau voor de Leefomgeving*) i zajedničkog istraživačkog centra EU (JRC - *Joint Research Centre*) [55,56].

EDGAR modelom se procenjuje emisija: direktnih gasova sa efektom staklene bašte, gasova prekursora ozona, acidifikujućih gasova, primarnih čestica i emisija supstanci koje oštećuju ozonski omotač. Proračun emisije je zasnovan na emisionim faktorima i drugim specifičnim parametrima za svaku državu i vrši se prema jednačini (3) [57]:

$$EM_C(x, y) = \sum_{i,j,k} \left[AD_{c,i}(y) * TECH_{c,i,j}(y) * EOP_{c,i,j,k}(y) * EF_{c,i,j}(y, x) * (1 - RED_{c,j,i,k}(y, x)) \right] \quad (3)$$

gde su:

c, x, y, i, j, k - država, zagađujuća materija, sektor emisije, vrsta tehnologije, procenat smanjenja emisije; EM - godišnja emisija određene zagađujuće materije; AD - stopa aktivnosti u određenom sektoru izvoru emisije; $TECH$ - vrsta tehnologije za svaki sektor; EOP - primenjena mera za smanjenje emisije za svaki sektor izvora emisije; EF - emisioni faktor za svaki sektor i tehnologiju; RED - efikasnost smanjenja emisije u odnosu na primenjenu meru.

Koristeći IPCC sektorski pristup (IPCC - *Intergovernmental Panel on Climate Change*), EDGAR metodologija razlikuje sedam glavnih sektora, izvora emisije, koji su dalje podeljeni na veliki broj, preko 50, podsektora. Glavni sektori iz kojih se procenjuje emisija su: energetska sektor (proces sagorevanja goriva i fugalna emisija); industrijski procesi; korišćenje rastvarača i drugih proizvoda; poljoprivreda; korišćenje zemljišta i promena namene zemljišta i šumarstvo; otpad i drugi antropogeni izvori [58–60]. EDGAR modelom je procenjena godišnja emisija iz 25 regiona sveta i država koje pripadaju tim regionima, po sektorima i ukupna emisija.

Ulazni podaci za procenu emisije, pre svega, stopa aktivnosti i emisioni faktori, su uglavnom preuzeti iz međunarodnih statističkih baza podataka, međunarodnih naučnih publikacija, industrijskih i energetska izveštaja. I upravo, primenjeni podaci o stopi aktivnosti, a posebno vrednosti emisija faktora koji se koriste za procenu emisije ovim modelom, a koji nisu specifični za određenu državu, primenjenu tehnologiju i sektor emisije, su osnovni faktori nesigurnosti EDGAR modela [61]. Ekspertska procena ovih parametara pokazala je da se u zavisnosti od sektora, nesigurnost kreće u opsegu od 10% pa do veoma velikih vrednosti nesigurnosti >100% [62]. Takođe, kao jedan od uzroka nesigurnosti procenjene emisije je i nedostatak statističkih podataka o stopama aktivnosti za pojedine države, pogotovu za one u razvoju.

2.3.2. EMEP/CORINAIR model

EMEP/CORINAIR (EMEP - *European Monitoring and Evaluation Programme* / CORINAIR - *CORe Inventory of AIR emissions*) metodologijom vrši se proračun emisija zagađujućih materija u cilju korišćenja podataka za donošenje odluka od međunarodnog interesa u pogledu zaštite kvaliteta vazduha. Ovom metodologijom se pomaže u pripremi nacionalnih inventara emisija preko EMEP/EEA priručnika po pitanju usklađenosti sa propisima zahtevanim od strane različitih međunarodnih organizacija, kao što su npr. EMEP, UNECE, UNFCCC, Eurostat i dr. [47]. Pomenuti inventar procenjuje emisije u okviru pan-Evropskog domena sa prostornom rezolucijom 50x50 km na godišnjem nivou [63].

Metodologijom se procenjuje emisija sledećih zagađujućih materija: NO_x, SO₂, NH₃, NMVOC, CO, CH₄, PM_{2.5}, PM₁₀, TSP, TSP, 10POPs (HCH, PCP, HCB, TCM, TRI, PER, TCB, TCE, dioksini, furani i PAH), 9 teških metala (As, Cd, Cr, Cu, Hg, Ni, Pb, Se i Zn). Ulazni parametri za procenu emisija ovim modelom, svrstani su prema SNAP nomenklaturi (SNAP - *Selected Nomenclature for Air Pollution*). Prema ovoj nomenklaturi, sektori izvori emisija zagađujućih materija razvrstani su na tri nivoa. Prvi nivo čine jedanaest glavnih sektora, koji su dalje detaljnije podeljeni na oko 50 kategorija izvora emisije u SNAP nivo 2, koje su dalje podeljene na oko 350 različitih kategorija u SNAP nivo 3, koji je najdetaljniji. SNAP nivo 3 može dalje da se razvrsta ako su podaci o aktivnostima i emisionim faktorima dostupni [64].

Jedanaest glavnih sektora izvora emisije, prirodnih i antropogenih, koji čine ulazne parametre za procenu emisije ovim modelom, prikazani su u Tabeli 2.

Tabela 2. Jedanaest glavnih sektora u CORINAIR metodologiji [65]

SNAP kod	Opis sektora
1	Sagorevanje goriva u termoenergetskim sektorima i postrojenjima za transformaciju energije
2	Sagorevanje u ne-industrijskim ložištima
3	Sagorevanje u industriji
4	Proizvodni procesi
5	Dobijanje i distribucija fosilnih goriva i geotermalne energije
6	Korišćenje rastvarača i ostalih proizvoda
7	Drumski saobraćaj
8	Ostali mobilni izvori i mašine
9	Obrada i odlaganje otpada
10	Poljoprivreda
11	Prirodni izvori

Proračun emisija i kod ovog modela se vrši pristupom zasnovanim na emisionim faktorima i stopama aktivnosti, prema jednačini (4).

$$E_{\text{zagađujuće materije}} = \sum AR \times EF_{\text{tehnologija, zagađujuća materija}} \quad (4)$$

gde su :

$E_{\text{zagađujuće materije}}$ - emisija određene zagađujuće materije; AR - stopa aktivnosti. Podaci se obično dobijaju iz privredne statistike, uključujući energetska statistiku, stope privrednog rasta, podaci o stanovništvu, itd. [66]; $EF_{\text{tehnologija, zagađujuća materija}}$ - emisioni faktor za određenu zagađujuću materiju i primenjenu tehnologiju.

EMEP/CORINAIR metodologija u okviru IPCC smernica definiše, u zavisnosti od nivoa detaljnosti ulaznih parametara, različite nivoe preciznosti procene emisije, od 1 do 3, a od izbora metoda zavisi nivo preciznosti procenjene emisije [60].

”Nivo 1” metoda koristi lako dostupne statističke podatke o intenzitetu procesa i podrazumevanih emisionih faktora. Ova metoda je najjednostavnija metoda, ima najviši stepen neizvesnosti.

”Nivo 2” koristi nešto specifičnije emisione faktore dobijene na osnovu poznavanja vrste procesa i specifičnim uslovima procesa koji su primenjeni u datoj zemlji čiji se inventar emisija priprema.

”Nivo 3” Procena emisije se vrši na sličan način kao i kod Nivo 2 metode (tj. mera aktivnosti x emisioni faktor) ali sa većim razdvajanjem mere aktivnosti i emisionih faktora.

Pored proračuna emisije, ovim modelom se vrši i projekcija budućih emisija. Predviđanje emisija je, kao i procena emisija, funkcija stope aktivnosti u kombinaciji sa emisionim faktorima. Ovi parametri se za predviđanje emisije moraju procenjivati ili modelovati korišćenjem različitih mogućih kretanja u privredi, ekonomiji, kretanja u tehnologijama, čime se projektuju različiti scenariji. EMEP/CORINAIR metodologija definiše tri grupe scenarija koje su detaljno opisane u odgovarajućoj literaturi [47,60,66].

Kao i kod EDGAR modela i EMEP/CORINAIR model pokazuje određeni nivo nesigurnosti podataka emisije. U IPCC smernicama [67] i EMEP/EEA priručniku, dat je detaljan opis uzroka i kvantifikacije nesigurnosti u modelima za procenu emisije zasnovanim na inventarima emisija. Uzimajući u obzir sve uticaje na nesigurnost u proceni emisija, pokazalo se da najveći doprinos u nesigurnosti imaju emisioni faktori, dok nesigurnost u podacima o aktivnosti dominira u predviđanjima emisija [68].

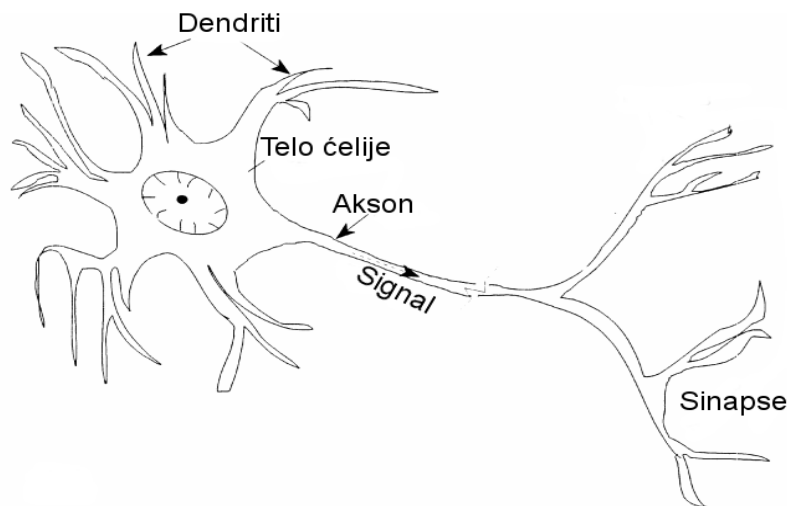
Kao što je već rečeno, podaci o emisijama zasnovani na inventarima su jedna od osnovnih komponenti u modelovanju kvaliteta vazduha i veoma je značajno da ti podaci budu što pouzdaniji. Međutim, brojna istraživanja su pokazala da ulazni podaci emisija značajno doprinose nesigurnosti modela kvaliteta vazduha [69–73]. U skladu sa tim, radi procene tačnosti, pouzdanosti i praktične vrednosti modela, podnošenje nacionalnih izveštaja o emisijama zagađujućih materija u vazduh, prema napred pomenutim smernicama, podrazumeva i obezbeđivanje izveštaja o verifikaciji modela. Poređenje sa drugim nezavisnim modelima (ukoliko su dostupni) je jedna od brzih opcija za verifikaciju postojećih modela.

Pored EDGAR i EMEP/CORINAIR modela, postoje i brojni drugi modeli za procenu emisija, koji su takođe zasnovani na emisionim faktorima i stopama aktivnosti, tj. inventarskom pristupu. Generalno, razlikuju se prema vremenskoj i prostornoj rezoluciji, izvorima emisije koje uključuju u procenu, kao i prema zagađujućim materijama čije se emisije procenjuju. Pored navedenih, jedan od značajnijih modela kojima se vrši procena i predviđanje emisije velikog broja zagađujućih materija u vazduh na globalnom nivou (48 država Evrope, Azija i Anex I države UNFCCC konvencije), je GAINS model (*Greenhouse Gas and Air Pollution Interactions and Synergies*). Detaljnije informacije o ovom modelu mogu se naći u naučnoj literaturi kao i na sajtu IIASA instituta [74].

3. Veštačke neuronske mreže

3.1. Biološke i veštačke neuronske mreže

Jedan od osnovnih elemenata nervnog sistema, koji ima ulogu prenosioca i skladištenja informacija je neuron. Pojednostavljeni izgled biološkog neurona je prikazan na Slici 3. Osnovne funkcionalne jedinice biološkog neurona su: dendriti, telo ćelije i akson. Dendriti imaju funkciju prihvatanja signala, u vidu elektro-hemijskih impulsa, iz drugih neurona, koje dalje prosleđuju telu ćelije. Nakon toga se aktivnost neurona nastavlja preko aksona, koji prenosi signal na dendrite drugih neurona. Veza između neurona uspostavljena je preko sinapsi, koje imaju značajnu ulogu u funkcionisanju celokupnog sistema [75]. Procenjuje se da svaki neuron može primiti stimulans od oko 10.000 drugih neurona. Grupa neurona je organizovana u podsisteme i integracija ovih podsistema čini mozak. Nervni sistem kod čoveka je sastavljen od oko 10^{11} neurona koji su međusobno povezani [76].



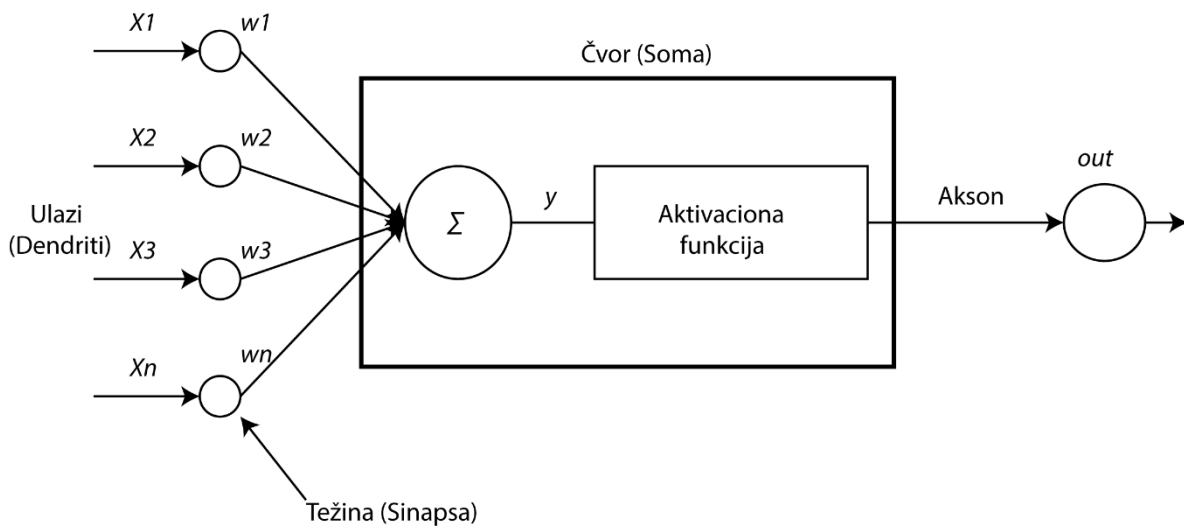
Slika 3. Biološki neuron

Dakle, procesiranje informacija koje koriste biološke neuronske mreže su jedan od načina na koji priroda rešava određeni problem, kao što su npr. procena, predviđanje, ili klasifikacija, na osnovu stečenog znanja i iskustva.

Veštačke neuronske mreže (ANN - *Artificial Neural Network*) predstavljaju matematičke modele kojima se na pojednostavljen način "imitira" biološki nervni sistem. Osnovni principi koji su preuzeti iz biološkog nervnog sistema su: i. nervni sistem obradu informacija obavlja uz pomoć paralelno distribuirane arhitekture sastavljene od jednostavnih funkcionalno povezanih procesnih jedinica (neurona). ANN su takođe sastavljene od više jednostavnih procesora (veštačkih neurona) koji su međusobno povezani; ii. nervni sistem je sposoban da uči [77]. Drugim rečima, ANN se ne programiraju, već uče, kao i ljudi, na osnovu "iskustva", tj. na osnovu predstavljenog izvesnog broja rešenih primera problema koji rešavaju. Osnovni delovi veštačkog neurona i njihova veza sa biološkim neuronom prikazani su na Slici 4. U ANN sistemima, veštački neuron prima aktivacioni signal od ulaza, koji predstavlja informacije iz sredine i svaki ulaz odgovara jednoj promenljivoj. Jedan od ključnih elemenata u ANN sistemima su težine veza ili težinski koeficijenti (*weights*), kojima se izražava relativna jačina veze kojom su neuroni povezani. Tokom procesiranja podataka kroz veštački neuron, svaki ulaz (x_i) se množi težinskim koeficijentom (w_i), a zatim sumira da bi se odredio nivo aktivacionog signala u neuronu. Izlazni signal (y) nakon sumiranja se posredstvom aktivacione funkcije procesira da bi se dobio izlazni signal iz neurona (*out*) [78–80]:

$$y = \sum_{i=1}^n w_i x_i \quad (5)$$

$$out = f(y) \quad (6)$$

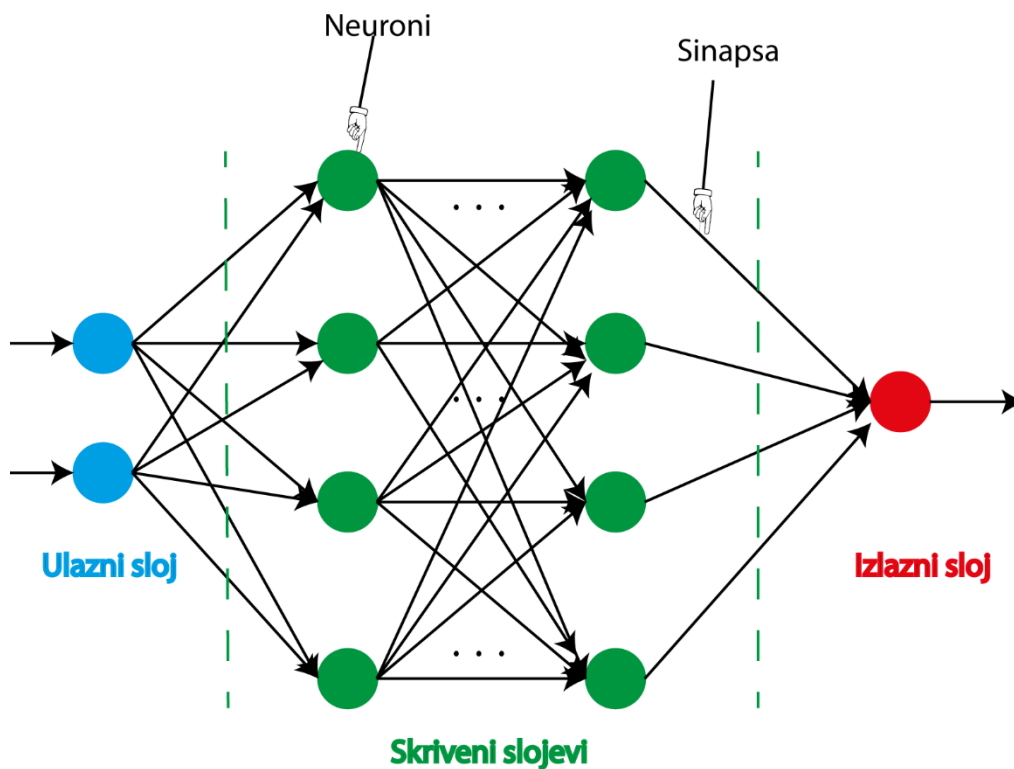


Slika 4. Šematski prikaz veštačkog neurona

3.2. Osnovni principi funkcionisanja veštačkih neuronskih mreža

Na osnovu prethodnih izlaganja se može uočiti da se veštačka neuronska mreža formira povezivanjem većeg broja nezavisnih neurona, grupisanih u slojeve, pri čemu su izlazi jednih sinapsama, tj. težinskim koeficijentima povezani sa ulazima drugih neurona. Neuroni u istom sloju uglavnom nisu međusobno povezani. Neuroni u ulaznom sloju primaju informacije iz spoljašne sredine, dok neuroni u izlaznom sloju generišu rešenje problema za koji se koriste ANN.

Najjednostavnija forma neuronskih mreža je perceptron, koji se sastoji od dva povezana sloja neurona (ulazni i izlazni). Međutim, ANN koje se danas koriste za rešavanje problema klasifikacije ili predviđanja, se obično sastoje iz jednog ulaznog sloja, jednog ili više skrivenih slojeva (*hidden layers*) i jednog izlaznog sloja [81]. Tipičan primer višeslojne mreže (MLP - *Multi-Layer Perceptron*) je prikazan na Slici 5.



Slika 5. Višeslojna neuronska mreža

Brojna istraživanja su pokazala da primena troslojne neuronske mreže sa jednim skrivenim slojem daje veoma zadovoljavajuće rezultate u rešavanju različitog spektra problema, uključujući klasifikaciju, predviđanje i aproksimaciju funkcija [82–85].

ANN može da se opiše kao paralelni procesor koji ima sposobnost čuvanja znanja stečenog iskustvom i njegovog korišćenja. Funkcionišu po principu "black box" modela, ne zahtevajući detaljne informacije o sistemu. Takođe, još jedna prednost ANN je njihova sposobnost da rešavaju probleme u sistemima u kojima postoji kompleksna, nelinearna povezanost između ulaznih i izlaznih parametara.

Proces obučavanja, ili treninga je izuzetno značajan segment u razvoju modela zasnovanih na ANN. ANN kroz proces obuke, ili treninga, uče iz iskustva, tj. primera prezentovanog u obliku numeričkih (ulaza-izlaza) podataka. Proces obučavanja se se zasniva na tome da se mreži zadaju ulazni i izlazni podaci, pri čemu mreža teži da dobije odgovarajuću izlaznu vrednost, tj. vrednost koja je približna očekivanoj vrednosti. Radi dobijanja željenog izlaza,

tokom obuke se kroz algoritam obuke menjaju težinski koeficijenti, dok mreža ne nauči da funkcioniše na odgovarajući način. Težine, tj. sinapse su objekat učenja mreža i u njima je "skladišteno" naučeno znanje mreže tokom procesa obuke [86].

Način na koji se, tokom procesa obučavanja mreže, podešavaju težinski koeficijenti zavisi od algoritma učenja. Jedna od najčešće korišćenih tehnika za postepeno smanjivanje greške je procedura poznata kao *gradient descent*. Jedan od najpoznatijih i najviše korišćenih algoritama za obučavanje mreža, koji je zasnovan na ovoj tehnici (*gradient descent*), je algoritam sa širenjem greške unazad (BP - *backpropagation*) i njegove varijante [87–89]. Detaljne informacije o načinu funkcionisanja i matematičko predstavljanje BP algoritma mogu se naći u odgovarajućoj literaturi vezanoj za ANN [90,91], dok će u ovom radu biti prikazane samo osnovne karakteristike ovog algoritma.

BP algoritam se koristi za obučavanje mreža sa kretanjem signala napred. Proces treninga mreže počinje sa postavljanjem inicijalnih nasumično odabranih težinskih koeficijenata i generalno je organizovan u dve faze. Najpre se u prvoj fazi mreži prezentuju ulazni parametri iz trening skupa podataka koji se prostiru kroz mrežu, kroz skriveni i izlazni sloj neurona, radi dobijanja izlaza, tj. računaju se odzivi svih neurona. Ova faza predstavlja fazu prostiranja "unapred" (*forward*). Poređenjem dobijenog (*no*) i željenog izlaza (*do*) računa se greška (*E*) mreže:

$$E = \frac{1}{2} \left[\sum_p \sum_i |no_{ip} - do_{ip}|^2 \right]^{1/2} \quad (7)$$

Nakon toga sledi faza modifikacije težinskih koeficijenata po gradijentnom postupku. Ova faza počinje od izlaznog sloja, pri čemu se signal greške prostire od izlaza ka ulazu i predstavlja fazu prostiranja unazad (*backward*). Faze prostiranja kroz mrežu koje su opisane, se vrše sekvencijalno za sve podatke iz trening skupa i ciklus obrade ovih podataka se naziva epoha. Princip izvršenja BP algoritma se može sumirati kroz sledeće korake [92,93]:

1. definišu se težinski koeficijenti,
2. prezentuje se mreži prvi ulazni vektor iz skupa podataka za trening,

3. ulazni vektor se prostire kroz mrežu radi dobijanja izlaza,
4. računa se signal greške poređenjem dobijenih i željenih izlaza iz mreže,
5. signal greške se zatim vraća unazad kroz mrežu,
6. podešavaju se težinski koeficijenti radi smanjenja ukupne greške,
7. ponavljaju se koraci 2-7 sa sledećim ulazom, sve dok se ne dobije zadovoljavajuća vrednost ukupne greške mreže.

Pronalaženje kombinacije težinskih koeficijenta radi minimiziranja greške mreže nije jednostavan proces. Problemi koji se uočavaju kod BP algoritma, kao i kod drugih algoritama koji koriste tehniku gradijentnog spusta, je da je učenje sporo, kao i da nije garantovano pronalaženje globalnog minimuma funkcije greške. Problem pronalaženja najmanje greške mreže je uslovljen pojavom "robusnosti" površine greške, tj. pojavom lokalnih minimuma funkcije greške. U tom smislu je poželjno da algoritam za obuku ne bude zarobljen u nekom od lokalnih minimuma, kao konačnim rešenjem, " ne znajući" da postoji globalni minimum [94]. Za rešavanje ovih problema, BP algoritam koristi dva optimizaciona parametra:

- impuls (*momentum*) i
- stopa obučavanja (*learning rate*).

Impuls definiše uticaj prethodnih na sadašnju promenu vrednosti težinskih koeficijenata (koji će biti veoma mali u lokalnom minimumu). Uvođenjem impulsa postignuto je preskakanje lokalnih minimuma.

Stopa obučavanja reguliše brzinu promene težinskih koeficijenata tokom iterativnog *gradient descent* procesa obučavanja. Ukoliko je zadata velika vrednost stope obučavanja, greška mreže će se menjati netačno zbog velike promene težinskih koeficijenata. S druge strane, ukoliko je zadata previše mala vrednost stope obučavanja, proces obučavanja traje dugo.

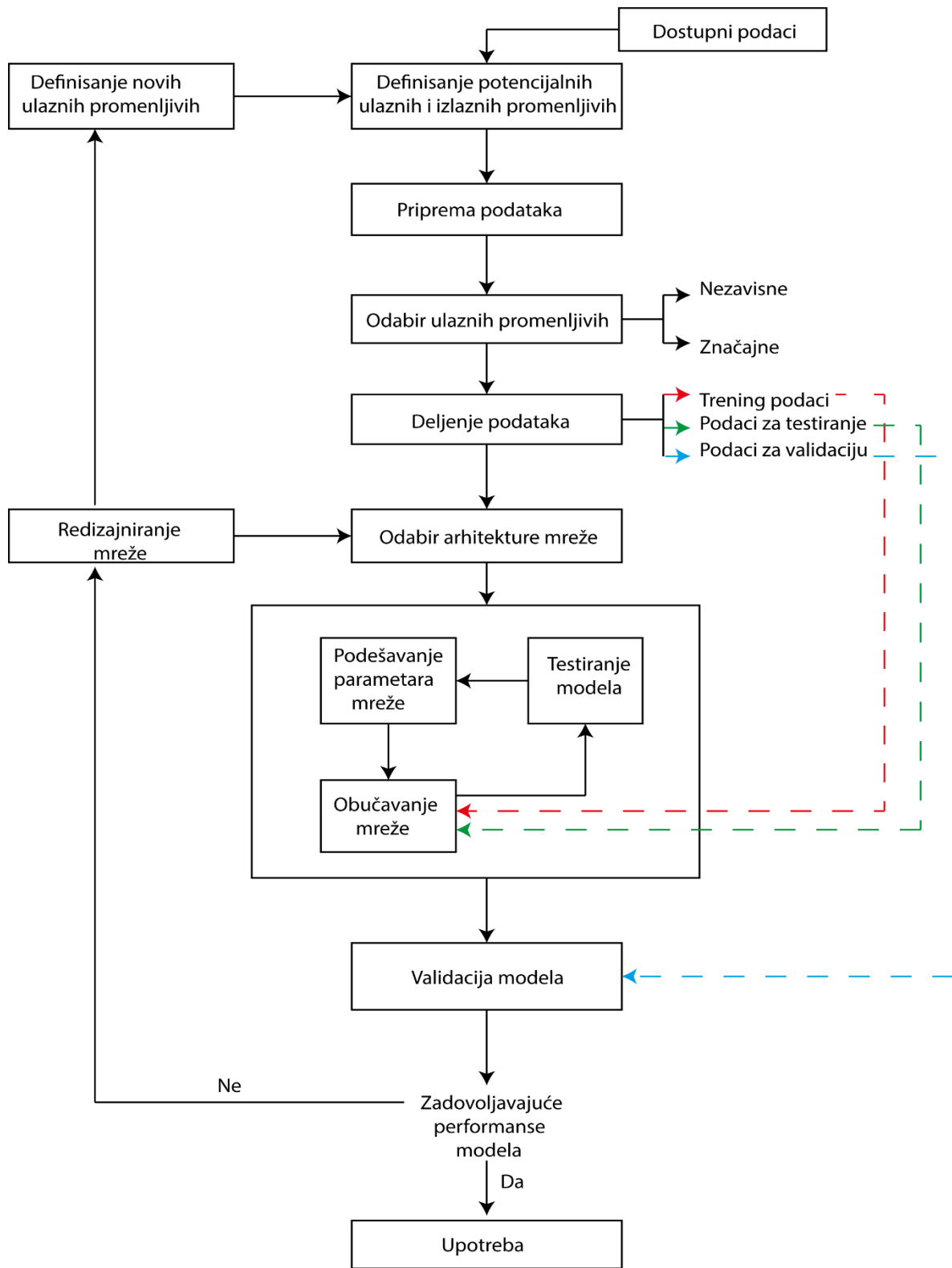
Oba optimizaciona parametra imaju vrednosti između 0 i 1 i nije jednostavno pronaći optimalnu kombinaciju vrednosti ovih parametra koja će dati najbolje rezultate [95]. Inicijalna vrednost stope obučavanja, kao i impulsa najčešće iznose 0.1 [96].

U praksi, radi postizanja zadovoljavajućih rezultata mreže, tj. kada je greška mreže dovoljno mala i kada daje zadovoljavajuće rezultate na izlazu, broj iteracija na skupu podataka za obučavanje je više hiljada. U tom smislu nije lako odrediti koliko dugo će trajati proces obučavanja mreže. Malim brojem iteracija dobijamo "lošiju" mrežu, koja nije u stanju da "nauči" zakonitosti problema, dok s druge strane, ukoliko je proces obučavanja sproveden kroz veliki broj iteracija (predugo obučavanje), to može dovesti do "pretreniranosti" mreže, koje se ogleda u tome da mreža ne uči zakonitosti problema, već memoriše podatke iz skupa za obučavanje mreže. Proces treninga bi trebalo zaustaviti tek kada se sagledaju mogućnosti generalizacije mreže na potpuno novim podacima za testiranje mreže.

I pored navedenih problema, kako osnovni, tako i unapređene varijante BP algoritma, koje koriste različite tehnike optimizacije (heurističke i standardne numeričke), našle su široku primenu u praksi [97,98].

3.3. Proces razvoja modela zasnovanih na veštačkim neuronskim mrežama

Sa ciljem dobijanja ANN modela koji će na zadovoljavajući (najbolji) način rešavati zadati problem, veoma je važno da se pre početka razvoja modela definišu osnovni koraci u razvoju modela. Definisane i podela procesa razvoja na odgovarajuće faze se ne može striktno odrediti, ali generalno, protokol razvoja ANN modela se može izvesti oslanjajući se na deset iterativnih koraka za razvoj modela iz oblasti zaštite životne sredine, predloženih od strane *Jakeman*-a i saradnika [99], koji obuhvataju celokupni naučni proces, uključujući formulisanje hipoteza, prikupljanje i opservaciju podataka, itd. Proces razvoja ANN modela je algoritamski predstavljen na Slici 6.



Slika 6. Šematski prikaz procesa razvoja neuronske mreže

3.3.1. Odabir ulaznih promenljivih za razvoj modela

Prvi korak u razvoju ANN modela je definisanje problema koji se rešava ovim pristupom, kroz odabir izlaznih veličina neuronske mreže. S druge strane, definisanje potencijalnih ulaznih promenljivih predstavlja veoma značajan segment, jer zahteva detaljno poznavanje problematike za koju se model razvija. Izborom potencijalnih ulaznih promenljivih, na osnovu dostupnih podataka, potrebno je da se mreži prezentuju sve one promenljive koje imaju uticaja na problem koji se rešava, budući da mreža uči zakonitosti između ulaza i izlaza, koje zatim koristi za generisanje rešenja.

U sledećoj fazi razvoja modela, pristupa se pripremi podataka na osnovu definisanih potencijalnih ulaznih i izlaznih promenljivih. Priprema podataka podrazumeva prikupljanje i procesiranje podataka, pri čemu je veoma značajno da broj skupova podataka (ulaznih i izlaznih promenljivih) za obučavanje mreže bude dovoljan da "pokrije" primerima ceo domen razmatranog problema. Jedno od često korišćenih empirijskih pravila je da broj skupova podataka za obučavanje neuronske mreže mora biti bar deset puta veći od broja ulaznih promenljivih [94].

Faza odabira ulaznih promenljivih podrazumeva selekciju odgovarajućih ulaznih promenljivih, od prvobitno izabranih, primenom različitih tehnika. Selekcija je veoma značajan deo u procesu razvoja ANN modela, jer performanse samog modela uglavnom zavise od prezentovanih podataka [100]. Naime, ukoliko se isključe jedna ili više ulaznih promenljivih iz polaznog skupa, to za posledicu može imati nemogućnost modela da nauči zakonitosti i "ulaz-izlaz" veze. S druge strane, ukoliko su u proces razvoja modela uključene irelevantane, ili suviše ulazne promenljive, to doprinosi smanjenju performansi modela, pojavi šuma u podacima, povećanju kompleksnosti modela, povećava se verovatnoća pretreniranosti mreže, itd. [101].

Selekcija ulaznih promenljivih (*IVS - Input Variable Selection*) se može definisati kao proces pronalaženja "odgovarajućeg" podskupa S značajnih varijabli iz izvornog skupa C potencijalnih ulaznih promenljivih [102]. Osnovni ciljevi selekcije ulaznih promenljivih su [103]:

- Pобољшanje predikcionih performansi prediktora (ulaznih promenljivih),
- Smanjenje dimenzionalnosti podataka,
- Brže i ekonomičnije modelovanje,
- Povećanje razumljivosti dobijenih podataka.

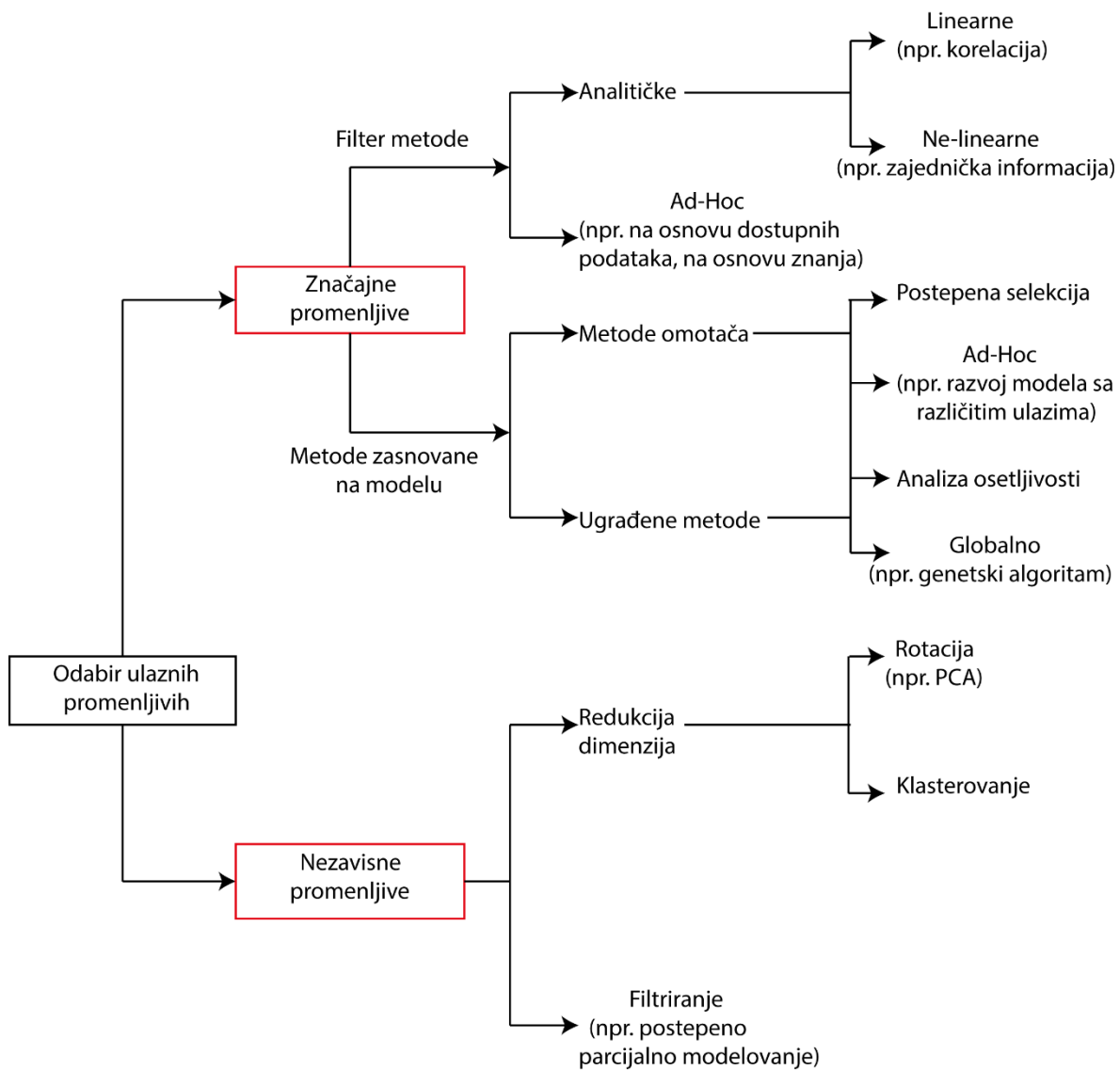
Za razvoj ANN modela IVS se zasniva na odabiru "značajnih" i "nezavisnih" promenljivih. Postoje brojne tehnike koje se koriste radi utvrđivanja značajnosti veze između potencijalnih ulaza i izlaza (Slika 7.). Ove tehnike se generalno mogu podeliti na:

- filter metode, koje nisu zasnovane na modelu (*model free*),
- metode zasnovane na modelu (*model-based*), u koje spadaju metode "omotača" (*wrappers*), tj. metode prethodnog učenja i ugrađene metode (*embedded*).

Filter metode ocenjuju značajnost ulaza nezavisno od modela, dok metode zasnovane na modelu zahtevaju procenu značajnosti ulaza na osnovu performansi modela kreiranih sa datim ulazima.

U zavisnosti od problema koji se rešava, neretko prilikom izbora potencijalnih ulaza dolazi do pojave multikolinearnosti. U tom smislu je, za razvoj ANN modela, značajno imati nekorelisane ulazne promenljive, jer korelisani podaci unose pometnju u neuronskim mrežama tokom procesa obuke [104]. IVS tehnike za odabir "nezavisnih" ulaznih promenljivih se mogu podeliti na dve kategorije (Slika 7.) [105]:

- metode redukcije dimenzija i
- tehnika filtriranja.



Slika 7. IVS tehnike za odabir ulaznih promenljivih

Metoda redukcije dimenzija se može postići primenom analize glavnih komponenti (PCA - *Principal Component Analysis*), ili tehnikom klasterovanja ulaznih varijabli, kojom se bira reprezentativni ulaz iz svakog klastera koji će se dalje razmatrati [106,107]. Drugi pristup za odabir "nezavisnih" ulaza je tehnika filtriranja. Najistaknutiji primer ove tehnike je postepen (*stepwise*) proces izgradnje modela, gde se na osnovu skupa statističkih kriterijuma postepeno biraju ulazne promenljive koje imaju najveći uticaj. Primeri ove tehnike za izbor ulaznih promenljivih su IVS algoritmi poput parcijalne korelacije (*partial correlation*) i

parcijalne zajedničke informacije (PMI - *Partial Mutual Information*) [108]. Treba napomenuti da ova podela IVS metoda nije striktno definisana, jer se mnoge metode koje se koriste za odabir "nezavisnih", takođe mogu koristiti i za odabir "značajnih" promenljivih, poput korelacione analize, PMI metoda i novijeg rekurzivnog algoritma za odabir ulaznih promenljivih (RVS - *Recursive Variable Selection*) [109].

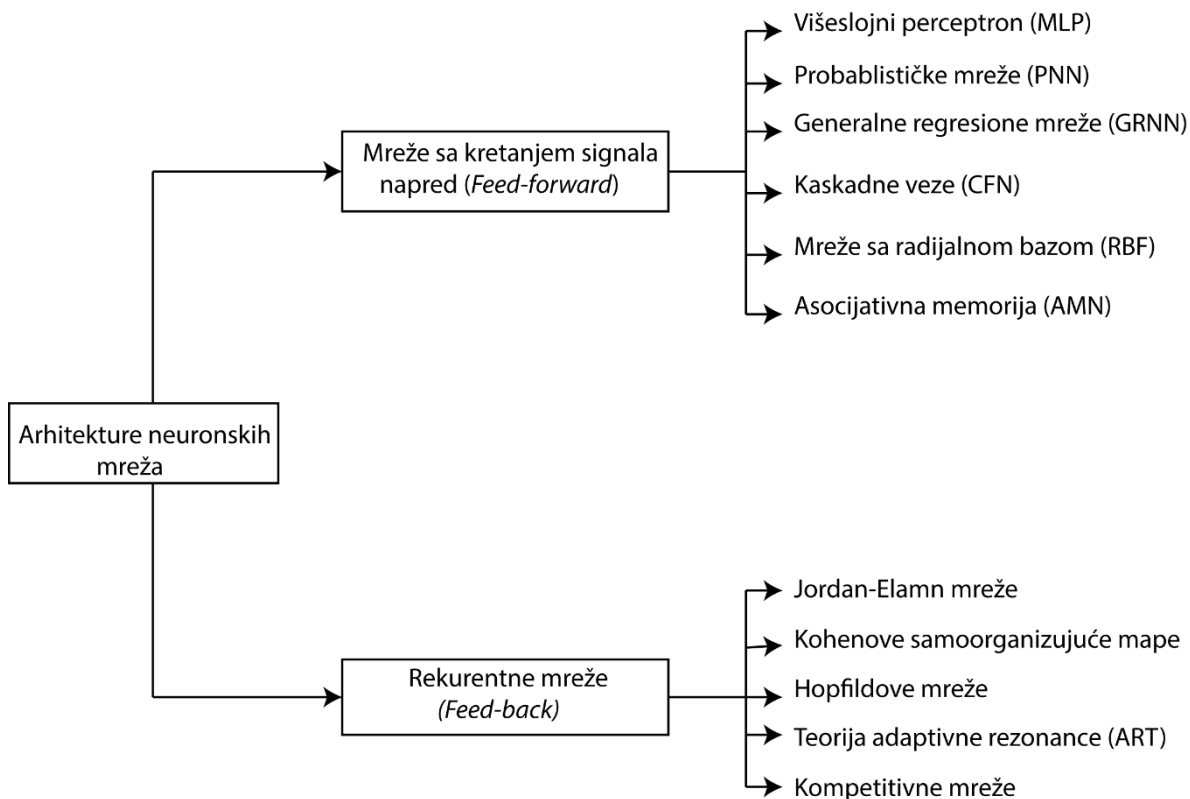
Detaljnije informacije o tehnikama za odabir relevantnih ulaznih promenljivih koje su primenjene u ovom radu biće prikazane u poglavlju 4.3.

3.3.2. Odabir arhitekture mreže

Sledeća faza u razvoju ANN modela je izbor odgovarajuće arhitekture mreže. Arhitektura mreže određuje celokupnu strukturu i protok informacija kod ANN modela, tako da ima značajnu ulogu u funkcionisanju veze između ulaznih i izlaznih promenljivih. Postoje brojne arhitekture mreža koje se mogu koristiti radi rešavanja različitih problema, neke od njih, podeljene u odnosu na tok signala, su prikazane na Slici 8. [85,110].

Mreže sa kretanjem signala napred (*feed-forward*) – kod ovih mreža signal se prenosi od ulaznog, preko različitog broja skrivenih do izlaznog sloja u jednom pravcu. Tipičan primer *feed-forward* arhitekture je MLP. Sposobnost MLP da koristi različite aktivacione funkcije u različitim slojevima neurona, omogućuje MLP da "hvata" kompleksne i ne-linearne veze unutar sistema. Iako danas postoje različite arhitekture neuronskih mreža, proteklih godina, u većini aplikacija neuronskih mreža MLP je primenjen za rešavanje problema [111].

Rekurentne mreže (*feed-back*) – kod ovih mreža signal se ne prenosi samo napred od ulaznog ka izlaznom sloju, već se može kretati i unazad od izlaznog sloja do ulaznog i/ili skrivenog sloja kroz povratne spege [112]. U poslednjih nekoliko godina rekurentne mreže sve više nalaze primenu. Veoma su pogodne za vremenske serije podataka [113].



Slika 8. Arhitekture neuronskih mreža

3.3.3. Obučavanje i validacija neuronskih mreža

Nakon odabira odgovarajuće arhitekture, sledeća faza u razvoju ANN modela je obučavanje neuronske mreže. Kao što se može videti na Slici 6., procesu obučavanja prethodi podešavanje parametara mreže. Određivanje parametara mreže uključuje određivanje parametara arhitekture, kao što je broj neurona u skrivenom sloju i način na koji su oni povezani (npr. broj neurona u slojevima, broj slojeva), način na koji oni procesiraju signale (npr. vrsta aktivacione funkcije) i definisanje parametara obučavanja. Proces obučavanja neuronske mreže je opisan u poglavlju 3.2.

Deo procesa razvoja ANN modela je deljenje prvobitnog seta podataka na trening podatke, podatke za testiranje i podatke za validaciju mreže. Trening set predstavlja grupu podataka koji se koriste za obučavanje mreže. Podaci za testiranje, koji se izdvajaju iz trening skupa, su grupa podataka koji se prezentuju mreži tokom procesa obučavanja radi evaluacije greške

modelovanja, ažuriranja najboljih težina i sprečavanja pretreniranosti mreže (oko 15% podataka trening seta). Podaci za validaciju predstavljaju potpuno nove podatke koji nisu prezentovani mreži u procesu obučavanja, radi procene generalizacije razvijenog modela [114].

Fazom validacije modela proveravaju se njegove mogućnosti za rešavanje određenog problema u realnim uslovima, poređenjem rezultata koje daje kreirani ANN model i stvarnih rezultata. Ukoliko rezultati razvijenog modela nisu zadovoljavajući, pristupa se ponovnoj analizi procesa razvoja. U tom slučaju, vrši se redizajniranje mreže koje se može sprovesti na neki od sledećih načina [113]:

- definisanjem novih ulaznih promenljivih;
- promenom arhitekture mreže;
- promenom strukture mreže (povećanje broja slojeva i neurona u skrivenom sloju);
- promenom parametara obučavanja mreže;
- promenom načina odabira podataka za testiranje mreže u procesu obučavanja;
- promenom veličine mreže (ukoliko je proces obučavanja previše dug, preporučuje se raščlanjivanje problema na nekoliko manjih problema, za koje će se razviti zasebne neuronske mreže).

3.4. Primena neuronskih mreža za modelovanje emisija zagađujućih materija

Sposobnost ANN da nauče i "proniknu" u često kompleksne i nepoznate veze među podacima, uslovlila je njihovu široku primenu na različitim poljima. ANN se koriste za predviđanje, aproksimaciju, klasifikaciju, obradu podataka, predviđanje kretanja na tržištu, simulacije iz oblasti zaštiti životne sredine, itd.

Primena ANN modela za predviđanja u oblasti zaštite životne sredine počela je početkom devedesetih godina prošlog veka i poslednjih godina je njihova primena značajno porasla. Pregledom dostupne literature, može se zaključiti da su ANN veoma korisne za razne simulacije u oblasti zaštite životne sredine [50,115–118].

Primena neuronskih mreža za simulaciju pojava i predviđanje u oblasti vazduha i klimatskih promena, počela je sa modelovanjem koncentracija različitih zagađujućih materija u vazduh [119–121]. Da su ANN modeli uspešni u predviđanju, pokazuju i novija istraživanja, u kojima je takođe, pokazana prednost ovih modela u odnosu na klasične regresione modele [87,122–127].

Od naročitog značaja za problem koji se analizira u ovom radu su rezultati istraživanja u kojima su modelovane emisije zagađujućih materija u vazduh. Postojeći literaturni podaci pokazuju da su ANN korišćene, uglavnom, za predikciju emisije različitih zagađujućih materija iz različitih izvora emisije [128–133], dok je broj istraživanja u kojima je modelovana emisija gasovitih zagađujućih materija na nacionalnom nivou, zanemarljivo mali [51,52,134,135].

Eksperimentalni deo

4. Materijali i metode

4.1. Prikupljanje ulaznih i izlaznih podataka

Da bi razvijeni ANN model dao zadovoljavajuće rezultate, neophodno je u proces razvoja modela uključiti sve relevantne ulazne promenljive. To podrazumeva detaljan i sistematičan pristup u pronalaženju veza između emisije određenih zagađujućih materija u vazduh i nezavisnih promenljivih. Budući da je kroz brojna istraživanja dokazano da ekonomski i industrijski razvoj kao posledicu ima povećanu emisiju polutanata u vazduh, u okviru ove teze su za razvoj ANN modela, upravo, posmatrane međusobne zavisnosti indikatora održivog razvoja, ekonomski i industrijski indikatori i emisija gasovitih zagađujućih materija.

U okviru ovog rada vršeno je modelovanje emisije gasovitih zagađujućih materija na nacionalnom nivou i to: amonijaka, nemetanskih organskih isparljivih materija, azotnih oksida, metana i gasova stalene bašte. Potencijalne ulazne promenljive za razvoj modela određene su i na osnovu literaturnih podataka o uticaju pojedinih sektora izvora emisije na emisiju određenog polutanta.

S obzirom na to da su u modelovanju emisija uključene različite države (države EU, Kina, SAD, Rusija, Japan i Indija), radi prikupljanja podataka pretraživane su brojne međunarodne statističke baze podataka koje prate, kako indikatore koji se odnose na zagađenje vazduha, tako i indikatore koji se odnose na ekonomski, industrijski i poljoprivredni stepen razvoja. Neke od međunarodnih baza podataka čiji su podaci korišćeni u okviru ove teze su: Eurostat, Evropska agencija za zaštitu životne sredine (EEA), baza podataka emisije za globalna atmosferska istraživanja (EDGAR), Svetska banka, nacionalni poljoprivredni statistički servis Sjedinjenih Država (USDA - *U.S. Department of Agriculture*), nacionalni statistički servis iz oblasti energetike Sjedinjenih Država (USEIA - *U.S. Energy Information*

Administration), agencija za zaštitu životne sredine Sjedinjenih Država (EPA - *United States Environmental Protection Agency*), nacionalni zavod za statistiku Kine i dr.

4.2. Arhitekture mreža primenjene za razvoj modela

4.2.1. Višeslojni perceptron (MLP)

MLP mreže su, kao što je već rečeno, neuronske mreže sa kretanjem signala napred i po tipu obučavanja spadaju u mreže sa nadgledanim obučavanjem (*supervised learning*) i sposobne su da rešavaju složene probleme. Tipičan MLP se sastoji iz: jednog ulaznog sloja, skrivenog (jedan ili više) sloja i izlaznog sloja neurona [136]. I pored toga što nije striktno ograničen broj skrivenih slojeva neurona, u praksi se uglavnom koriste MLP mreže sa jednim ili dva skrivena sloja neurona. Topologija MLP mreže korišćene u ovom radu je standardna troslojna MLP mreža, ili neuronska mreža sa propagacijom greške unazad (BPNN - *Backpropagation Neural Network*), koju čini ulazni, skriveni i izlazni sloj neurona. Prilikom razvoja MLP modela, broj neurona u ulaznom i izlaznom sloju je jednak broju ulaznih i izlaznih promenljivih, dok se optimalan broj neurona u skrivenom sloju najčešće određuje po principu "probe i greške" [137].

Za postizanje dobrih performansi MLP modela, ključni parametri koji se moraju odrediti tokom razvoja modela su broj neurona i težinski koeficijenti u skrivenom sloju. Pored toga, u skrivenom sloju ulazi se sumiraju i procesiraju posredstvom aktivacione funkcije, koja takođe predstavlja jedan od parametra koji utiče na rezultate MLP modela. Postoje različite vrste aktivacionih funkcija, neke od njih koje se koriste su: logistička, linearna, sinusna, Gausova. Izbor optimalnog skupa težinskih koeficijenata, broja skrivenih neurona i aktivacione funkcije tokom procesa obučavanja, pokazao je da MLP može aproksimirati bilo koju merljivu funkciju između ulaznih i izlaznih vektora [82,83].

Za obučavanje MLP mogu se koristiti različiti algoritmi i kod većine njih je proces učenja zasnovan na iterativnom postupku, međutim BP algoritam i unapređeni algoritmi proistekli iz standardnog BP algoritma su najčešće primenjivani [87]. Varijante BP algoritma koje se mogu koristiti za obučavanje MLP su zasnovane na iskustvenim i standardnim numeričkim

tehnikama optimizacije [98]. Kod ovih algoritama traženje minimuma greške je bazirano na gradijntnom pristupu nelinearne optimizacije i tu spadaju:

- i. Pristup sa opadajućim gradijntom stope obučavanja (*learning rate gradient descent*) – težinski koeficijenti (w) se postepeno ažuriraju u zavisnosti od promena stope obučavanja i gradijnta greške. Brzina konvergencije ove varijante BP je znatno veća u odnosu na standardni BP algoritam i omogućena je promena stope obučavanja tokom procesa obuke mreže [138]:

$$w_{k+1} = w_k - \alpha_k g_k \quad (8)$$

gde je w_k vektor trenutnih težinskih koeficijenata, α_k je stopa obučavanja (*learning rate*) i g_k je trenutni gradijnt greške.

- ii. Pristup sa konjugovanim gradijntom – za razliku od standardnog BP algoritma kod koga se težinski koeficijenti podešavaju u smeru negativnog gradijnta (smer najbržeg spusta), kod algoritama sa konjugovanim gradijntom, minimum greške se traži u suprotnom smeru, u odnosu na prvi prikazani algoritam, čime se povećava brzina konvergencije u odnosu na osnovni BP algoritam [97]:

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

gde je y_k konjugovani gradijnt greške čija se vrednost dobija na osnovu vrednosti prethodnog gradijnta greške :

$$y_k = -g_k + \beta_k y_{k-1} \quad (10)$$

U zavisnosti od načina na koji se određuje vrednost konstante β_k (*momentum*) postoje više varijante ovog algoritma: *Fletcher - Reeves*, *Polak - Ribiere* i *Powell - Beale* procedure izračunavanja konstante [98,139].

- iii. Njutnov pristup optimizacije – za razliku od prva dva, Njutnov pristup optimizacije je znatno brži. Kod ovog algoritma težinski koeficijenti se računaju prema sledećoj jednačini:

$$w_{k+1} = w_k - A_k g_k \quad (11)$$

gde A_k predstavlja *Hessian* matricu drugog izvoda funkcije greške za trenutne vrednosti težinskih koeficijenta [97,98]. Proračun *Hessian* matrice je složen proces i može se vršiti *Quasi-Newton*, ili *Levenberg-Marquardt* metodama [140,141].

U okviru *Neuroshell 2* softvera, koji je u ovom radu korišćen za razvoj ANN modela, implementirane su tri varijante BP algoritma za obučavanje MLP mreže [142]:

1. *Vanilla* algoritam koji pripada BP algoritmu sa opadajućim gradijentom stope obučavanja kod koga se za ažuriranje težinskih koeficijenata koristi samo stopa obučavanja, ali ne i *momentum* (impuls);
2. *Momentum* algoritam koji pripada BP algoritmu sa konjugovanim gradijentom kod koga se za ažuriranje težinskih koeficijenata koristi ne samo stopa obučavanja već i *momentum* (impuls);
3. *TurboProp* algoritam predstavlja način obuke koji radi mnogo brže u "paket" režimu nego ostali BP algoritmi i ima prednosti u odnosu na prethodna dva algoritma, jer nije osetljiv na parametre, stopu obučavanja i impuls.

4.2.2. Generalne regresione neuronske mreže (GRNN)

GRNN su razvijene početkom devedesetih godina prošlog veka, sa ciljem rešavanja problema aproksimacije nelinearnih funkcija [143,144]. GRNN koristi funkciju raspodele verovatnoće (*probability distribution function*) za predviđanje diskretnih vrednosti verovatnoća određenog događaja [96]. Neuronska mreža sa opštom regresijom (GRNN), srodna mrežama sa radijalnom osnovom (RBF), zasnovana je na standardnoj statističkoj tehnici nazvanoj *Gaussian kernel regression* [145], kojom se može proceniti povezanost u bilo kom uzorku podataka, kao i proceniti mogućnost konvergencije ka optimalnoj regresionoj površini sa ograničenim brojem podataka [126]. GRNN funkcioniše po principu

merjenja koliko je daleko dati skup podataka koje mreža obrađuje u odnosu na trening set podataka u N -dimenzionalnom prostoru, pri čemu je N broj ulaza [86]. U okviru *Neuroshell 2* softvera se za izračunavanje rastojanja u N - dimenzionalnom prostoru mogu koristiti *Vanilla* i *City block* algoritami. U ovom radu je korišćen *Vanilla* algoritam koji rastojanja između podataka koje mreža obrađuje i skupa za obučavanje računa kao u Euklidovom prostoru, prema jednačini:

$$r = \sqrt{\sum_{i=1}^N x_i^2} \quad (12)$$

gde je: r - ukupno rastojanje, x_i – rastojanje po jednoj koordinati i N – broj dimenzija.

Kao i MLP, GRNN spada u neuronske mreže sa kretanjem signala napred i koja po tipu obučavanja spada u nadgledane mreže. Prednost GRNN u odnosu na MLP se ogleda u njihovoj jednostavnosti, ne zahtevaju iterativnu proceduru obuke radi postizanja željenog rešenja, brzom obučavanju i sposobnosti da daju dobro predviđanje sa znatno manje ulaznih podataka [126,146]. GRNN je naime, "one-pass" mreža sa jednom iteracijom tokom procesa obuke i sastoji se od četiri sloja neurona, pri čemu svaki sloj ima različitu ulogu: ulazni, skriveni (obrazac), sumarni i izlazni sloj. Broj neurona u svakom GRNN sloju je određen brojem ulaznih/izlaznih promenljivih i brojem serija podataka u skupu za obučavanje:

- Ulazni sloj - svaki neuron u ovom sloju predstavlja jednu ulaznu promenljivu,
- Skriveni sloj (obrazac) – broj neurona u ovom sloju je ekvivalentan broju serija podataka iz skupa za obučavanje. Ulazne vrednosti primljene iz ulaznog sloja su nelinearno transformisane u ovom sloju i neuroni ovog sloja mogu da memorišu vezu između ulaznih neurona i odgovarajućeg odziva skrivenog sloja [147],
- Sumarni sloj – broj neurona u ovom sloju je ekvivalentan broju izlaznih neurona plus jedan,
- Izlazni sloj – broj neurona u ovom sloju je jednak broju izlaznih promenljivih.

Za razliku od MLP kod kojih se proces obuke mreža podešava parametrima stope obučavanja i impulsom, kod GRNN mreže se za podešavanje koristi parametar faktor ujednačavanja (*smoothing factor*), koji se mora odrediti tokom procesa obučavanja mreže. Ovaj parametar

definiše predikcione sposobnosti GRNN i može se predstaviti kao širina Gausove krive za svaku funkciju raspodele verovatnoće određene tokom procesa obuke [148]. Određivanje optimalne vrednosti faktora ujednačavanja se može vršiti: metodom "probe i greške" dok se ne dobije vrednost koja daje najbolje rezultate, ili primenom jednog od *NetPerfect* algoritma, iterativnog ili genetskog algoritma. U ovom radu je za određivanje optimalne vrednosti faktora ujednačavanja korišćen genetski algoritam.

4.2.3. Rekurentne neuronske mreže (RNN)

RNN spadaju u mreže kod kojih se signal može prostirati u oba smera, tj. za razliku od standardnih (*feedforward*) kod kojih je svaki sloj neurona povezan samo sa slojem koji mu neposredno prethodi, kod rekurentnih mreža se uspostavljaju povratne veze među slojevima [149,150]. RNN se takođe, kao i MLP, nadgledano obučavaju BP algoritmom. Razlika u strukturi, u odnosu na standardnu troslojnu mrežu, je postojanje dodatnog sloja kojim se ostvaruje povratna veza. Ovaj dodatni skriveni sloj se naziva još i "dugoročnom" memorijom mreže, jer se u njemu memoriše izlaz mreže iz prethodne iteracije [51].

U okviru rekurentnih mreža, postoje više različitih topologija. U ovom radu je korišćena *Džordan-Elman* rekurentna mreža, kod koje je povratna veza uspostavljena u okviru skrivenog sloja. Kod navedenog tipa RNN se karakteristike uočene kod određenih ulaznih podataka paralelno obrađuju sa karakteristikama podataka koji im prethode, što doprinosi preciznosti ove mreže [86]. Broj neurona u ulaznom sloju je jednak broju ulaznih promenljivih, broj neurona u izlaznom sloju je jednak broju izlaznih promenljivih, dok se broj neurona u skrivenom sloju (B_{sn}) određuje po principu probe i greške kao kod MLP, ili primenom formule:

$$B_{sn} = \frac{B_u + B_i}{2} + \sqrt{B_{tr}} \quad (13)$$

gde su: B_u – broj ulaznih promenljivih, B_i – broj izlaznih promenljivih i B_{tr} – broj serija podataka u setu podataka za obučavanje. Ostali parametri za obučavanje RNN mreža su isti kao kod MLP.

4.3. Odabrane metode primenjene za selekciju ulaznih promenljivih

4.3.1. Analiza glavnih komponenti (PCA)

Analiza glavnih komponenti predstavlja jednu od najjednostavnijih multivarijantnih metoda, koja je našla široku primenu u oblasti zaštite životne sredine. Osnovni cilj PCA je redukcija broja prediktivnih promenljivih i njihova transformacija u nove nekorelisane promenljive, koje se nazivaju glavne komponente (PC - *Principal Components*) [116,151,152]. Nove promenljive, dobijene primenom PCA predstavljaju nezavisne linearne kombinacije originalnih podataka, uz zadržavanje maksimalno moguće varijanse originalnih promenljivih i minimalan gubitak informacija [153,154]. Korelisani ulazni podaci se za PCA prezentuju u vidu matrice podataka u kojoj se red odnosi na objekat (n -objekata), dok se kolona odnosi na određenu promenljivu (p -promenljivih), da bi se dobile p linearnih kombinacija originalnih promenljivih u vidu glavnih komponenti:

$$PC_1 = w_{11}O_1 + w_{12}O_2 + \dots + w_{1p}O_p \quad (14)$$

$$PC_2 = w_{21}O_1 + w_{22}O_2 + \dots + w_{2p}O_p \quad (15)$$

.

$$PC_p = w_{p1}O_1 + w_{p2}O_2 + \dots + w_{pp}O_p \quad (16)$$

gde su: $PC_1, PC_2, \dots, PC_p, p$ glavnih komponenta, $w_{11}, w_{21}, \dots, w_{pp}$ su koeficijenti tj. konstante koje čine linearne koeficijente određene originalne promenljive (O_1, O_2, \dots, O_p) za određenu glavnu komponentu. Konstante su procenjene tako da prva glavna komponenta PC_1 objašnjava maksimum varijanse iz podataka, druga glavna komponenta PC_2 objašnjava maksimum varijanse koja je ostala neobjašnjena prvom komponentom, itd. Deo te varijanse koja je objašnjena jednom glavnim komponentom naziva se svojsvena vrednost i njena vrednost je najveća u prvoj glavnoj komponenti, dok je u svakoj narednoj glavnoj komponenti njena vrednost manja. Tokom PCA cilj je da se iteracijskim postupkom izdvoji što veći deo ukupne varijanse u što manji broj glavnih komponenti, čime će se redukovati

broj originalnih promenljivih. Naime, tokom PCA pretpostavlja se da će većina novih glavnih komponenti predstavljati "šum" i imati zanemarljivo malu varijansu u odnosu na varijansu koja je sadržana u prvih nekoliko glavnih komponenti, čime se iz velikog broja originalnih promenljivih kreira nekoliko glavnih komponenti u kojima je sadržan veći deo informacija [155]. Da bi rezultati primene analize glavnih komponenti bili što bolji, poželjno je da originalne promenljive koje su uključene u analizu budu visoko pozitivno ili negativno korelisane.

Da bi sam proces PCA bio što efikasniji, predlažu se tri osnovna koraka [156]:

- i. Ocena prikladnosti podataka za analizu glavnih komponentata;
- ii. Izdvajanje glavnih komponenti;
- iii. Rotacija vektora.

Ocena prikladnosti podataka za PCA uključuje razmatranje dva pitanja, a to je veličina uzorka i jačina korelacije između originalnih promenljivih. Kada je reč o veličini uzorka za PCA, preporuke autora su različite, međutim većina se slaže da veći uzorak podataka daje bolje rezultate [157,158]. Drugo pitanje, koje treba razmotriti pre primene PCA, jeste jačina korelacije između originalnih promenljivih. Preporučuje se da je primena PCA prikladna ukoliko se u matrici korelacija nalazi veći broj koeficijenata korelacije čija je vrednost veća od 0.3 [159]. Za ocenu opravdanosti primene PCA, u literaturi najčešće se koriste dva testa: Bartletov test sferičnosti (*Bartlett's Sphericity test*) [160] i Kajzer-Majer-Olkinov (KMO - *Kaiser-Meyer-Olkin*) pokazatelj pogodnosti uzorka [161]. Bartletov test sferičnosti se može predstaviti jednačinom [162]:

$$\chi_k^2 = \left[n - k - \frac{2(p-k)+7}{6} + \frac{2}{(p-k)} + \sum_{j=1}^k \left(\frac{\bar{\lambda}}{\lambda_j - \bar{\lambda}} \right)^2 \right] \times \left[-\ln \prod_{j=k+1}^p \lambda_j + (p-k) \ln \bar{\lambda} \right] \quad (17)$$

gde p - predstavlja broj glavnih komponenti, λ_j - predstavlja svojstenu vrednost k glavne komponente, n - je broj opservacija u uzorku dok se $\bar{\lambda}$ dobija prema sledećoj jednačini:

$$\lambda = \sum_{j=k+1}^p \frac{\lambda_j}{p-k} \quad (18)$$

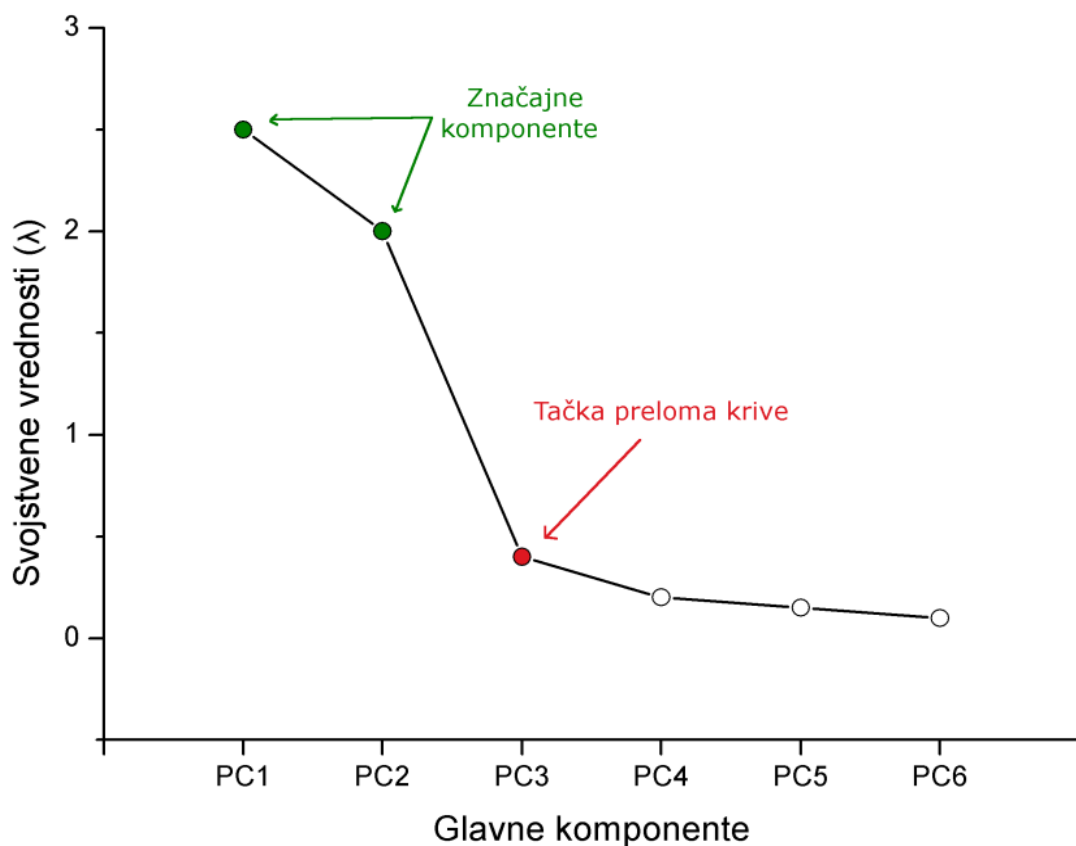
Nulta hipoteza ovog testa pretpostavlja da su sve dobijene promenljive međusobno nekorelisane i ukoliko je Bartletov test sferičnosti značajan ($p < 0.005$), polazna hipoteza se prihvata i primena PCA je opravdana [163].

Kajzer-Majer-Olkinov pokazatelj pogodnosti uzorka za primenu PCA je statistički parameter koji sumarno pokazuje koliko je mala parcijalna korelacija u odnosu na originalne (nultog reda) korelacije. Parcijalne korelacije za svaki par promenljivih u PCA se zasnivaju na povezanosti tih promenljivih, nakon što se isključe uticaji ostalih promenljivih u analizi. Ukoliko dve promenljive dele zajedničku glavnu komponentu sa ostalim promenljivima, njihova parcijalna korelacija će biti mala, ukazujući na to da dele zajedničku varijansu. Ukoliko parcijalna korelacija ima vrednosti približne 0, onda će vrednost KMO parametra biti bliska 1 i obrnuto. Kao što se može videti, KMO parametar može imati vrednosti između 0 i 1. Da bi PCA bila opravdana, u literaturi se preporučuje vrednost KMO testa iznad 0.6 [157].

Sledeći korak u realizaciji PCA je izdvajanje glavnih komponenti, koja predstavlja određivanje najmanjeg broja glavnih komponenti koje dobro predstavljaju međusobne veze u originalnom skupu promenljivih. To znači da se sa minimalnim brojem glavnih komponenti objasni što veći deo varijanse originalnog skupa promenljivih. Postoje različiti pristupi u određivanju broja glavnih komponenti:

- Istraživački pristup – podrazumeva da se proba sa različitim brojem glavnih komponenti dok se ne dobiju zadovoljavajući rezultati;
- Kajzerova tehnika ili kriterijum svojstvenih vrednosti – svojstvena vrednost komponente je ukupna varijansa svih promenljivih objašnjena tom komponentom. Prema ovoj tehnici se za dalji rad koriste samo one glavne komponente čije su svojstvene vrednosti veće od 1. Nedostatak ove tehnike je u tome što u nekim slučajevima zadržava previše glavnih komponenti;

- Tehnika dijagram prevoja (*scree test*) – broj glavnih komponenti se može odrediti i Katelovim kriterijumom, odnosno kriterijumom dijagrama prevoja [164]. Kod ovog pristupa se na osnovu dijagrama "svojtvene vrednosti/glavne komponente", nalazi tačka u kojoj se oblik krive menja i ona prelazi u horizontalu (Slika 9.). Prema Katelu, trebaju se zadržati one glavne komponente koje se nalaze iznad prevoja tog dijagrama, jer te glavne komponente najviše doprinose objašnjavanju varijanse u datom skupu podataka;
- Paralelna analiza – sve popularnija tehnika za izdvajanje glavnih komponenti. Paralelna analiza predstavlja upoređivanje iznosa svojstvenih vrednosti sa iznosima dobijenim na jednako velikom skupu slučajno generisanih podataka. Pri tome se zadržavaju samo one glavne komponente čije su svojstvene vrednosti veće od odgovarajućih vrednosti dobijenih na nasumično generisanim podacima [165,166].



Slika 9. Katelov kriterijum za određivanje značajnog broja glavnih komponenti [96]

Na kraju analize glavnih komponenti, nakon što je određen broj glavnih komponenti, sledeći korak bi bio "rotacija" komponenti. Rotacija se vrši radi lakšeg tumačenja glavnih komponenti. Naime, glavne komponente su često definisane tako da se jedna originalna promenljiva javlja u više glavnih komponenti. Rotacijom komponenti se struktura faktorskih težina (*factor loadings*), tj. koeficijenta korelacije između promenljivih i glavnih komponenti predstavlja na taj način da bi se dobila "jednostavna struktura". To bi značilo da svaka promenljiva jako korelira sa samo jednom glavnom komponentom, odnosno da samo toj glavnoj komponenti daje veliku težinu, a svaka komponenta bi bila definisana sa više sa njom jako koreliranih promenljivih, odnosno da svakoj glavnoj komponenti više promenljivih da veliku težinu [167]. U praksi transformacija glavnih komponenti može biti ortogonalna (nekorelirane komponente) ili kosa (korelirane komponente). U te dve kategorije rotacija spada više različitih tehnika među kojima su najpoznatije: *ortogonalna - Varimax* (koja pokušava da smanji broj promenljivih sa visokim apsolutnim vrednostima faktorskih težina) i *kosa - Direct Oblimin* rotacija koja se najčešće upotrebljava. I jedan i drugi pristup često daju slična rešenja. Detaljnije informacije o navedenim vrstama rotacije se mogu naći u literaturi [157].

Za realizaciju analize glavnih komponenti, u ovom radu je korišćen *SPSS 19* softver [168].

4.3.2. Korelaciona i Faktor povećanja varijanse (VIF) analiza

Korelaciona analiza spada u statističke metode na osnovu kojih se može, merenjem jačine veze između dve promenljive, odrediti kako značajnost, tako i nezavisnost date promenljive. Dakle, korelaciona analiza opisuje smer i jačinu linearne veze dve promenljive. Za merenje jačine veze najčešće se koriste koeficijent Pirsonove korelacije i koeficijent Spirmanove korelacije ranga. Koeficijent Pirsonove korelacije se izračunava prema formuli:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x}_n)(y_i - \bar{y}_n)}{\sqrt{\sum_{i=1}^n (x_i - \bar{x}_n)^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y}_n)^2}} \quad (19)$$

gde su: x i y vrednosti promenljivih čija se veza ispituje, \bar{x}_n i \bar{y}_n su srednje vrednosti. Koeficijent Pirsonove korelacije može imati vrednosti koje se kreću u opsegu od -1 do +1. Predznak ispred koeficijenta pokazuje da li je korelacija između promenljivih pozitivna (i jedna i druga promenljiva zajedno rastu ili opadaju), ili je korelacija negativna (kada jedna promenljiva raste tada druga opada i obrnuto). Apsolutna vrednost koeficijenta Pirsonove korelacije ukazuje na jačinu veze između datih promenljivih. U literaturi se mogu naći različita tumačenja jačine veze između dve promenljive [169,170]:

- Ukoliko r ima vrednosti od 0.1 do 0.29 korelacija je mala,
- Ukoliko r ima vrednosti od 0.3 do 0.49 korelacija je srednja,
- Ukoliko r ima vrednosti od 0.50 do 1.0 korelacija je velika.

Koeficijent Spirmanove korelacije rangova predstavlja neparametrijski ekvivalent Pirsonovom koeficijentu korelacije, jer ne pretpostavlja normalnu raspodelu promenljivih čija se veza ispituje. Kod ovog koeficijenta korelacije se vrednosti ne određuju na osnovu numeričkih vrednosti datih promenljivih, već se vrednost određuje na osnovu njihovih relativnih odnosa, tj. rangova. Koeficijent Spirmanove korelacije rangova se računa prema formuli:

$$\rho = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)} \quad (20)$$

gde su: d razlika između rangova x i y ($d = x_i - y_i$) i n veličina uzorka (broj rangova promenljivih x i y). Koeficijent Spirmanove korelacije takođe može imati vrednosti od -1 do +1. Veza između promenljivih je jača, pa je samim tim i vrednost Spirmanovog koeficijenta veća, ukoliko je razlika između rangova promenljivih koje se posmatraju manja [171]. Spirmanov koeficijent je pogodan za podatke koji su dati u ordinalnoj skali (ako se podaci rangiraju po veličini) i koristi se u slučajevima kada podaci ne zadovoljavaju kriterijume koji su potrebni za Pirsonovu korelaciju.

Statistička značajnost između Pirsonovog i Spirmanovog koeficijenta je znatno veća kod Pirsonovog koeficijenta korelacije i on se najčešće u literaturi koristi za procenu veze među promenljivama.

Variance Inflation Factor (VIF), ili faktor povećanja varijanse predstavlja jednu od statističkih tehnika koja se može primeniti za odabir nezavisnih ulaznih promenljivih pri razvoju modela zasnovanog na neuronskim mrežama. Naime, VIF analiza je pored korelacione analize, jedan od najčešće korišćenih statističkih pokazatelja za procenu stepena multikolinearnosti u podacima. VIF pokazuje koliki je deo stvarne varijanse određene nezavisne promenljive određen varijansama drugih nezavisnih promenljivih, tj. predstavlja odnos stvarne varijanse određene promenljive i varijanse koju bi ta promenljiva imala kada bi bila u potpunosti nezavisna u odnosu na ostale promenljive [172]. Vrednost VIF određene promenljive (prediktora) x_j se računa na osnovu linearne zavisnosti između prediktora x_j i ostalih nezavisnih promenljivih $[x_1, x_2, \dots, x_j, x_{j+1}, \dots, x_m]$, po formuli [173,174]:

$$VIF_j = \frac{1}{1 - R_j^2} \quad (21)$$

gde R_j^2 predstavlja koeficijent determinacije višestruke regresije.

Što je veća vrednost VIF to je veća multikolinearnost među datim promenljivama [175]. To bi značilo da se iz modela trebaju ukloniti sve one ulazne promenljive koje imaju visoku vrednost VIF. Međutim, ne postoji jasno definisana kritična vrednost VIF koja bi se smatrala "velikom" [174], tako da se u literaturi sreću vrednosti za VIF iznad 10, kao i one oko 4 koje su smatrane značajnim u pogledu postojanja multikolinearnosti [172].

Ako uporedimo ove dve metode, može se uočiti da VIF metoda ima blagu prednost u odnosu na korelacionu analizu. Razlog je taj što je korelaciona analiza ograničena time što može posmatrati vezu samo dve nezavisne promenljive u isto vreme, dok to nije slučaj kod VIF analize [175].

4.4. Regresioni modeli za komparaciju rezultata ANN modela

Radi komparacije rezultata razvijenih ANN modela za predviđanje emisije, u ovom radu će se koristiti konvencionalni linearni regresioni modeli, koji se u oblasti istraživanja vezanih za atmosferu najčešće koriste za predviđanje.

Višestruka linearna regresija (MLR - *Multiple Linear Regression*) je skup statističkih tehnika pomoću kojih se može istražiti međusobna zavisnost jedne zavisne promenljive i više nezavisnih promenljivih. Drugim rečima, MLR pokazuje koliko efikasno određeni skup nezavisnih promenljivih predviđa određeni ishod. Predstavlja jednu od najčešće korišćenih metoda za predviđanje koja je našla široku primenu u oblasti životne sredine kao i u oblasti istraživanja vezanih za atmosferu [136,162].

MLR tehnika je zasnovana na pronalaženju odgovarajućeg modela, tj. linearne kombinacije, koju čine skup nezavisnih promenljivih X_i , $i= 1, \dots, q$, koje procenjuju nepoznatu regresijsku funkciju zavisne promenljive Y . U MLR, linearna zavisnost između nezavisnih i zavisne promenljive je iskazana putem matematičke jednačine [176]. Generalni pristup MLR se može prikazati sledećom jednačinom [177]:

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_q X_q + \varepsilon_i \quad (22)$$

gde su: Y – zavisna promenljiva; $X_1, X_2, X_3 \dots X_q$ - nezavisne promenljive; β_0 – konstanta; $\beta_1, \beta_2 \dots \beta_q$ – koeficijenti za nezavisne promenljive koji se generalno određuju metodom najmanjih kvadrata; ε - residual.

U zavisnosti od problema koji se rešava, mogu se koristiti različite vrste regresije. Na osnovu načina na koji se unose nezavisne promenljive u jednačinu, razlikujemo tri glavne vrste [157]:

- Standardna MLR, odnosno istovremena – sve nezavisne promenljive se unose u jednačinu odjednom, tj. istovremeno. Ocenjuje se prediktivna sposobnost svake od unetih promenljivih na osnovu njenog doprinosa poboljšanju modela koji je

sastavljen od drugih nezavisnih promenljivih. Ova vrsta MLR, koja je primenjena i u ovom radu, je najčešće korišćena,

- Hijerarhijska MLR, odnosno sekvencijalna – nezavisne promenljive se unose u jednačinu jedna po jedna, ili po skupovima promenljivih, redosledom koji je određen na osnovu teorijskog modela zadatog od strane istraživača. Tokom procesa sekvencijalnog unošenja promenljivih u jednačinu, ocenjuje se doprinos svake nezavisno promenljive predikciji zavisne promenljive, ali tako da se uticaj prethodnih promenljivih ukloni. Nakon unošenja svih promenljivih, procenjuje se sposobnost finalnog modela, kao i relativan doprinos svakog skupa unetih promenljivih.
- Postepena MLR (*stepwise*) – nezavisne promenljive koje se unose u jednačinu, kao i njihov redosled, se određuje na osnovu određenih statističkih kriterijuma, npr. F-test. Postoje tri pristupa ove vrste MLR: promenljive se unose jedna po jedna (*forward selection*), unete su sve nezavisno promenljive u model, a zatim se jedna po jedna uklanjaju iz modela na osnovu uticaja na zavisnu promenljivu (*backward delation*) i postepena selekcija (*stepwise regression*) koja predstavlja kombinaciju prethodna dva pristupa.

Da bi se efikasno sprovedla MLR osnovne pretpostavke o podacima, u koje spadaju: normalnost, linearnost, homogenost varijanse, veličina uzorka i nekorelisane ulazne promenljive, trebaju biti ispunjene.

Uprkos brojnim literaturnim podacima koji su pokazali da MLR daje dobre rezultate u predviđanju, pristup zasnovan na regresiji nailazi na poteškoće u rešavanju problema kada se ulazne promenljive u modelu međusobno korelišu [178]. U slučajevima postojanja multikolinearnosti među nezavisnim promenljivima, teško je proceniti koja od datih promenljivih ima uticaja na zavisnu promenljivu i koliki je taj uticaj. PCA je jedna od statističkih tehnika koja se često koristi za uklanjanje multikolinearnosti među nezavisnim ulaznim promenljivima pre razvoja MLR modela. Naime, nakon sprovođenja PCA na originalnim ulaznim podacima, kao što je to opisano u poglavlju 4.3.1, dobijene glavne komponente (PC) se koriste kao nezavisne promenljive u regresionoj jednačini. Ovaj pristup, koji kombinuje MLR i PCA, se zove regresija glavnih komponenti (PCR - *Principal*

Component Regression) i zasniva se na odnosu između zavisne promenljive i PC generisanih od originalnih ulaznih promenljivih [151,179].

4.5. Statistički indikatori performansi modela

Performanse razvijenih modela u ovom radu su ocenjivane korišćenjem seta statističkih indikatora pokazatelja performansi. Statistička analiza performansi modela je uglavnom zasnovana na poređenju rezultata, vrednosti predviđenih modelom i stvarnih ili aktuelnih vrednosti. Set indikatora koji će se koristiti u analizi nije striktno definisan, ali je poželjno da taj broj indikatora bude što veći kako bi se stekla što šira slika o mogućnostima kreiranog modela.

Statistički indikatori za analizu performansi modela koji su korišćeni za procenu mogućnosti disperzionih modela [180,181], mogu se takođe primeniti i kod razvoja ANN modela za predviđanje emisije zagađujućih materija [182–184].

Najvažniji statistički indikatori koji su korišćeni u ovom radu za procenu performansi razvijenih modela su:

Greška korena srednjeg kvadrata (*RMSE - Root Mean Squared Error*) i srednja vrednost apsolutnih grešaka (*MAE - Mean Absolute Error*) predstavljaju meru rezidualne greške i daju globalnu sliku razlike između stvarnih i modelovanih vrednosti [162]. I *RMSE* i *MAE* su praktične za upotrebu s obzirom na to da se vrednost greške izražava u jedinicama promenljive koja je predmet predviđanja. Što su vrednosti ovih indikatora niže, to su performanse kreiranog modela bolje.

$$RMSE = \left[\overline{(C_p - C_o)^2} \right]^{1/2} \quad (23)$$

$$MAE = \frac{1}{n} \sum |C_p - C_o| \quad (24)$$

Koeficijent korelacije (r - *the correlation coefficient*) pokazuje stepen linearne korelacije između stvarnih i modelovanih vrednosti. Kreće se u rasponu -1 i +1 i što su vrednosti ovog indikatora bliže ovim vrednostima to je model precizniji.

$$r = \frac{\overline{(C_o - \bar{C}_o)(C_P - \bar{C}_P)}}{\sigma_{C_o}\sigma_{C_P}} \quad (25)$$

Indeks slaganja (IA - *the Index of Agreement*) je indikator stepena do kojeg je model za predviđanja oslobođen greške. Zadovoljavajuće vrednosti IA su one koje teže jedinici [185–187].

$$IA = 1 - \frac{\overline{(C_P - C_o)^2}}{\left[|C_P - \bar{C}_o| + |C_o - \bar{C}_o|\right]^2} \quad (26)$$

Faktor $FAI.25$ predstavlja procenat približno tačnih rezultata koje daje model u odnosu na stvarne vrednosti gde je odnos Predviđene/Stvarne vrednosti u opsegu [0.8 -1.25]. Za dobar model $FAI.25$ teži jedinici [188].

$$0.8 < \frac{C_P}{C_o} < 1.25 \quad (27)$$

Srednja apsolutna procentualna greška ($MAPE$ – *the Mean Absolute Percentage Error*), predstavlja meru tačnosti rezultata koje daje model u odnosu na stvarne vrednosti i izražava se u procentima. Ovaj indikator je pogodan za ocenu performansi modela za svaku državu pojedinačno koja je uključena u razvoj modela.

$$MAPE = 100 \frac{1}{k} \sum \frac{|C_o - C_P|}{C_o} \quad (28)$$

Indeks performansi modela (d_r - *the index of model performance*) je reformulisan IA i ukazuje na sumu magnituda razlika između modelom predviđenih i stvarnih devijacija u odnosu na stvarne srednje vrednosti [189].

$$d_r = 1 - \frac{\sum_{i=1}^n |C_{P_i} - C_{O_i}|}{2 \sum_{i=1}^n |C_{O_i} - \bar{C}_o|} \quad (29)$$

NSE (The Nash-Sutcliffe Efficiency) predstavlja normalizovan statistički pristup koji određuje relativnu veličinu rezidualne varijanse ("šum") u odnosu na varijansu stvarnih vrednosti ("informacija"). *NSE* pokazuje koliko dobro dijagram izmerenih - modelovanih vredosti fituje 1:1 liniju [190].

$$NSE = 1 - \frac{\sum_{i=1}^n (C_{O_i} - C_{P_i})^2}{\sum_{i=1}^n (C_{O_i} - \bar{C}_o)^2} \quad (30)$$

RSR (RMSE observation standard deviation ratio) predstavlja odnos *RMSE* i standardne devijacije stvarne vrednosti. Kao i kod *RMSE* niže vrednosti ovog indikatora pokazuju bolje performanse modela.

$$RSR = \frac{RMSE}{C_{O_{st.dev.}}} \quad (31)$$

Srednja kvadratna greška (*MSE – Mean Squared Error*) se izražava u jedinicama ispitivane promenljive i što je vrednost ove greške manja to su performanse modela bolje.

$$MSE = \frac{1}{n} \sum_{i=1}^n (C_{O_i} - C_{P_i})^2 \quad (32)$$

Koeficijent determinacije (R^2), čije vrednosti se kreću od 0 do 1, definiše udeo varijanse u zavisnoj promenljivoj koji je predviđen nezavisnom promenljivom. Što je vrednost koeficijenta bliža jedinici to su performanse kreiranog modela bolje.

$$R^2 = \frac{[\sum (C_p - \bar{C}_p)(C_o - \bar{C}_o)]^2}{\sum (C_o - \bar{C}_o)^2 \sum (C_p - \bar{C}_p)^2} \quad (33)$$

Rezultati i diskusija

5. Predviđanje emisije amonijaka u SAD i EU

Predviđanje emisije amonijaka, na nacionalnom nivou, na osnovu literaturnih podataka, je zasnovana na inventarima emisija i predviđana je primenom emisionih faktora, koje je teško i finansijski zahtevno odrediti. Emisioni faktori u mnogim slučajevima su uglavnom procenjene, a ne određene vrednosti, tako da predstavljaju i osnovne izvore nesigurnosti procenjene emisije amonijaka. U ovom poglavlju je opisan razvoj, optimizacija i evaluacija modela zasnovanog na ANN za predviđanje emisije amonijaka na nacionalnom nivou u SAD i državama EU, korišćenjem ekonomskih, industrijskih i parametara održivog razvoja, kao ulaznih promenljivih u modelu. Višeslojni perceptron (MLP) arhitektura je korišćena za razvoj modela. MLP je optimizovana primenom proba-i-greška procedure, koja uključuje izbor optimalnog broja neurona u skrivenom sloju, aktivacionu funkciju i BP algoritam. Analiza glavnih komponenti (PCA) je primenjena radi uklanjanja multikolinearnosti među ulaznim promenljivama. Dobijeni rezultati su pokazali da je primena PCA doprinela poboljšanim performansama ANN modela. Razvijeni ANN model je poređen sa modelom zasnovanom na regresiji glavnih komponenti (PCR). Na osnovu dobijenih rezultata može se zaključiti da se razvijeni ANN model može koristiti kao alternativni model u odnosu na modele zasnovane na inventarima emisija za predviđanje emisije amonijaka na nacionalnom nivou. Najvažnija prednost ovog modela je značajno manji broj ulaznih parametara za model, kao i njihova široka dostupnost, što je od posebnog značaja za države u razvoju kod kojih je primena konvencionalnih modela ograničena zbog nedostatka podataka [191].

5.1. Ulazni i izlazni podaci

Predviđanje emisije određenih zagađujućih materija u vazduh primenom pristupa zasnovanog na neuronskim mrežama i indikatorima održivog razvoja kao ulaznih parametara modela, dalo je veoma dobre rezultate kod predviđanja [52,192]. U skladu sa tim, kao ulazni parametri za razvoj MLP modela u ovom radu su odabrana četiri indikatora održivog razvoja za koje je procenjeno da doprinose emisiji amonijaka (Tabela 3.). Bruto domaći proizvod

(GDP – *Gross Domestic Product*) po stanovniku je široko dostupan ekonomski indikator, koji je izabran jer ukazuje na stepen razvoja za svaku pojedinačnu državu uključenu u istraživanje. Kako je emisija amonijaka dominantna iz sektora poljoprivrede, dva indikatora koji obuhvataju izvore emisije iz ovog sektora: broj grla stoke (NC – *number of cattle*) i potrošnja azotnih đubriva (CNF – *Consumption of Nitrogen Fertilizers*) su izabrana, jer doprinose najviše emisiji amonijaka u okviru poljoprivrednih aktivnosti. Sektor saobraćaja, predstavlja još jedan od izvora emisije koji značajno doprinosi emisiji amonijaka [193]. Četvrti ulazni parametar koji "pokriva" sektor saobraćaja kao izvora emisije amonijaka je indikator finalne potrošnje energije u saobraćaju (FECT – *Final Energy Consumption by Transport*), jer sumira potrošnju energije u svim vrstama saobraćaja [194].

Tabela 3. Lista ulaznih promenljivih za razvoj ANN modela i osnovni statistički pokazatelji

Ulazni indikator	Jedinica ^a	Srednja vrednost	Standardna devijacija	Minimalna vrednost	Maksimalna vrednost
GDP	_ ^b	0.99	0.46	0.143	1.76
NC	grla po st.	0.19	0.15	0.051	1.43
CNF	t po st.	0.03	0.05	0.006	0.51
FECT	toe po st. ^c	0.87	0.42	0.204	2.44

^anakon normalizacije

^bbez jedinice zbog normalizacije po proseku EU27

^ctona ekvivalenta nafte

Izabrani ulazni parametri su normalizovani po stanovniku i/ili po proseku vrednosti GDP EU27, kako bi mreža dala što bolje rezultate i kako bi se mogle porediti države različitih veličina. Ulazni i izlazni podaci za svaku Evropsku državu uključenu u ovom istraživanju su prikupljeni sa sajta evropske statističke baze podataka, Eurostata [195], dok su podaci za SAD preuzeti iz nacionalne statističke baze podataka za poljoprivredu (USDA), baze podataka za energetiku (USEIA) i agencije za zaštitu životne sredine (EPA) [196–198]. Statistička analiza izabranih ulaznih parametara za period 2000 – 2010. je prikazana u Tabeli 3., dok su koeficijenti korelacije i VIF prikazani u Tabeli 4.

Tabela 4. Pirsonov koeficijent korelacije između ulaznih promenljivih i odgovarajući faktor povećanja varijanse (VIF)

	Pirsonov koeficijent korelacije				VIF
	FECT	NC	GDP	CNF	
FECT	1.000				2.062
NC	0.493	1.000			2.585
GDP	0.716	0.642	1.000		2.917
CNF	0.046	0.478	0.049	1.000	1.519
NH ₃	0.467	0.759	0.459	0.633	

ANN model emisije amonijaka je obučavan, testiran i proveren korišćenjem podataka za SAD i 21 državu Evropske unije. Pri tome su podaci od 2000. do 2008. korišćeni za razvoj modela, dok su podaci od 2009. do 2010. korišćeni za validaciju modela. Statistička analiza podataka emisije amonijaka na nacionalnom nivou za svaku državu i različite setove podataka je prikazana u Tabeli 5.

Tabela 5. Statistički pokazatelji podataka emisije amonijaka za period od 2000. do 2010. godine

Država	NH ₃ emisija (kg po st.)			
	Sr. vr.	St. Dev.	Min.vr.	Maks.vr.
SAD	12.767	1.060	11.744	15.777
Bugarska	7.149	0.439	6.699	7.614
Češka	6.636	0.735	5.823	7.706
Danska	15.496	1.353	13.475	17.413
Estonija	7.563	0.285	7.287	8.001
Grčka	6.063	0.357	5.575	6.606
Španija	8.750	0.635	7.780	9.491
Francuska	12.033	0.880	9.978	13.244
Kipar	6.725	0.204	6.510	7.011
Letonija	7.259	0.309	6.846	7.711
Litvanija	9.432	1.012	8.473	10.735
Mađarska	7.040	0.583	6.530	8.036
Holandija	8.713	0.904	7.346	10.259
Austrija	7.736	0.207	7.457	8.080
Poljska	7.373	0.222	7.102	7.590
Portugal	5.071	0.513	4.502	5.964
Rumunija	8.713	0.735	7.516	9.396
Slovenija	8.833	0.270	8.530	9.246
Slovačka	4.756	0.225	4.502	4.998
Finska	7.151	0.165	6.887	7.403
Velika Britanija	5.132	0.362	4.585	5.665
Skup podataka				
Svi podaci	8.390	3.313	4.502	26.706
Trening podaci	8.611	3.534	4.630	26.706
Podaci za proveru	7.639	2.299	4.502	13.937

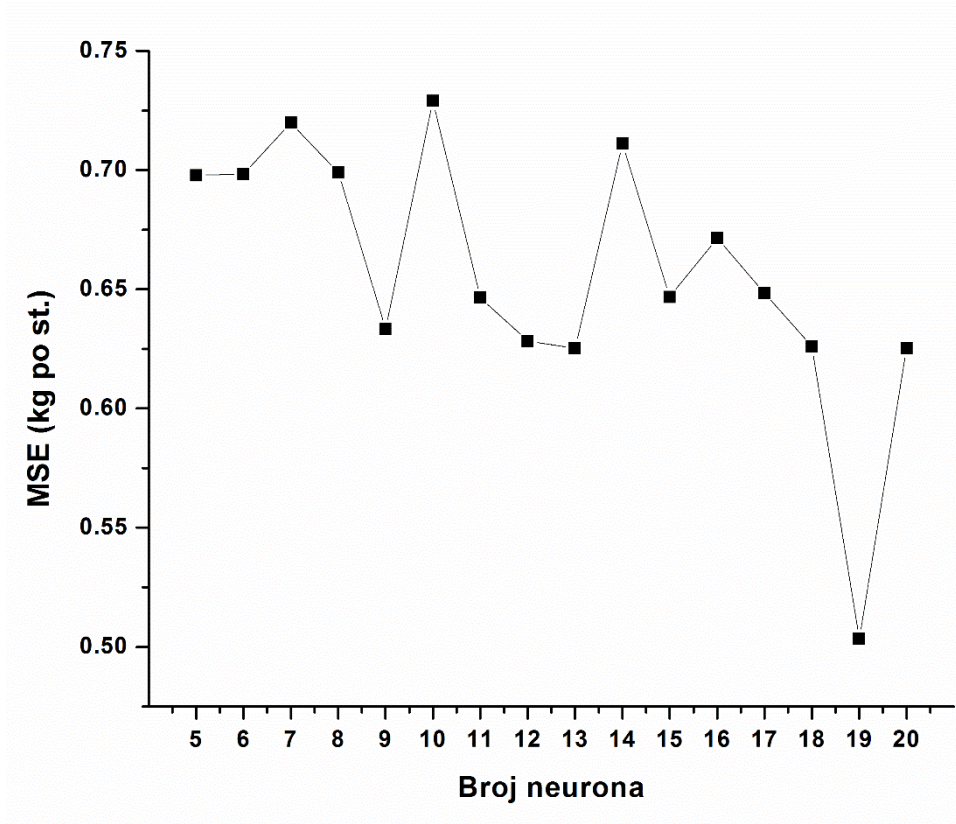
5.2. Rezultati optimizacije MLP modela

U cilju dobijanja MLP modela sa što boljim performansama, neophodno je izvršiti optimizaciju parametara obuke tokom procesa razvoja modela. Proces optimizacije, primenjen u ovom radu je zasnovan na principu proba-i-greška procedure, koja je sprovedena u tri koraka:

- i. Određivanje optimalnog broja neurona u skrivenom sloju;
- ii. Određivanje aktivacione funkcije i

iii. Određivanje BP algoritma koji daje model sa najboljim performansama.

U prvom koraku optimizacije MLP mreže, određivan je broj skrivenih neurona. Proces odabira optimalnog broja neurona je bio izveden, kao što je već napred rečeno, po principu probe i greške: formirano je 16 različitih $4 - H - 1$ MLP arhitektura, gde H predstavlja broj skrivenih neurona koji varira između 5 i 20. Razvijeni modeli su testirani i ocenjeni na osnovu MSE vrednosti. Dobijene MSE vrednosti za MLP modele dobijene tokom optimizacije broja neurona su prikazane na Slici 10. Utvrđeno je da je optimalan broj neurona u skrivenom sloju 19, tako da je finalna MLP arhitektura bila $4 - 19 - 1$.



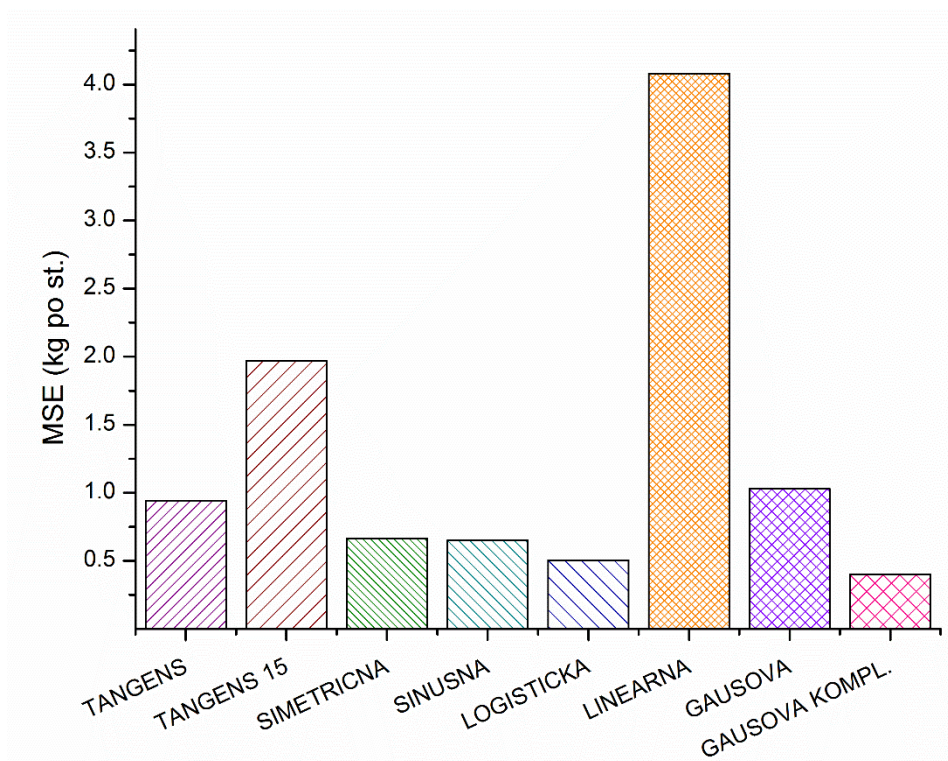
Slika 10. Optimizacija MLP mreže – određivanje optimalnog broja skrivenih neurona

Sledeći korak u procesu optimizacije je bio određivanje odgovarajuće aktivacione funkcije, između osam različitih varijanti koje su razmatrane (Tabela 6.). Kao i kod optimizacije broja skrivenih neurona, u cilju pronalaženja optimalne aktivacione funkcije, razvijeno je osam

modela sa različitim aktivacionim funkcijama, koji su testirani i ocenjeni na osnovu *MSE* vrednosti (Slika 11.)

Tabela 6. Testirane aktivacione funkcije u procesu optimizacije MLP mreže

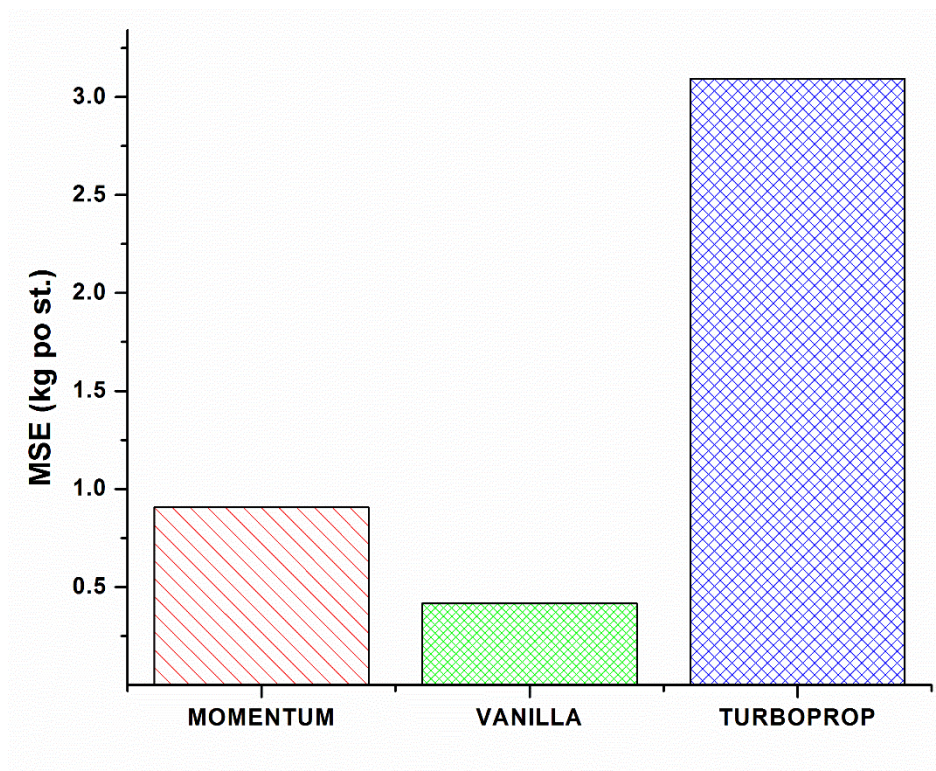
Aktivaciona funkcija	Jednačina
Linearna	$f(x) = x$
Logistička (Sigmoidna)	$f(x) = \frac{1}{1 + \exp(-x)}$
Simetrična logistička	$f(x) = \frac{2}{(1 + \exp(-x)) - 1}$
Sinusna	$f(x) = \sin(x)$
Tangens (hiperbolička tangens)	$f(x) = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)}$
Tangens15 (hiperbolična tangens 1.5)	$f(x) = \tanh(1.5x)$
Gausova	$f(x) = \exp(-x^2)$
Gausova komplementarna	$f(x) = 1 - \exp(-x^2)$



Slika 11. Optimizacija MLP modela – određivanje optimalne aktivacione funkcije

Kao što se može videti, MLP model sa Gausovom komplementarnom aktivacionom funkcijom daje najmanje *MSE* vrednosti, tj. pokazuje najbolje performanse.

Na kraju procesa optimizacije, određivan je najpogodniji BP algoritam za obuku mreže. U ovom radu su testirana tri BP algoritma, koji su ponuđeni od strane *Neuroshell* softvera koji je korišćen za razvoj MLP modela, *Vanilla*, *Momentum* i *TurboProp*. Kao što se može videti na Slici 12., najbolje performanse pokazuje MLP model treniran *Vanilla* (V) algoritmom.



Slika 12. Optimizacija MLP modela – određivanje pogodnog BP algoritma

Na kraju procesa optimizacije MLP modela, može se zaključiti da je najbolje performanse dao model sa 19 neurona u skrivenom sloju, Gausovom komplementarnom aktivacionom funkcijom i *Vanilla* BP algoritmom.

5.3. Rezultati MLP modela kreiranih nakon primene analize glavnih komponenti

Dodatno poboljšanje performansi MLP modela je postignuto primenom analize glavnih komponenti na ulaznim podacima. Pre primene analize glavnih komponenti, veoma je značajno utvrditi da li dati podaci ispunjavaju uslove za sprovođenje analize glavnih komponenti. Primenljivost PCA je testirana KMO testom pogodnosti podataka i Bartletovim testom sferičnosti (Tabela 7.). Dobijena vrednost KMO - 0.618, kao i vrednost Bartletovog testa sferičnosti ($p < 0.0001$), su pokazali da je primena PCA na datim ulaznim podacima opravdana. Vrednosti zajedničkog varijabiliteta – komunalitet (*communality*), koje su

rangirane između 0 i 1 i koje pokazuju deo varijanse za svaku promenljivu koji je objašnjen glavnim komponentama, je takođe prikazan u Tabeli 7. Niske vrednosti zajedničkog varijabiliteta ukazuju na to da se određena promenljiva ne "uklapa" dobro u svoju glavnu komponentu. Kao što se može videti, u ovom slučaju, sve promenljive imaju vrednost zajedničkog varijabiliteta veću od 0.8.

Tabela 7. Rezime primenjene PCA: komunalitet, KMO vrednost i Bartletov test

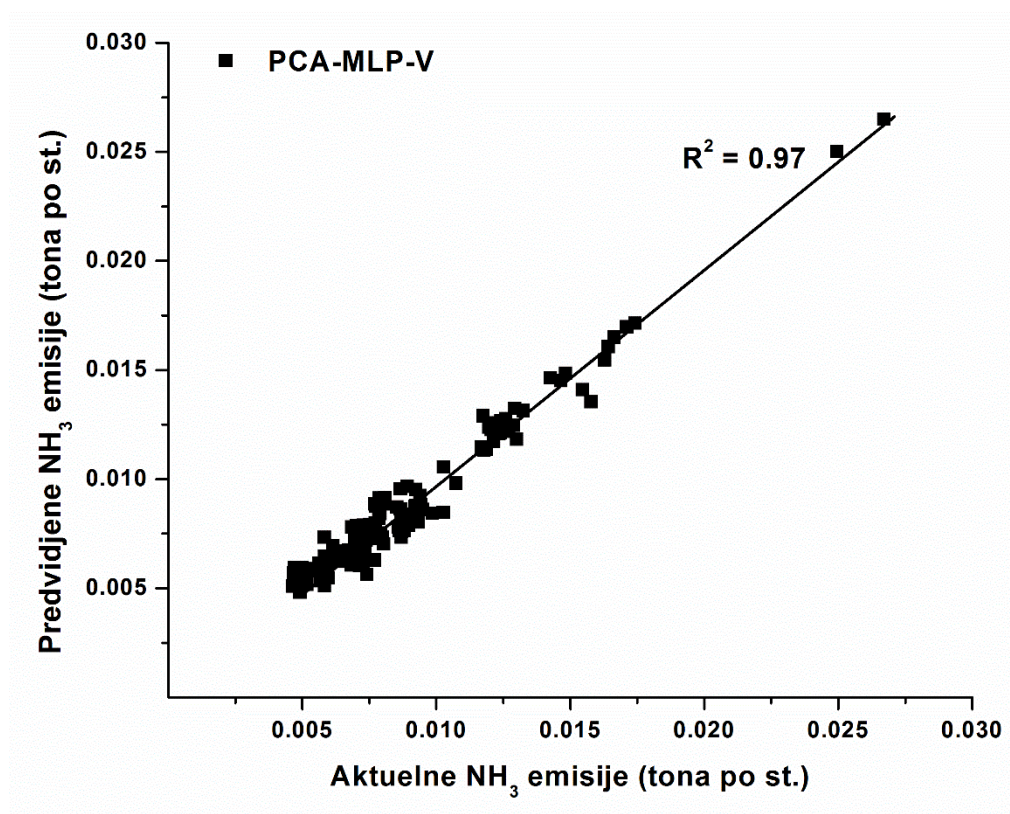
Kaiser-Meyer-Olkin (KMO)		0.618
Bartletov test sferičnosti	Chi-kvadrat test	225.911
	Stepeni slobode	6
	Značajnost. (<i>p</i>)	< 0.0001
Ulazna promenljiva	Komunalitet	
FECT	0.905	
GDP	0.895	
CNF	1.000	
NC	0.999	

Transformacija glavnih komponenti je vršena *Direct Oblimin* metodom. Faktorske težine svih promenljivih nakon rotacije glavnih komponenti, strukturna matrica koja daje korelaciju između promenljivih i glavnih komponenti, svojstvene vrednosti, kumulativna varijansa i ukupna varijansa objašnjena izdvojenim glavnim komponentama su prikazane u Tabeli 8. Na osnovu analize su izdvojene tri glavne komponente koje objašnjavaju ukupno 94.97% ukupne varijanse.

Tabela 8. Faktorske težine, strukturna matrica, svojstvene vrednosti, varijansa po komponenti, kao i ukupna varijansa objašnjena izdvojenim glavnim komponentama

Ulazna promenljiva	Komponenta					
	Faktorske težine			Strukturna matrica		
	PC1	PC1	PC3	PC1	PC2	PC3
FECT	0.970	-0.012	-0.041	0.951	0.073	0.392
GDP	0.921	0.014	0.049	0.945	0.107	0.464
CNF	0.000	1.000	-0.001	0.093	1.000	0.133
NC	0.002	-0.001	0.999	0.450	0.134	1.000
Svojstvene vrednosti				2.151	0.984	0.664
Varijansa po komponenti (%)				53.76	24.61	16.60
Kumulativna varijansa (%)				53.76	78.37	94.97

Izdvojene glavne komponente (PC) su zatim korišćene kao ulazne promenljive za MLP model razvijen korišćenjem PCA – transformisanih ulaznih promenljivih, tj. kreiran je PCA – MLP – V model, sa istim trening parametrima kao i kreirani MLP – V model sa originalnim podacima. Dobijena vrednost *MSE* za kreirani PCA – MLP – V model je 0.45 (kg po st.), dok je komparacija aktuelnih (izmerenih) i PCA – MLP – V modelom predviđenih vrednosti emisije amonijaka prikazana na Slici 13.



Slika 13. Poređenje aktuelnih emisija amonijaka i PCA – MLP – V modelom predviđenih vrednosti

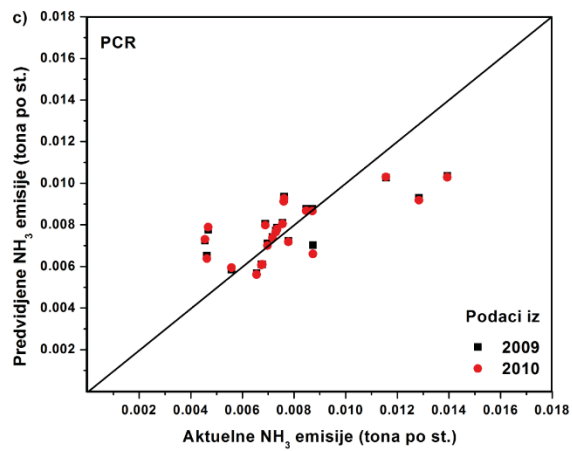
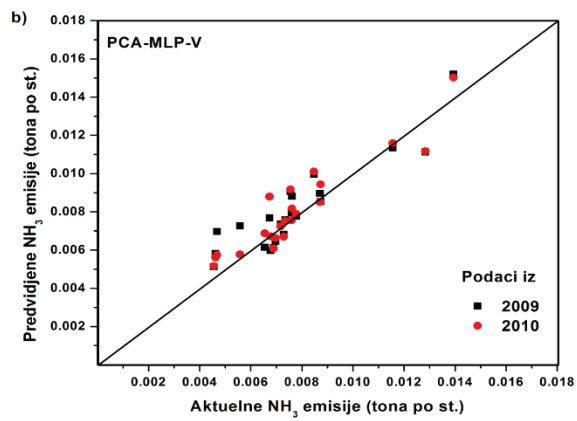
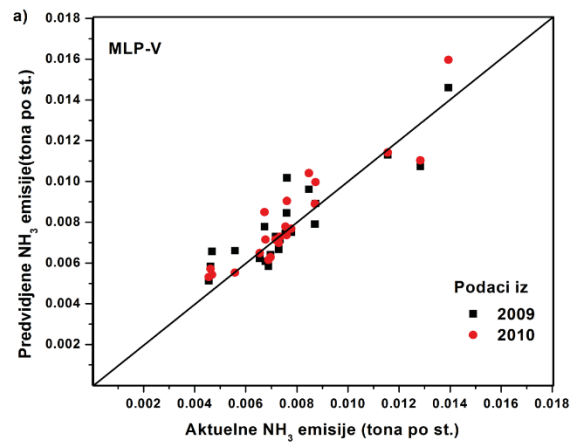
5.4. Rezultati validacije MLP modela

Validacija kreiranih MLP modela (MLP – V i PCA – MLP – V) je vršena sa potpuno novim podacima (podaci za period 2009 – 2010.) za sve države koje su uključene u studiji. Modeli su proveravani za rezultate koje daju za jednogodišnja i dvogodišnja predviđanja.

Za komparaciju rezultata MLP modela, razvijen je PCR model (34) sa istim glavnim komponentama, dobijenim nakon analize glavnih komponenti, koje su korišćene kao ulazi pri razvoju PCA – MLP – V modela.

$$NH_3 = 7.782 \cdot PC1 + 0.264 \cdot PC2 + 15.382 \cdot PC3 + 5.124 \quad (34)$$

Poređenje aktuelnih i predviđenih vrednosti za MLP – V, PCA – MLP – V i PCR modele su prikazani na Slici 14.



Slika 14. Rezultati validacije modela. Dijagram aktuelna – predviđena emisija amonijaka za a) MLP – V; b) PCA – MLP – V; c) PCR

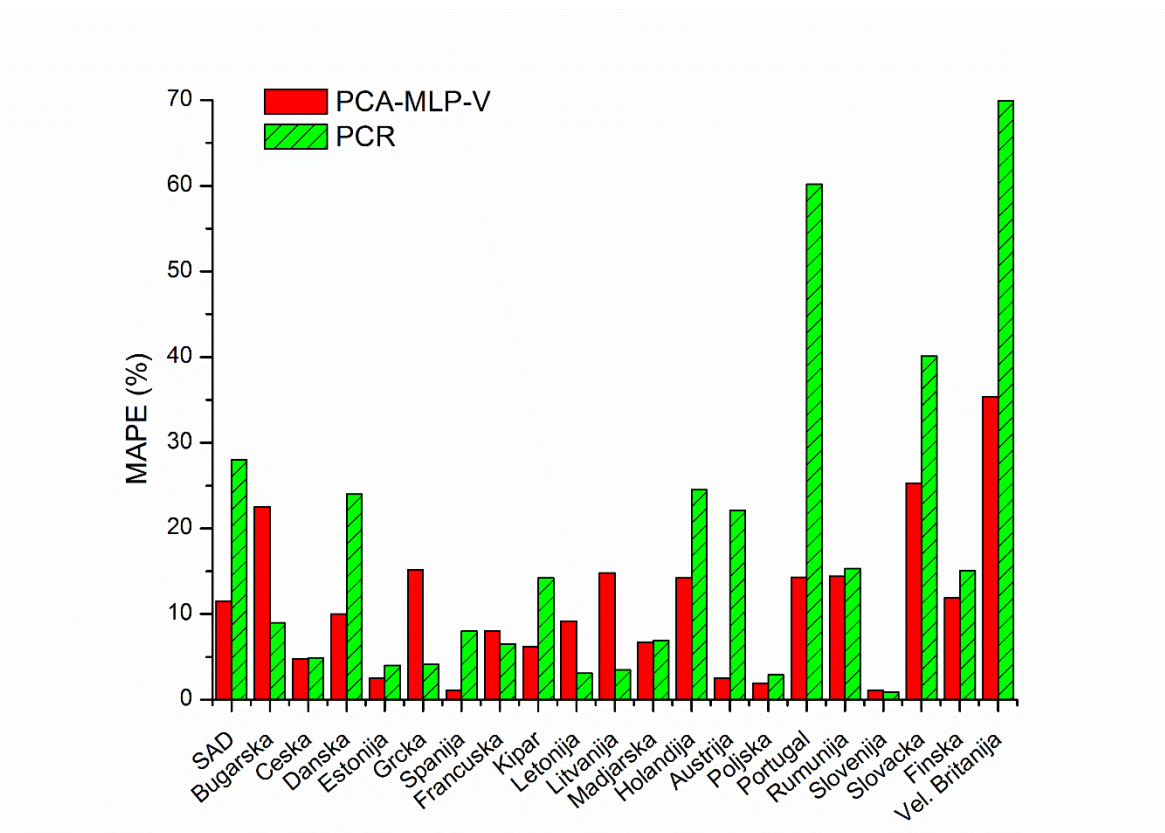
Rezultati validacije su pokazali da i MLP - V i PCA – MLP - V modeli pokazuju značajno bolje rezultate u predviđanju u odnosu na linearni PCR model, na osnovu vrednosti svih statističkih pokazatelja performansi modela, koji su prikazani u Tabeli 9.

Tabela 9. Vrednosti statističkih pokazatelja performansi modela (faza validacije modela)

Indikatori performansi	MLP-V model		PCA-MLP-V model		PCR model	
	2009	2009-2010	2009	2009-2010	2009	2009-2010
d_r	0.74	0.73	0.76	0.76	0.62	0.62
NSE	0.73	0.70	0.77	0.76	0.41	0.42
$RMSE$ [kg po st.]	1.07	1.13	1.00	1.02	1.69	1.66
MAE [kg po st.]	0.86	0.90	0.80	0.80	1.29	1.24
RSR	0.46	0.48	0.43	0.43	0.72	0.71
$FA\ 1.25$ [%]	86	86	86	86	74	76

Takođe, dobijeni rezultati pokazuju da PCA – MLP – V model ima malo bolje performanse u odnosu na MLP – V model, kako za jednogodišnje, tako i za dvogodišnje predviđanje. Primena analize glavnih komponenti na originalni set ulaznih podataka, rezultirala je u smanjenju $RMSE$ i MAE vrednosti za oko 10%, dok vrednost $FAI.25$ pokazuje iste vrednosti od 86%, što znači da na validacionom setu podataka PCA – MLP – V model predviđa sa relativnom greškom većom od 25% za 14% slučajeva.

Vrednosti srednje apsolutne procentualne greške ($MAPE$), za predviđanja emisije amonijaka postignuta sa PCA – MLP – V i PCR modelima na validacionom setu podataka (2009 - 2010.) za sve države pojedinačno su prikazane na Slici 15. Kao što se može videti, PCA – MLP – V model pokazuje bolje rezultate predviđanja u odnosu na PCR model za celokupni validacioni set podataka ($MAPE = 11\%$), za SAD i za većinu evropskih država. Samo u slučaju Bugarske, Slovačke i Velike Britanije, PCA – MLP – V model daje predviđanja sa $MAPE$ vrednostima većim u odnosu na $MAPE$ vrednosti PCR modela, iznad 18%. U slučaju SAD, predviđanje postignuto PCA – MLP – V modelom je sa relativnom greškom od 12%, što je oko 2.5 puta preciznije u odnosu na predviđanje postignuto linearnim PCR modelom.



Slika 15. MAPE vrednosti predviđanja PCA- MLP – V i PCR modela za pojedinačne države

5.5. Rezultati analize osetljivosti (značajnosti) ulaznih promenljivih

Nakon procesa obuke MLP mreže, postoji mogućnost ispitivanja značajnosti svake ulazne promenljive za razvijeni model. U slučaju troslojne MLP, kod koje se koristi BP algoritam za obučavanje mreže, analiza značajnosti se može sprovesti određivanjem faktora doprinosa za svaku od ulaznih promenljivih. Faktor doprinosa (*CF - contribution factor*) predstavlja sumu apsolutnih vrednosti težinskih koeficijenata, određenih tokom treninga mreže, za određenu promenljivu [142]. Rezultati ove analize su pokazali da ulazni parametar "broj grla stoke" (NC) ima najveći uticaj na izlaz tj. na emisiju amonijaka ($CF=14.05$), zatim po značajnosti sledi promenljiva GDP ($CF=10.52$). Ulazne promenljive FECT ($CF=5.67$) i potrošnja azotnih veštačkih đubriva – CNF ($CF=5.16$) imaju manje uticaja na izlaz modela.

Visoka vrednost faktora doprinosa za NC potvrđuje činjenicu da je sektor poljoprivrede jedan od dominantnih sektora izvora emisije amonijaka.

6. Predviđanje emisije nemetanskih isparljivih organskih jedinjenja u EU i Kini

U ovom poglavlju je opisano kako je pristup zasnovan na neuronskim mrežama prethodno primenjen kod predviđanja emisije amonijaka, primenjen za predviđanje godišnje emisije na nacionalnom nivou nemetanskih isparljivih organskih jedinjenja (NMVOCs) u Kini za period 2005 - 2011. godine. NMVOCs emisija u Kini je predviđena ANN modelom koji je razvijen sa dostupnim podacima za devet EU država, čije emisije NMVOCs po glavi stanovnika približno odgovaraju emisijama NMVOCs u Kini, za period 2004 - 2012. Izbor unapred, tj. *forward* metoda je korišćena radi procene i poređenja značajnosti potencijalnih ulaznih promenljivih kako za devet EU država, tako i za Kinu. Finalni ANN model koji je razvijen korišćenjem samo pet ulaznih promenljivih, pokazao je dobre performanse u predviđanju, kako za EU države, čiji su podaci korišćeni za obučavanje mreže, tako i za Kinu, čiji su podaci bili potpuno novi set podataka koji je prezentovan razvijenom ANN modelu [199].

6.1. Prikupljanje i selekcija podataka

Kao što je već pomenuto, izbor odgovarajućih promenljivih u odnosu na potencijalno odabrane, je od velikog značaja za funkcionisanje i performanse razvijenog ANN modela. Ukoliko u model nisu uključene značajne promenljive, to može doprineti lošijem predviđanju. Međutim, ukoliko su u model uključene suvišne promenljive, to može smanjiti računsku efikasnost modela, proces obuke traje duže, i javlja se niz drugih problema koji smanjuju njegovu efikasnost [200].

Selekcija ulaznih promenljivih

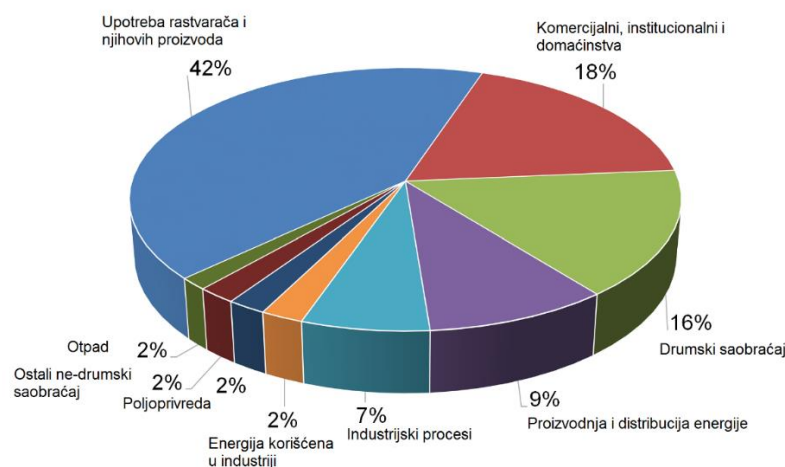
Selekcija ulaznih promenljivih je u ovom slučaju vršena "forward" pristupom, kojim se selektuje svaka od potencijalnih promenljivih pojedinačno. Ovaj pristup za selekciju ulaznih promenljivih spada u metode kod kojih je selekcija zasnovana na modelu (*model based*). *Forward* selekcija je linearna inkrementalna strategija odabira, kojom se biraju individualno potencijalne ulazne promenljive, jedna po jedna. Proces selekcije počinje sa kreiranjem ANN

modela za svaku n pojedinačnu potencijalnu ulaznu promenljivu [108]. Na kraju ovog procesa sa za kreiranje finalnog ANN modela koriste one ulazne promenljive čiji su n pojedinačni ANN modeli dali najbolje rezultate.

Ovaj pristup za selekciju ulaznih promenljivih pokazao se pogodnim za poređenje značaja ulaza za slučaj kada je model razvijen u jednoj državi, ili definisanoj grupi država, a onda zatim treba da se primeni u drugim državama.

Dostupni podaci o ulaznim i izlaznim promenljivama

Doprinos svakog pojedinačnog sektora izvora NMVOCs emisije u Evropi dobijenih na osnovu nacionalnih izveštaja, prema Konvenciji o prekograničnom zagađenju vazduha na velikim udaljenostima (LRTAP Convention), su prikazani na Slici 16. [201].



Slika 16. Doprinos različitih sektora emisiji NMVOCs u EU državama

Indikatori održivog razvoja, ekonomski, industrijski i poljoprivredni indikatori su i u ovom slučaju primenjeni za razvoj modela za predviđanje emisije NMVOCs zasnovanom na neuronskim mrežama. Sa ciljem da se obuhvate ključni izvori emisije, sledeće promenljive su odabrane kao potencijalni ulazi:

1. Električna energija dobijena iz obnovljivih izvora (EGRS - *Electricity Generated from Renewable Sources*),

2. Potrošnja električne energije u domaćinstvima (ECH - *Electricity Consumption by Households*),
3. Površina poljoprivrednog zemljišta (AL - *Agricultural Land*),
4. Proizvodnja iz hemijske industrije (CS - *Chemical Shipments*),
5. Proizvodnja vozila (MVP - *Motor Vehicle Production*),
6. Ukupna potrošnja naftnih derivata (CTPP - *Consumption Total Petroleum Products*),
7. Ukupna potrošnja energije u transportu (FECT - *Final Energy Consumption in Transport*),
8. Bruto domaći proizvod (GDP - *Gross Domestic Product*),
9. Generisanje komunalnog otpada (MWG - *Municipal Waste Generation*),
10. Bruto unutrašnja potrošnja energije (GIEC - *Gross Inland Energy Consumption*).

Pored indikatora nivoa ekonomske aktivnosti, GDP je izabran i kao indikator koji kvantifikuje upotrebu kontrolnih mera i tehnologija koje se primenjuju radi smanjenja emisije, a koje mogu značajno da utiču na ukupno smanjenje emisije zagađujućih materija u vazduh. EU države koje su bile uključene kod razvoja modela, birane su na osnovu dva kriterijuma: jedan se odnosi na to da je ukupna emisija NMVOC po glavi stanovnika približna emisiji u Kini i drugi je da su dostupni podaci za sve potencijalne ulazne promenljive. Lista devet EU država koje su uključene u ovom istraživanju, koji su dobijeni iz baze podataka EDGAR projekta [202], kao i osnovni statistički podaci o emisijama NMVOC su prikazani u Tabeli 11., dok su osnovni statistički pokazatelji potencijalnih ulaznih promenljivih prikazani u Tabeli 10.

Tabela 10. Statistička analiza potencijalnih ulaznih promenljivih za odabrane EU države

Ulazna promenljiva ^a	Jedinica nakon normalizacije	Statistička analiza		
		Srednja vrednost	Minimalna vrednost	Maksimalna vrednost
EGRS	%	17.9	1.7	60.0
ECH	toe po st.	0.18	0.10	0.41
AL	ha po st.	0.37	0.11	1.07
CS ^b	\$ po st.	3514	1140	14942
MVP ^c	broj ^d	39.74	3.33	88.70
CTPP	toe po st.	1.59	1.01	2.36
FECT	toe po st.	0.89	0.66	1.35
GDP	€ po st.	29627	20200	41000
MWG	kg po st.	554	455	792
GIEC	toe po st.	4.13	2.72	5.78

^a Eurostat [195]

^b American Chemistry Council [203]

^c The International Organization of Motor Vehicle Manufacturers [204]

^d na 1000 stanovnika

Radi postizanja boljih performansi modela, kao i radi poređenja različitih država, ulazni i izlazni podaci su normalizovani po stanovniku. NMVOC model je obučavan sa podacima devet EU država za period 2004 – 2011., dok je validacija modela vršena podacima iz 2012. godine. Takođe, 20% slučajno izabranih podataka, iz seta podataka za obuku mreže, je izdvojeno radi unutrašnjeg testiranja modela sa ciljem sprečavanja pretreniranosti mreže. Nakon što je kreirani NMVOC model proveren validacionim podacima EU država, pristupilo se primeni kreiranog modela na potpuno nove podatke za Kinu, za period 2005 – 2011. godine (Tabela 12.).

Tabela 11. Statistička analiza emisije NMVOC (kg po st.) za izabrane EU države za period 2004 – 2012

Država	NMVOC emisija		
	Srednja vrednost	Minimalna vrednost	Maksimalna vrednost
Francuska	15.53	11.04	22.02
Nemačka	12.72	11.23	14.07
Italija	18.13	14.49	22.09
Velika Britanija	16.51	13.30	20.81
Belgija	11.93	9.48	14.43
Irska	11.85	9.57	14.63
Holandija	9.81	8.88	10.79
Španija	16.79	13.44	21.18
Švedska	21.13	19.85	22.72

Tabela 12. Podaci za Kinu

Ulaz/Izlaz ^a	Jedinica	Godina						
		2005	2006	2007	2008	2009	2010	2011
EGRS ^b	%	18.1	17.4	17.0	19.5	19.5	21.2	19.0
ECH	toe po st.	0.019	0.022	0.027	0.029	0.031	0.033	0.036
AL	ha po st.	0.100	0.099	0.099	0.092	0.091	0.090	0.090
CS ^c	\$ po st.	225	281	380	508	564	732	970
MVP ^d	broj ^e	4.4	5.5	6.7	7.0	10.4	13.7	13.7
CTPP	toe po st.	0.27	0.29	0.29	0.30	0.33	0.36	0.37
FECT	toe po st.	0.08	0.10	0.11	0.12	0.12	0.14	0.15
GDP ^f	€ po st.	1380	1650	2114	2722	2989	3535	4343
MWG ^g	kg po st.	119	113	115	117	118	122	125
GIEC	toe po st.	1.27	1.38	1.49	1.54	1.61	1.70	1.81
NMVOC ^h	kg po st.	14.5	15.3	15.6	15.6	16.3	17.1	-

^a National Bureau of Statistics of China [205]

^b US Energy Information Administration [197]

^c American Chemistry Council [203]

^d The International Organization of Motor Vehicle Manufacturers [204]

^e na 1000 stanovnika

^f The World Bank [206]

^g The Organisation for Economic Co-operation and Development [7]

^h Zhao i saradnici [207]

6.2. Razvoj i rezultati razvijenog MLP modela

Kao što je već ranije rečeno, troslojna neuronska mreža sa kretanjem signala napred koja je obučavana nekim od BP algoritama je najčešće korišćena arhitektura koja je pokazala veoma dobre rezultate na širokom spektru problema [86]. Zbog dobrih performansi koje je pokazao MLP model za predviđanje emisije amonijaka i u ovom slučaju je za razvoj ANN modela korišćena troslojna BPNN arhitektura.

Za skaliranje ulaznih podataka izabrana je linearna funkcija, dok je sigmoidna (logistička) funkcija izabrana kao aktivaciona funkcija. Obučavanje MLP mreže je vršeno BP algoritmom sa konjugovanim gradijentom, sa vrednostima parametara stope obučavanja i momentuma 0.1, dok su vrednosti inicijalnih težina 0.3. Kao što je ranije pomenuto, broj neurona u ulaznom i izlaznom sloju je jednak broju ulaznih i izlaznih promenljivih, dok je broj neurona u skrivenom sloju, u ovom slučaju, određen primenom formule (9).

Najpre se pristupilo određivanju najznačajnijih ulaznih promenljivih, od prvobitno određenih potencijalnih. Performanse kreiranih modela su određivane primenom statističkog parametra, koeficijenta determinacije R^2 . Dobijene vrednosti R^2 za kreirane ANN modele za svaku potencijalnu ulaznu promenljivu pojedinačno, za validacioni set podataka država EU i njihova primena na podatke za Kinu, su prikazane u Tabeli 13., dok su dijagrami aktuelne – predviđene vrednosti za najznačajnije ulazne promenljive prikazani na Slici 17. Kao što se može videti iz Tabele 13. u slučaju EU država, svi potencijalni ulazi, osim FECT, GDP, MWG i GIEC, se mogu smatrati značajnim za predviđanje NMVOC emisije. Pomenuti ulazi su manje značajni, verovatno zbog sličnog nivoa razvijenosti, kao i ekoloških standarda koji su primenjeni u izabranim EU državama.

S druge strane, primena podataka za Kinu na kreiranim modelima je pokazala da su samo dva kreirana modela, sa EGRS i MVP ulaznim parametrima, dovoljno osetljivi da pokažu značajne razlike u rezultatima kao posledicu varijacije ulaza za različite godine (Slika 17. g i h).

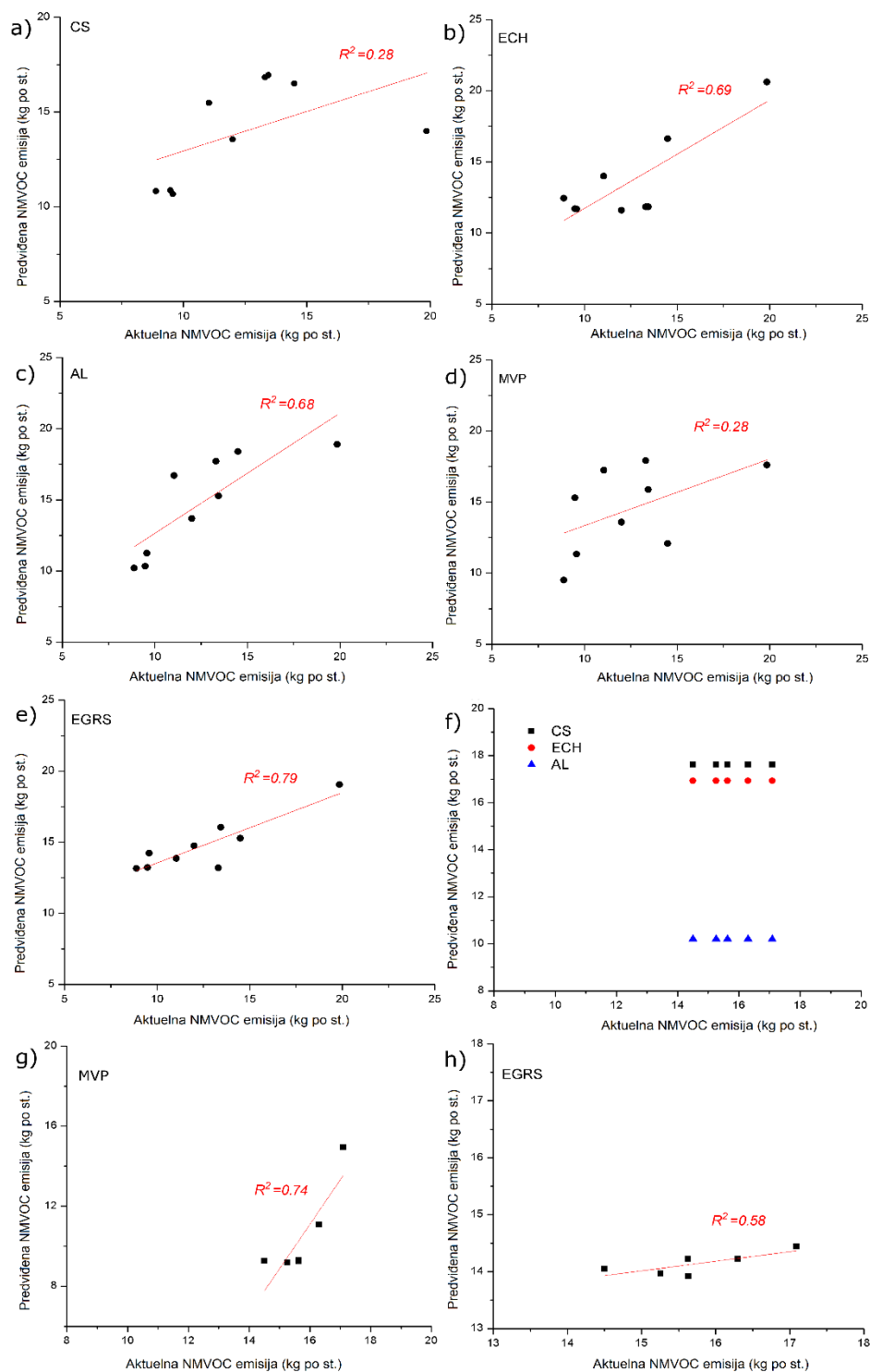
Tabela 13. R^2 vrednosti za kreirane ANN modele za svaku ulaznu promenljivu pojedinačno

Ulaz	EU države ^a	Kina ^b
	R^2	R^2
EGRS	0.79	0.58
ECH	0.69	-
AL	0.68	-
CS	0.28	-
MVP	0.28	0.74
CTPP	0.26	-
FECT	0.05	-
GDP	0.01	-
MWG	0.00	-
GIEC	0.00	-

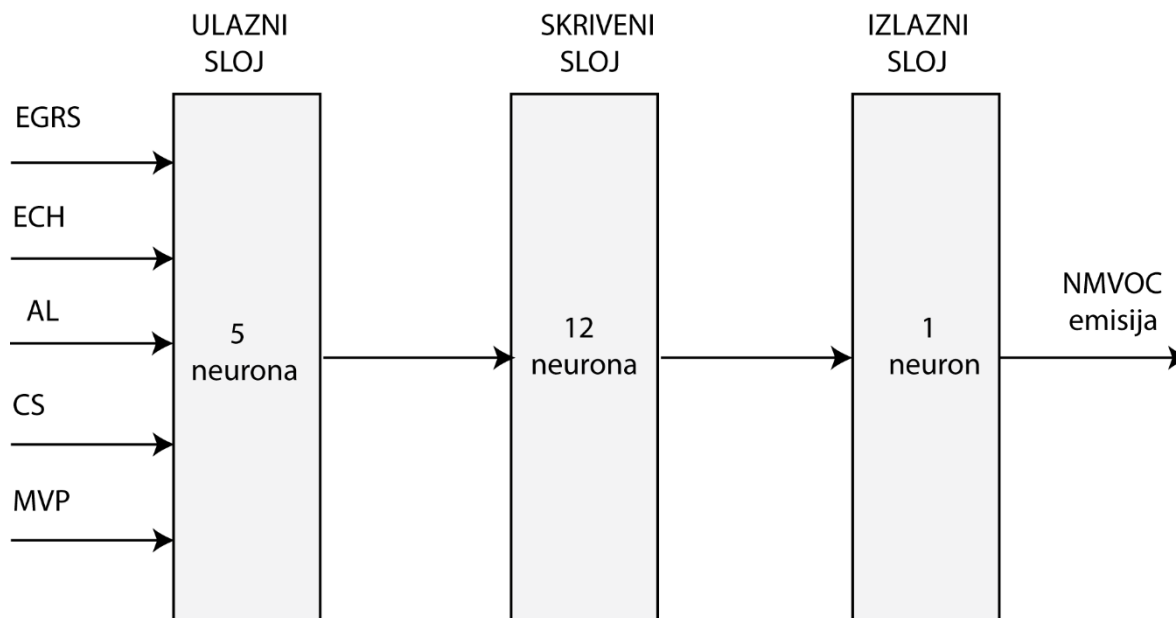
^a kada je model primenjen na podatke iz 2012.godine

^b kada je model primenjen na podatke 2005 – 2011.

Shodno tome, finalni ANN model je kreiran korišćenjem dva ulazna parametra koji su se pokazali značajnim za Kinu (EGRS i MVP), ali su takođe uključeni i ulazni parametri koji su se pokazali značajnim za EU države (ECH, AL, CS), kako bi se poboljšale performanse modela za EU države, a time implicitno i za Kinu. Šematski prikaz troslojnog MLP modela kreiranog sa pet ulaznih promenljivih za predviđanje NMVOC emisije je prikazan na Slici 18.



Slika 17. Poređenje aktuelne i predviđene vrednosti emisije NMVOC razvijenih ANN modela: **a-e** testiranje sa EU podacima iz 2012. godine i **f-h** primena na podacima za Kinu 2005 – 2010.

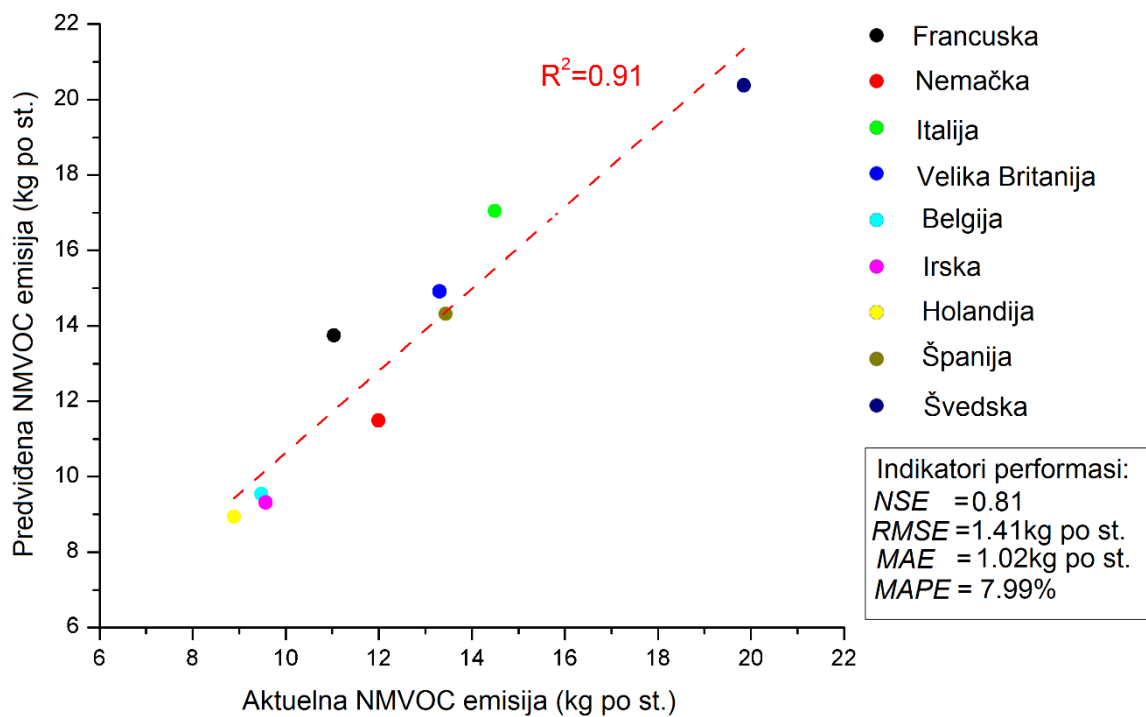


Slika 18. Šematski prikaz NMVOC MLP modela

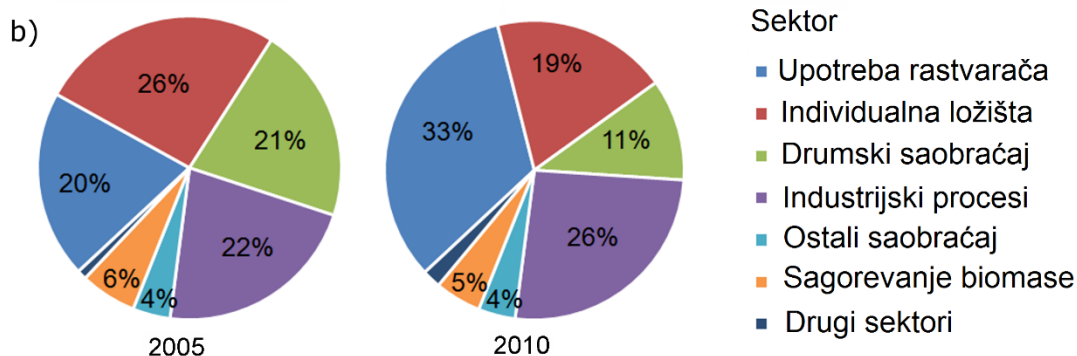
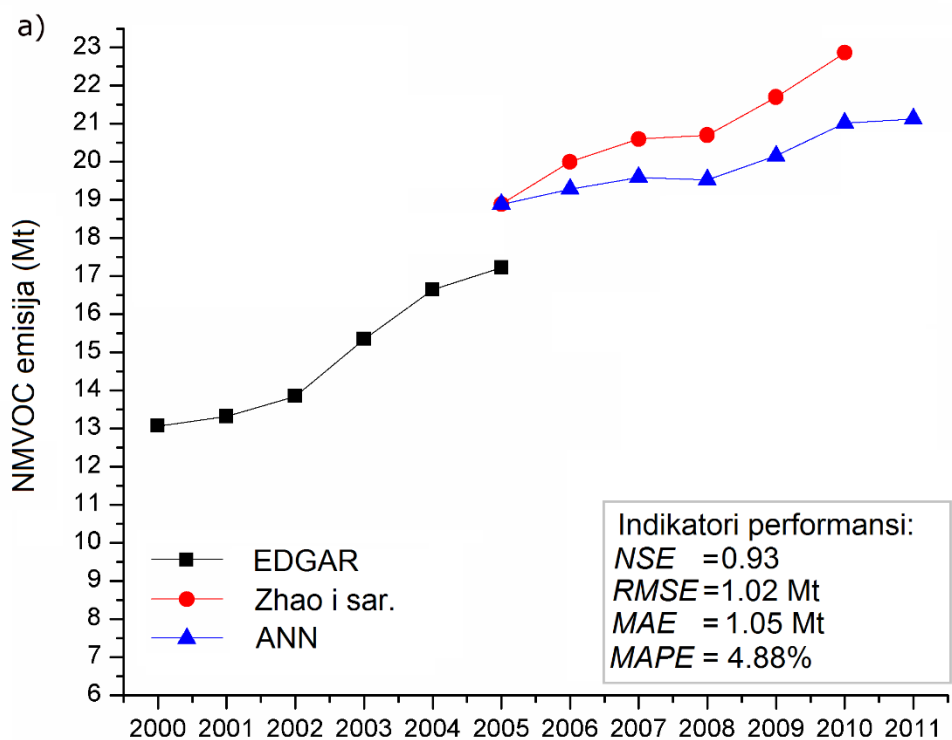
Poređenje aktuelne i modelom predviđene NMVOC emisije za EU države primenjene na potpuno novim podacima za 2012. godinu je prikazan na Slici 19. Kao što se može videti, dobijeni rezultati se mogu smatrati kao veoma dobri, sa R^2 vrednošću 0.91 i $MAPE$ vrednošću od 8%.

Predviđanje emisije NMVOC za Kinu kreiranim MLP modelom za period 2005 – 2011., zajedno sa emisijom NMVOC za period 2000 – 2005. dobijene iz EDGAR baze podataka emisije zagađujućih materija [202] i emisijom NMVOC za period 2005 – 2010. procenjenu od strane Zhao i sar. [207], je predstavljeno na Slici 20. Veća $MAPE$ vrednost dobijena za EU države ($MAPE$ - 8%) u poređenju sa $MAPE$ vrednošću dobijenom za Kinu ($MAPE$ - 4.88%) je verovatno posledica veće heterogenosti u EU podacima.

Kao što se može videti na Slici 20., NMVOC emisija za 2005. godinu dobijena od strane EDGAR i Zhao i sar. [207], se razlikuje oko 10%. Razlika u procenjenoj emisiji između ova dva inventara emisija može se pripisati različitim emisionim faktorima korišćenim pri proračunu, ili razlike u korišćenim nacionalnim statističkim podacima koji se mogu razlikovati u odnosu na internacionalne statističke baze podataka [59].



Slika 19. Poređenje aktuelne i modelom predviđene NMVOC emisije za EU države



Slika 20. Poređenje a) aktuelne i modelom predviđene NMVOC emisije za Kinu i b) doprinos pojedinih sektora emisiji NMVOC u Kini [207]

S druge strane, predviđenje emisije NMVOC dobijeno primenom ANN modela, kao što se može videti značajno korespondira sa procenom emisije Zhao i sar. [207], sa *MAPE* vrednošću od 4.88%. Razlika između ove dve procene se ogleda u tome što je predviđanje emisije ANN modelom pokazalo sporiji rast emisije u 2006. i 2007. godini. Naime, Zhao i sar. [207] procenjuju rast emisije NMVOC od 1.1 i 0.6 Mt u 2006. i 2007. godini, dok ANN model predviđa rast emisije od samo 0.4 i 0.3 Mt tokom ove dve godine. Stagnacija emisije

NMVOC u 2008. godini je predviđena i od strane jednog i drugog modela, kao i progresivno povećanje emisije tokom 2009. i 2010. godine. Stagnacija emisije u Kini u 2008. godini je verovatno uslovljena povećanjem proizvodnje električne energije iz obnovljivih izvora od 2.5% u poređenju sa 2007. godinom (Tabela 12.). Međutim, uprkos razlikama u procenjenoj emisiji za neke od godina, trend promene emisije tokom godina je generalno konzistentan između ova dva pristupa (ANN i Zhao i aut.). Takođe, kreiranim ANN modelom u ovom slučaju je predviđena slična stagnacija NMVOC emisije u 2011. godini kao i u 2008. godini.

Za period za koji je vršeno modelovanje NMVOC emisije, sektor "upotreba rastvarača" je postao dominantni izvor NMVOC emisije u Kini, sa porastom od 20% na 33% udela u ukupnoj emisiji (Slika 20.) [207]. S obzirom na to da je neizvesnost procenjene emisije NMVOC iz ovog sektora visoka, kreće se oko $\pm 30\%$ [208], sa povećanjem udela ovog sektora u ukupnoj emisiji, može se pretpostaviti da će se neizvesnost procenjene emisije u kasnijim godinama takođe povećavati, što je dovelo do razlike u proceni između pristupa zasnovanog na neuronskim mrežama i inventarskom pristupu predloženom od strane Zhao i sar. [207] (Slika 20.).

7. Predviđanje emisije metana u EU

U ovom poglavlju je opisan razvoj i rezultati modela zasnovanih na neuronskim mrežama za predviđanje emisije metana (CH₄) na nacionalnom nivou. Kao i u slučaju amonijaka i nemetanskih isparljivih organskih jedinjenja, kao ulazni parametri za razvoj modela korišćeni su široko dostupni ekonomski, industrijski i indikatori održivog razvoja. U ovom slučaju su za razvoj modela korišćene dve arhitekture neuronskih mreža, i to troslojna MLP (BPNN - *Backpropagation Neural Network*) i generalna regresiona neuronska mreža (GRNN). Pored modela zasnovanih na neuronskim mrežama, za komparaciju rezultata razvijenih ANN modela, razvijen je i konvencionalni MLR model.

ANN i MLR modeli su razvijeni i provereni sa godišnjim podacima za 20 EU država. Na osnovu dobijenih rezultata, zaključeno je da razvijeni ANN model daje značajno bolja predviđanja u odnosu na MLR model. Navedenim istraživanjem se došlo do zaključka da ANN model može uspešno i precizno predvideti nacionalnu emisiju metana za period od dve godine, čime se otvara mogućnost primene ovog pristupa modelovanja kao podrške za implementaciju strategije održivog razvoja i politike upravljanja životnom sredinom [209].

7.1. Ulazne i izlazne promenljive i priprema podataka za razvoj modela

Kao što je već naglašeno, jedan od veoma bitnih koraka u razvoju ANN modela je pronalaženje odgovarajućih ulaznih promenljivih koje imaju značajan uticaj na izlaznu promenljivu [210], u ovom slučaju na emisiju metana. Smatra se da ekonomske aktivnosti u velikoj meri doprinose ukupnoj emisiji metana, različiti industrijski procesi, poljoprivreda kao i proizvodnja i potrošnja energije [40].

Za razvoj ANN modela za predviđanje emisije metana je kao jedna od ulaznih promenljivih odabran bruto domaći proizvod (GDP - *Gross Domestic Product*). GDP je specifičan parametar za svaku državu i u ovom slučaju on generalno predstavlja indikator veličine ekonomske i privredne aktivnosti. GDP je već korišćen kao ulazni parametar za modelovanje emisije ANN pristupom i pokazao se kao značajan [51,52].

Količina generisanog komunalnog otpada (MWG - *Municipal Waste Generation*) kao i odlaganje komunalnog otpada na ili ispod površine zemlje (WDL - *Waste Deposit onto or into Land*) su takođe odabrani kao ulazne promenljive za model, budući da se metan emituje sa deponija usled anaerobne biodegradacije vlažnog otpada [211].

Još jedan od značajnih antropogenih emitera metana su i obradive površine prekrivene pirinčem. Razlog za značajnu emisiju metana sa pirinačnih polja leži u činjenici da je nivo organskog ugljenika na ovim površinama veliki, kao i produženi anaerobni uslovi u toku rasta pirinča [212,213]. Iz navedenih razloga je parametar koji pokazuje površinu zemljišta korišćenog za uzgajanje pirinča (LR - *Land use - Rice*) odabran kao još jedna ulazna promenljiva za razvoj ANN modela.

Emisija metana nastala usled crevne fermentacije kod životinja i iz prirodnih đubriva, su još dva značajna izvora emisije [214]. Da bi se obuhvatio ovaj sektor emisije metana, kao ulazna promenljiva je odabran broj grla stoke (NC - *Number of Cattle*). Pored toga, metan je sastavni deo prirodnog gasa, pa je u tom smislu, primarna proizvodnja prirodnog gasa (PPG - *Primary Production of Gas*) u direktnoj vezi sa emisijom metana, tako da je i taj parametar odabran kao ulazna promenljiva za razvoj modela. Podaci za sve ulazne promenljive na nacionalnom nivou, za svih 20 EU država uključenih u ovom istraživanju su preuzeti iz Evropske statističke baze podataka – Eurostat- a [195]. Podaci za izlaznu promenljivu, u ovom slučaju nacionalna emisija metana, preuzeti su iz baze podataka emisije zagađujućih materija za globalna atmosferska istraživanja – EDGAR [215].

Ulazne promenljive koje su izabrane za razvoj modela, kao i izlazna promenljiva, sa jedinicama nakon normalizacije i izvor podataka su dati u Tabeli 14. Odabrane ulazne promenljive su, kao i u prethodnim slučajevima, normalizovane po stanovniku i/ili proseku GDP za EU27, kako bi se omogućilo poređenje različitih država.

Tabela 14. Lista ulaznih i izlaznih promjenljivih

Ulazna promjenljiva	Jedinica ^a	Izvor podataka
GDP	_ ^b	Eurostat [195]
WDL	kg po st.	
MWG	kg po st.	
LR	ha po st.	
NC	broj grla po st. pc ^c	
PPG	toe po st. ^d	
Izlazna promjenljiva		
CH ₄ (emisija)	kg po st.	EDGAR [215]

^aJedinica nakon normalizacije

^bBez jedinice zbog normalizacije po proseku GDP za EU27

^c Broj grla stoke po stanovniku

^dTona ekvivalenta nafte (toe) po stanovniku

ANN modeli za predviđanje emisije metana su obučavani, testirani i validirani sa dostupnim podacima za 20 EU država za period od 2000. do 2008., budući da podaci posle 2008. godine nisu dostupni. Dostupni podaci su podeljeni u tri seta:

- Podaci za obučavanje – za obučavanje mreža su korišćeni podaci od 2000. do 2006; primer podataka za obučavanje mreža za 2006. godinu je prikazan u Tabeli 15.
- Podaci za testiranje – grupa podataka izdvojenih iz trening skupa kojima se testira mreža u procesu obučavanja. Oko 15% podataka iz trening skupa, nasumično izabranih, je korišćeno za testiranje mreže.
- Podaci za validaciju mreže – za validaciju mreža su korišćeni podaci za period 2007 - 2008.

Tabela 15. Podaci za obučavanje za 2006. godinu

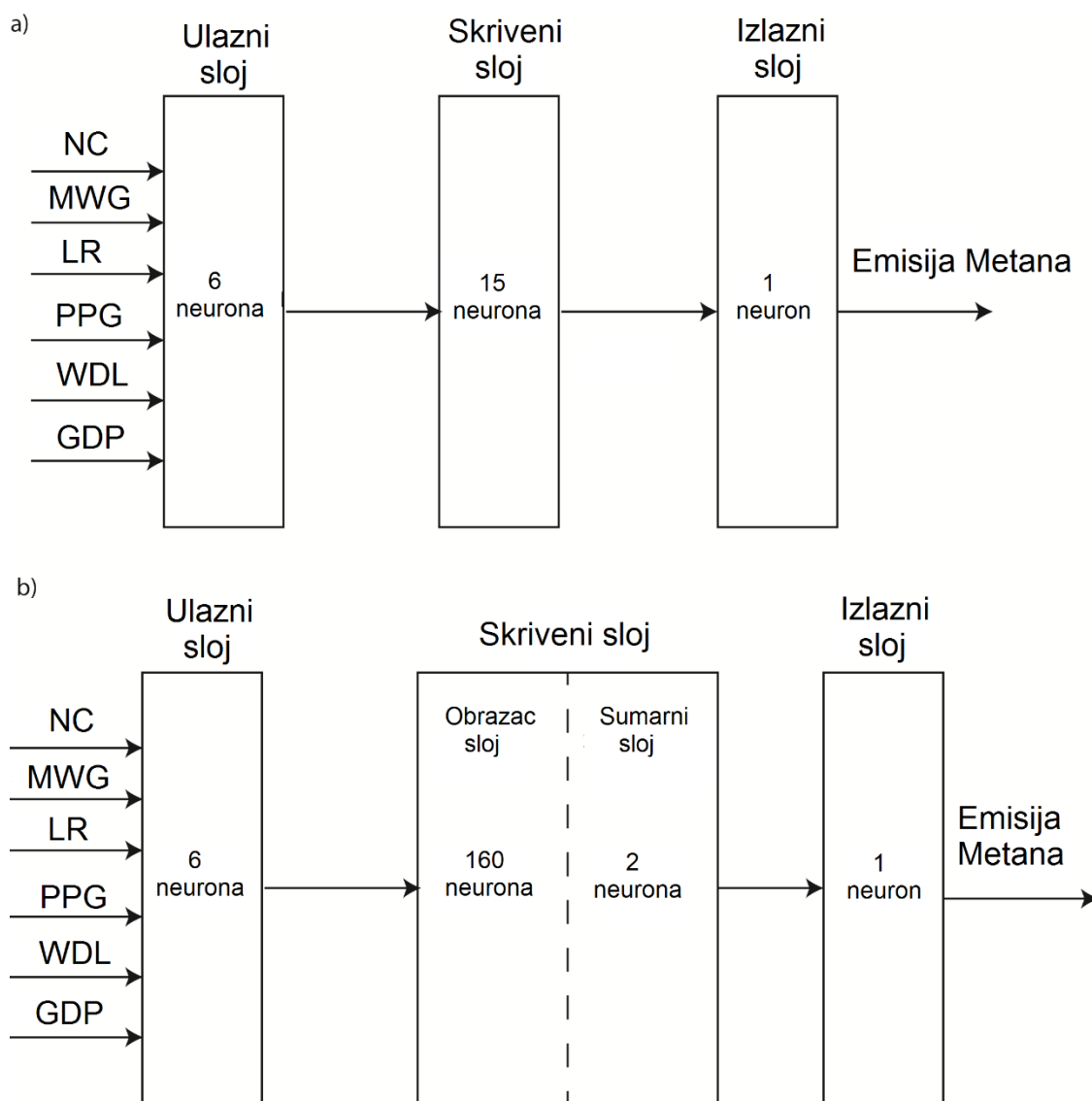
Država	GDP	WDL [kg st.]	MWG [kg st.]	NC [gr. st.]	LR x10 ⁻⁷ [ha st.]	PPG [toe st.]	CH ₄ [kg st.]
Bugarska	0.14	357	461	0.082	5.83	0.048	78.05
Češka	0.49	199	296	0.136	0	0.014	55.77
Danska	1.70	37	740	0.291	0	1.718	68.83
Estonija	0.42	278	399	0.182	0	0.000	78.68
Grčka	0.79	385	442	0.061	20.8	0.002	35.41
Španija	0.95	355	594	0.141	24.3	0.001	39.25
Francuska	1.20	194	536	0.299	27.4	0.017	62.68
Letonija	0.30	293	412	0.164	0	0.000	65.68
Litvanija	0.30	357	391	0.246	0	0.000	71.30
Luksenburg	3.03	130	683	0.397	0	0.000	108.65
Mađarska	0.38	376	468	0.070	2.38	0.236	36.81
Holandija	1.40	15	622	0.225	0	3.391	60.85
Austrija	1.32	25	653	0.243	0	0.189	48.64
Poljska	0.30	236	321	0.138	0	0.102	87.27
Portugal	0.64	297	463	0.133	24.0	0.000	58.38
Rumunija	0.19	292	389	0.136	2.59	0.442	58.41
Slovenija	0.65	361	431	0.227	0	0.001	70.65
Slovačka	0.35	234	301	0.094	0	0.033	36.12
Finska	1.33	286	494	0.177	0	0.000	88.82
Velika Britanija	1.36	352	586	0.171	0	1.192	49.67

7.2. Razvoj i rezultati modela

Za razvoj ANN modela su korišćene dve arhitekture neuronskih mreža čiji su rezultati predviđanja poređeni: troslojna mreža sa kretanjem signala napred (BPNN) koja je najčešće korišćena i generalna regresiona neuronska mreža (GRNN) koja je pokazala dobre rezultate u modelovanju emisije [51,52].

Kod BPNN je broj neurona u ulaznom i izlaznom sloju jednak broju ulaznih i izlaznih promenljivih, dok je broj neurona u skrivenom sloju određen korišćenjem formule (9). Broj skrivenih neurona kod razvijenog BPNN modela je 15, sa logističkom aktivacionom funkcijom, dok su vrednosti parametara stope obučavanja i momentum 0.1. Obučavanje mreže je trajalo 30, minuta tokom kojih je izvedeno 73952 epoha. Šematski prikaz BPNN modela je prikazan na Slici 21.a.

Kod GRNN je takođe broj neurona u ulaznom i izlaznom sloju jednak broju ulaznih i izlaznih promenljivih, dok je broj neurona u skrivenom sloju jednak broju serija podataka iz skupa za obučavanje, tj. 160 neurona. Ulazni podaci su skalirani linearnom funkcijom dok je faktor ujednačavanja određen genetskim algoritmom. Šematski prikaz GRNN modela je prikazan na Slici 21.b.



Slika 21. Šematski prikaz ANN arhitektura: a) BPNN i b) GRNN

Uporedo sa razvojem ANN modela, razvijen je i regresioni MLR model sa istim podacima koji su korišćeni i kod razvoja ANN, sa ciljem komparacije rezultata. MLR model je kreiran primenom *SPSS 19* softvera [168] i dobijeni MLR koeficijenti su predstavljeni jednačinom:

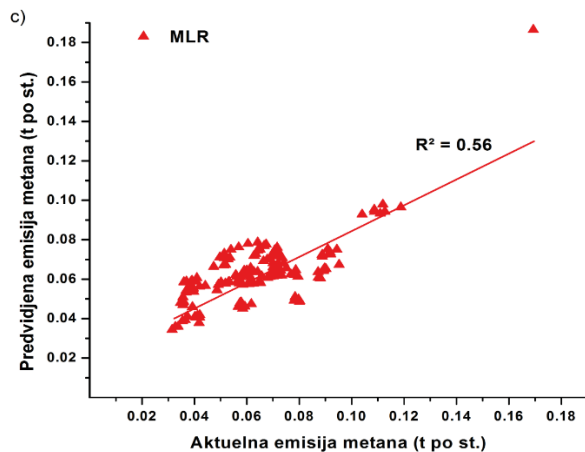
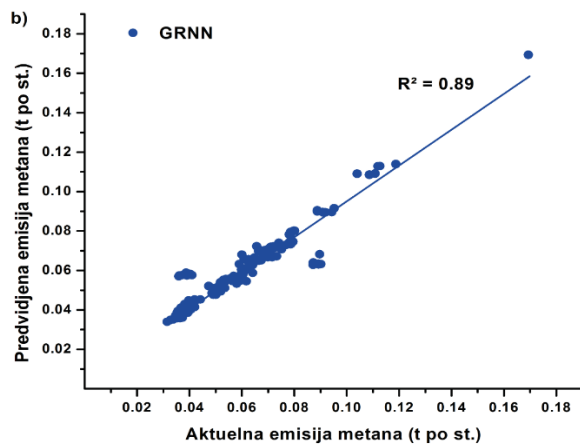
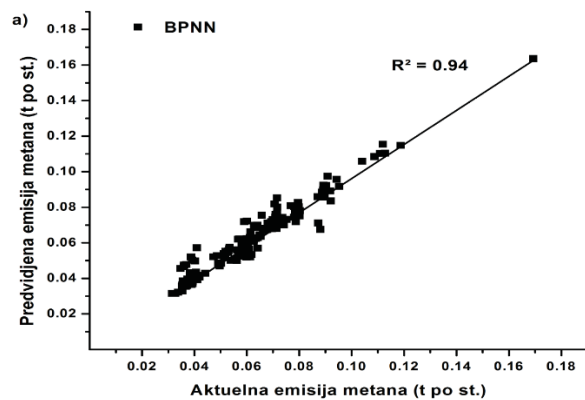
$$CH_4 \text{ emisija} = 0.0480 + 0.0001WDL - 0.0001MWG + 0.0922NC - 6465.1928LR + 0.0008PPG + 0.0131GDP \quad (35)$$

gde su: WDL – odlaganje otpada na ili ispod površine zemlje, MWG – količina generisanog komunalnog otpada, NC – broj grla stoke, LR – površina zemlje pod pirinčem, PPG – proizvodnja primarnog gasa i GDP – bruto domaći proizvod.

U Tabeli 16. date su vrednosti indikatora performansi razvijenih modela, kako za fazu obučavanja, tako i za fazu validacije. Na Slici 22. je prikazana komparacija između aktuelnih i modelom predviđenih vrednosti za trening podatke (2000 - 2006.).

Tabela 16. Vrednosti indikatora performansi kreiranih modela

Model	Indikatori performansi				
	<i>IA</i>	<i>FAI.25</i> [%]	<i>MAE</i> [kg st.]	<i>RMSE</i> [kg st.]	<i>R</i>
Podaci za obučavanje					
BPNN	1.00	94	3.4	5.0	0.97
GRNN	0.97	92	3.6	7.0	0.94
MLR	0.83	64	11.3	14.0	0.75
Podaci za validaciju					
BPNN	1.00	90	5.6	7.7	0.95
GRNN	0.98	93	4.8	7.3	0.96
MLR	0.90	58	11.9	252.5	0.83

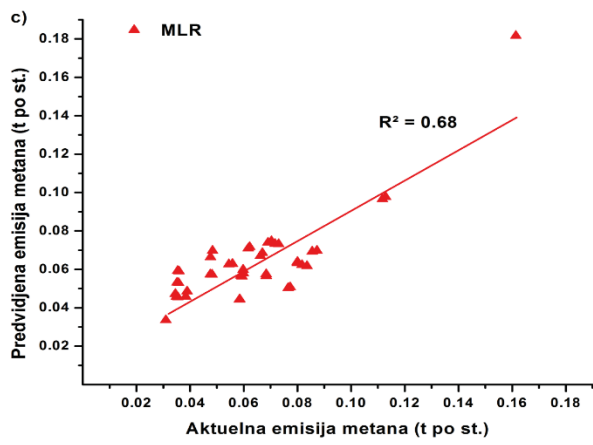
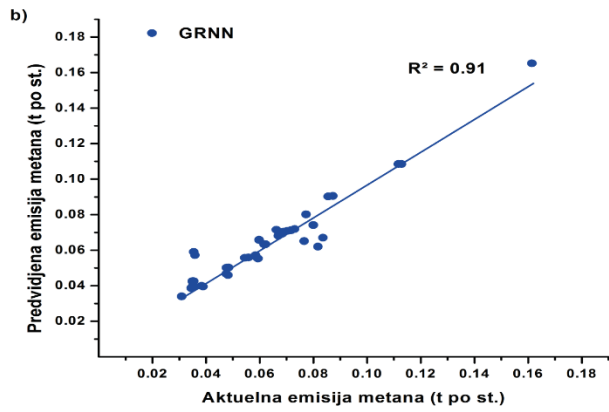
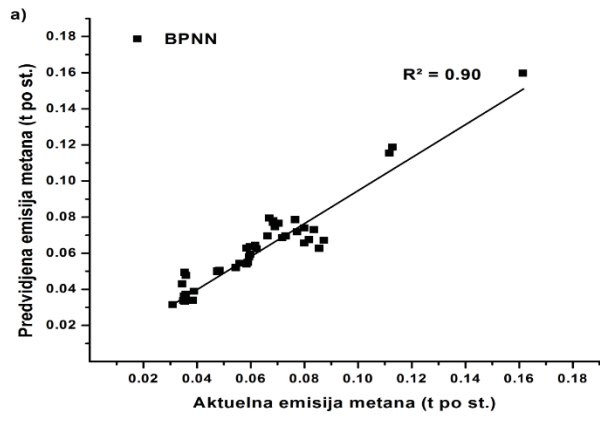


Slika 22. Poređenje aktuelnih i predviđenih vrednosti emisije metana za trening podatke a) BNN, b) GRNN, c) MLR

Podaci za validaciju modela od 2007. do 2008. godine, predstavljaju potpuno nove podatke koji nisu prezentovani modelima u fazi obučavanja. Validacioni set podataka koji je predstavljen u Tabeli 17. je korišćen za procenu mogućnosti modela da pravi dobra predviđanja. Rezultati kreiranih modela BPNN, GRNN i MLR na podacima za validaciju su prikazani na Slici 23., dok su vrednosti indikatora performansi modela prikazane u Tabeli 16. Aktuelne emisije metana, GRNN predviđanja i relativne greške za 2007. i 2008. godinu su prikazane u Tabeli 18.

Tabela 17. Validacioni set podataka (vrednosti za 2007. i 2008. godinu)

Država	GDP	WDL [kg st.]	MWG [kg st.]	NC [gr. st.]	LR x10 ⁻⁷ [ha st.]	PPG [toe st.]
Bugarska	0.16	389	433	0.080	8.59	0.031
Češka	0.51	205	293	0.133	0	0.016
Danska	1.67	37	790	0.284	0	1.518
Estonija	0.48	291	449	0.179	0	0.000
Grčka	0.80	357	447	0.061	23	0.002
Španija	0.94	347	583	0.148	21.6	0.000
Francuska	1.18	194	543	0.300	2.59	0.014
Letonija	0.37	323	378	0.175	0	0.000
Litvanija	0.34	369	401	0.233	0	0.000
Luksenburg	3.12	126	695	0.406	0	0.000
Mađarska	0.40	341	457	0.070	2.58	0.199
Holandija	1.40	13	629	0.234	0	3.328
Austrija	1.32	20	596	0.241	0	0.192
Poljska	0.33	239	322	0.142	0	0.102
Portugal	0.64	299	468	0.136	25.4	0.000
Rumunija	0.23	284	379	0.131	3.9	0.428
Slovenija	0.68	341	439	0.239	0	0.001
Slovačka	0.41	240	309	0.093	0	0.020
Finska	1.36	267	506	0.171	0	0.000
Velika Britanija	1.35	323	570	0.166	0	1.068
Bugarska	0.18	441	474	0.075	10.2	0.020
Češka	0.59	197	305	0.131	0	0.016
Danska	1.71	32	830	0.292	0	1.646
Estonija	0.49	248	391	0.177	0	0.000
Grčka	0.83	372	452	0.061	24.1	0.001
Španija	0.96	287	556	0.133	21.1	0.000
Francuska	1.20	171	541	0.303	2.56	0.013
Letonija	0.40	311	332	0.167	0	0.000
Litvanija	0.39	368	408	0.229	0	0.000
Luksenburg	3.23	123	697	0.406	0	0.000
Mađarska	0.42	333	454	0.070	2.49	0.200
Holandija	1.45	8	624	0.244	0	3.651
Austrija	1.36	19	599	0.240	0	0.158
Poljska	0.38	228	320	0.146	0	0.097
Portugal	0.65	332	515	0.136	24.8	0.000
Rumunija	0.26	301	392	0.125	4.6	0.418
Slovenija	0.74	339	457	0.234	0	0.001
Slovačka	0.48	250	328	0.090	0	0.016
Finska	1.40	265	521	0.171	0	0.000
Velika Britanija	1.17	287	544	0.162	0	1.025



Slika 23. Poređenje aktuelnih i predviđenih vrednosti emisije metana za validacioni set podataka: a) BPNN, b) GRNN, c) MLR

Tabela 18. Aktuelne i GRNN modelom predviđene vrednosti emisije metana sa relativnim greškama za 2007. i 2008. godinu

Država	Godina 2007.			Godina 2008.		
	Aktuelne emisije [kg st.]	GRNN [kg st.]	Relativna greška [%]	Aktuelne emisije [kg st.]	GRNN [kg st.]	Relativna greška [%]
Bugarska	77.28	80.17	3.74	76.54	65.12	-14.92
Češka	55.80	55.94	0.28	54.45	55.73	2.37
Danska	68.33	69.30	1.42	68.39	69.32	1.36
Estonija	79.95	74.12	-7.29	79.97	74.22	-7.19
Grčka	35.44	42.54	20.02	34.87	42.49	21.85
Španija	38.87	39.53	1.71	38.37	39.96	4.15
Francuska	62.16	63.34	1.90	61.67	63.32	2.67
Letonija	66.19	71.49	8.01	66.93	68.16	1.84
Litvanija	71.69	71.22	-0.65	72.98	71.92	1.46
Luksenburg	111.64	108.51	-2.81	112.85	108.51	-3.84
Mađarska	35.73	39.35	10.13	35.12	40.18	14.39
Holandija	59.80	65.75	9.95	59.77	65.87	10.21
Austrija	48.20	45.97	-4.64	47.54	46.72	-1.72
Poljska	83.57	67.05	-19.77	81.71	62.03	-24.09
Portugal	58.45	56.48	-3.39	58.44	57.12	-2.25
Rumunija	58.91	55.93	-5.06	59.35	55.34	-6.75
Slovenija	70.33	70.88	0.78	69.00	70.43	2.07
Slovačka	35.86	57.23	59.60	35.40	59.08	66.90
Finska	87.25	90.57	3.80	85.49	90.24	5.55
Velika Britanija	48.39	50.26	3.86	47.61	50.12	5.27

Na osnovu rezultata prikazanih u Tabeli 16. i na Slici 22., dobijenih od strane kreiranih modela u trening fazi, može se videti da su BPNN i GRNN modeli pokazali mnogo bolje rezultate u odnosu na regresioni MLR model. Dakle, može se reći da modeli zasnovani na ANN pokazuju veoma dobre performanse predviđanja.

Rezultati ocene (validacije) modela, njihovih stvarnih mogućnosti predviđanja, su pokazali da GRNN arhitektura pokazuje malo bolje rezultate u odnosu na drugu primenjenu BPNN arhitekturu i značajno bolje rezultate u odnosu na MLR model (Slika 23.). Najznačajniji statistički indikatori performansi modela *FA1.25* i *IA* pokazuju bolje rezultate ANN modela na validacionom setu podataka, tj. za GRNN model *FA1.25* iznosi 93% i 90% za BPNN

model, dok je vrednost *FAI.25* za MLR model samo 58%. Takođe, za GRNN model *IA* ima vrednost 0.98, za BPNN je 1.00, dok je vrednost *IA* za MLR model 0.90. Za GRNN model se može zaključiti da je optimalno obučen, jer pokazuje veoma dobre rezultate sa sličnim vrednostima *FAI.25* i drugih indikatora greške, kako na trening tako i na validacionom setu podataka. S druge strane BPNN model pokazuje slične, ali malo lošije rezultate na validacionom setu podataka u odnosu na trening podatke (Tabela 16.), što bi značilo da bi se BPNN mogla redizajnirati, u cilju optimizacije rezultata predviđanja.

GRNN model je postigao 75% predviđanja uspešno, sa relativnom greškom manjom od 10%, drugih 10% predviđanja je postigao sa relativnom greškom između 10% i 20%, dok je samo 15% predviđanja postigao sa relativnom greškom većom od 20%. U slučaju Mađarske, Bugarske i Holandije, predviđanja GRNN modela su sa relativnom greškom između 10% i 20%, dok je u slučaju Slovačke, Poljske i Grčke relativna greška modela iznad 20% za obe godine (Tabela 18.).

Nesigurnost koja se odnosi na ulazne parametre, kod modela zasnovanih na inventarskom pristupu (emisioni faktori i stope aktivnosti), je značajna, tj. od 20-30% pa do 50% u slučaju Austrije, što za posledicu ima neizvesnost podataka procenjene emisije metana za 20-40% [216]. S obzirom na to da GRNN model koristi značajno manji broj ulaznih promenljivih i postiže predviđanja sa relativnom greškom manjom od 20% u većini ispitivanih država, čini ANN pristup odgovarajućom alternativom za predviđanje emisije metana na nacionalnom nivou, pogotovo u slučajevima kada je procena emisije inventarskim pristupom onemogućena zbog nedostatka podataka.

Problem odstupanja između aktuelnih i predviđenih vrednosti za države kod kojih je relativna greška iznad 10%, može se pripisati kvalitetu ulaznih i izlaznih podataka korišćenih tokom kreiranja modela.

U slučaju Slovačke, podaci za parameter WDL za 2011. godinu su procenjeni. Takođe, podaci za parameter MWG do 2001. godine su prikupljeni prema nacionalnom katalogu otpada (*National Waste Catalogue*), dok se od 2002. godine prikupljaju prema Evropskoj listi otpada (*The European List of Waste*) [195]. U tom smislu, primena procenjenih, ili

trening podataka koji nisu dobijeni primenom standardnih metoda, može usloviti veliku relativnu grešku za 2007. i 2008. godinu (Tabela 18.). Relativna greška u slučaju Holandije za 2008. godinu je 10.21% i može se pripisati zabrani direktnog odlaganja mešanog komunalnog otpada u 2003. godini, što je dovelo do smanjenja WDL od 50 kg po stanovniku u 2002. do samo 11 kg po stanovniku u 2004. godini [195].

U slučaju Grčke, trening podaci za parameter GDP za 2005. godinu su nesigurni zbog netačnog izveštavanja o stanju ekonomskog razvoja od strane vlade. Uticaj ovog parametra se može smatrati razlogom za veće vrednosti relativne greške i za 2007. i za 2008. godinu.

MWG podaci za Poljsku i Mađarsku su procenjene, za 2003. godinu u slučaju Mađarske i od 2005. do 2008. godine u slučaju Poljske, što može objasniti dobijene velike relativne greške (14.39% Mađarska i -24.09 Poljska).

U slučaju Bugarske, GRNN model je dao relativnu grešku predviđanja od -14.92% za 2008. godinu. Deponija u Sofiji je bila zatvorena u periodu od oktobra 2005. do decembra 2007. godine, što je uslovalo privremeno skladištenje sakupljenog komunalnog otpada od 2.5% u 2005. i 10% u 2006. i 2007. godini. Ovi iznosi nisu bili uključeni u podacima o količinama generisanog otpada, a to je takođe uslovalo i pad količine deponovanog otpada u 2006. i 2007. godini [195]. Uzimajući u obzir sve ove faktore uticaja, promene u MWG za 2005. i 2006. godinu mogu biti uzrok velike relativne greške za 2008. godinu.

8. Predviđanje emisije azotnih oksida u EU, SAD, Kini, Indiji, Japanu i Rusiji

Emisija azotovih oksida (NO_x) u atmosferu ima višestruke negativne efekte na životnu sredinu i direktne ili indirektno efekte na ljudsko zdravlje. U ovom poglavlju je opisan razvoj i rezultati modela zasnovanog na neuronskim mrežama, za predviđanje emisije azotovih oksida na nacionalnom nivou, primenom, kao i u prethodnim slučajevima predviđanja, ekonomskih, industrijskih, poljoprivrednih i indikatora održivog razvoja, kao ulaznih promenljivih modela. Za razvoj modela su korišćeni dostupni podaci za 17 evropskih država, SAD, Kinu, kao i podaci (za razliku od prethodnih modela) za Japan, Rusiju i Indiju za period 2001 - 2008. godine. Za kreiranje ANN modela korišćena je GRNN arhitektura. Inicijalni GRNN model je razvijen sa svih 11 potencijalnih ulaznih promenljivih, a nakon toga su za izbor nezavisnih i značajnih ulaznih promenljivih na originalnim podacima primenjene korelaciona i VIF analiza. Najbolji rezultati selekcije ulaznih promenljivih su dobijeni na osnovu korelacije između ulaznih promenljivih. Naime, GRNN modeli kreirani sa manjim brojem ulaznih promenljivih su pokazali mnogo bolje predviđanje u poređenju sa GRNN modelom kreiranim sa svim potencijalnim ulaznim promenljivama [217].

8.1. Ulazni i izlazni podaci

Generalno govoreći, na globalnom nivou, proces sagorevanja u termoelektranama, industrija, drumski saobraćaj, domaćinstva i poljoprivredne aktivnosti su najveći emiteri oksida azota. Drugim rečima, rast ekonomskog i industrijskog razvoja vode ka povećanoj emisiji zagađujućih materija u vazduh, među kojima spada i emisija NO_x . U tom smislu je, u ovom slučaju predviđanja emisije, odabrano 11 potencijalnih ulaznih promenljivih za koje se smatra da obuhvataju najznačajnije izvore emisije azotnih oksida:

1. Proizvodnja električne energije iz uglja (EPCS - *Electricity Production from Coal Sources*),
2. Proizvodnja električne energije iz nafte (EPOS - *Electricity Production from Oil Sources*),

3. Potrošnja dizel goriva u drumskom sektoru (RDFC - *Road sector diesel fuel consumption*),
4. Potrošnja energije u drumskom sektoru (RSEC - *Road Sector Energy Consumption*),
5. Bruto domaći proizvod (GDP - *Gross Domestic Product*),
6. Upotreba energije (EU - *Energy Use*),
7. Potrošnja energije iz fosilnih goriva (FFEC - *Fossil Fuel Energy Consumption*),
8. Električna energije proizvedena iz obnovljivih izvora (EPRS - *Electricity Production from Renewable Sources*),
9. Broj proizvedenih vozila (NV - *Number of Vehicles*),
10. Potrošnja električne energije (EPC - *Electric Power Consumption*),
11. Poljoprivredno zemljište (AL - *Agricultural Land*)

Podaci za svih 11 ulaznih promenljivih za razvoj modela su preuzeti iz tri statističke baze podataka: Svetska banka (*World Bank*), OICA – proizvodna statistika (*Organisation Internationale des Constructeurs d'Automobiles*) i internacionalne statističke baze podataka NationMaster [204,206,218]. Spisak svih 11 odabranih promenljivih, kao i njihovi statistički pokazatelji su prikazani u Tabeli 19.

Tabela 19. Statistička analiza 11 inicijalnih ulaznih promenljivih

Ulazna promenljiva	Jedinica	Minimalna vrednost	Maksimalna vrednost	Srednja vrednost	Standardna devijacija
EPCS	kwh po st.	147.570	7288.78	1987.83	1657.65
EPOS	kwh po st.	7.920	1535.48	234.20	281.17
RDFC	toe po st.	0.010	0.69	0.30	0.17
RSEC	toe po st.	0.027	1.80	0.59	0.36
GDP	€ po st.	371.722	40400.00	20978.89	12432.13
EU	toe po st.	0.438	782.83	9.21	59.60
FFEC	% od ukupne	0.325	0.96	0.77	0.15
EPRS	kwh po st.	23.330	9366.16	1417.97	1947.67
NV	br.vozila po st.	0.001	0.12	0.04	0.03
EPC	kwh po st.	399.571	17212.95	6780.86	3931.43
AL	%	0.073	0.74	0.44	0.19

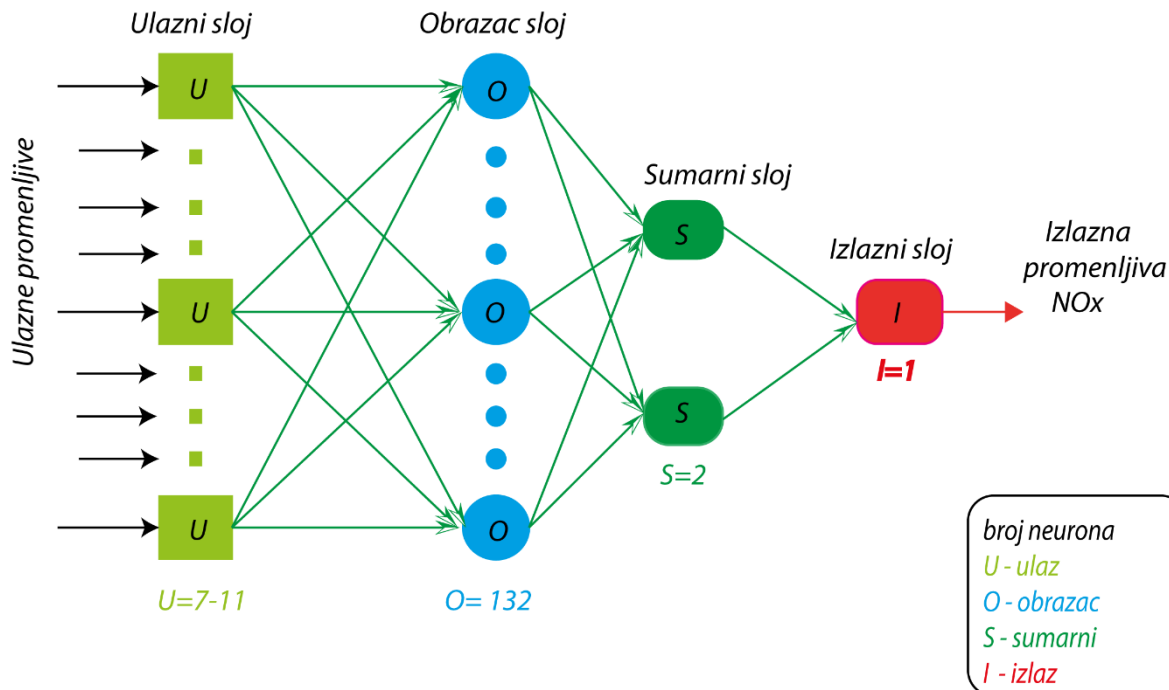
Kao i u ostalim slučajevima razvoja ANN modela, opisanih u prethodnim poglavljima i u slučaju NOx se podaci za razvoj modela normalizuju po stanovniku radi postizanja boljih rezultata predviđanja i mogućnosti komparacije različitih država. ANN modeli su obučavani podacima od 2001. do 2006. godine za 17 EU država, za SAD, Kinu, Rusiju, Indiju i Japan, dok su podaci 2007 - 2008. godine korišćeni za validaciju modela. Vrednost statističkih pokazatelja za NOx emisiju na nacionalnom nivou preuzete sa EDGAR baze podataka emisija zagađujućih materija [215] za svaku od država uključenih u razvoj NOx modela i za različite setove podataka su prikazane u Tabeli 20.

Tabela 20. Statistička analiza NOx emisije na nacionalnom nivou

Država	NOx emisija (kg po st.)			
	Min. vr.	Maks. vr.	Sr. vr.	St. dev.
Belgija	24.30	31.92	29.22	2.47
Češka	25.98	32.68	29.62	2.64
Nemačka	21.07	25.37	23.14	1.42
Španija	41.13	47.84	46.38	2.24
Francuska	21.31	28.26	25.51	2.34
Italija	21.06	26.97	24.27	2.08
Mađarska	15.94	20.77	18.17	2.05
Holandija	25.35	31.24	28.50	2.01
Austrija	25.78	29.78	28.25	1.34
Poljska	21.16	22.69	22.12	0.51
Portugal	24.33	29.15	27.40	1.57
Rumunija	12.86	14.87	13.95	0.73
Slovenija	24.20	29.76	26.18	1.66
Slovačka	17.41	20.04	18.42	0.82
Finska	37.07	50.57	44.30	4.79
Švedska	31.83	36.46	34.85	1.57
Velika Britanija	28.21	34.54	31.64	1.83
SAD	46.81	62.36	55.55	5.02
Kina	9.09	15.66	12.12	2.33
Indija	5.04	6.07	5.49	0.35
Japan	18.74	21.55	20.50	0.94
Rusija	29.81	43.17	34.74	4.98
Skup podataka				
Podaci za obučavanje	5.04	62.36	27.90	11.85
Podaci za validaciju	5.88	51.97	25.45	9.97

8.2. Razvoj i optimizacija GRNN modela

Za razvoj ANN modela za predviđanje emisije NO_x u ovom radu je korišćena GRNN arhitektura (Slika 24.). Pre početka obučavanja GRNN mreže, nakon odabira potencijalnih ulaznih promenljivih, sledeći korak je bio izdvajanje podataka za testiranje mreže iz grupe podataka za obučavanje. Podaci za testiranje mreže su izdvojeni nasumičnim odabirom 20% od ukupnog broja (132) serija podataka za obučavanje mreže. GRNN modeli su obučavani korišćenjem linearne funkcije (u opsegu od 0 do 1) za skaliranje ulaznih podataka, dok je faktor ujednačavanja određen genetskim algoritmom. Dakle, parametri obučavanja GRNN modela su isti kao i kod modelovanja emisije metana budući da je taj model pokazao dobre rezultate. Nakon toga su različite kombinacije ulaznih promenljivih, dobijenih primenom nekoliko pristupa za selekciju ulaznih promenljivih, korišćene za kreiranje četiri GRNN modela (Tabela 21.). Naime, osnovni cilj ovog koraka optimizacije modela je da se odabere manji broj ulaznih varijabli koje najbolje karakterišu osnovni ulazno – izlazni odnos, istovremeno smanjujući suvišne ulazne promenljive [103].



Slika 24. GRNN arhitektura

Tabela 21. Detaljne informacije kreiranih GRNN modela

Model	Selekcija ulaza	Kritična vrednost ^a	Uklonjeni ulazi	Arhitektura ^b
GRNN1	Svi ulazi	-	-	11-106-2-1
GRNN2	Korelaciona analiza (ulaz-izlaz korelacija)	$r < 0.1$	FFEC	10-106-2-1
GRNN3	Korelaciona analiza (korelacija između ulaza)	$r > 0.7$	a: RSEC, GDP, EPC, EPRS b: RDFC, GDP, FFEC c: RSEC, FFEC, EPC d: RDFC, RSEC, EPRS, EPC	a: 7-106-2-1 b: 8-106-2-1 c: 8-106-2-1 d: 7-106-2-1
GRNN4	VIF analiza	$VIF > 5$	a: EPC b: EPC, GDP c: EPC, GDP, RSEC	a: 10-106-2-1 b: 9-106-2-1 c: 8-106-2-1

^aza uklanjanje ulaza

^bbroj neurona po slojevima

Prvi pristup za selekciju ulaznih promenljivih je zasnovan na korelacionoj analizi (Tabela 22.). Model označen kao GRNN2 je kreiran na osnovu "ulaz - izlaz" korelacije, i sadrži one ulazne promenljive koje su u korelaciji sa emisijom NO_x većom od 0.1. U ovom slučaju, kao što se može videti u Tabeli 22., samo jedna ulazna promenljiva (FFEC) od potencijalnih 11 se treba ukloniti iz modela.

S druge strane, GRNN3 model je kreiran na osnovu "ulaz - ulaz" korelacije. Ulazne promenljive sa međusobnom korelacijom većom od 0.7 su uklonjene iz polaznog seta podataka. Na osnovu rezultata korelacione analize, može se videti da postoji šest parova visoko korelisanih ulaznih promenljivih (RSEC-RDFC, GDP-RDFC, GDP-RSEC, EPC-RSEC, GDP-EPC i EPRS-FFEC). Dakle, sa ciljem dobijanja nezavisnog seta ulaznih promenljivih, sa različitim kombinacijama nezavisnih ulaznih promenljivih, kreirana su četiri GRNN modela (označena kao GRNN3a, GRNN3b, GRNN3c i GRNN3d).

Tabela 22. Rezultati korelacione analize

	EPCS	EPOS	RDFC	RSEC	GDP	EU	FFEC	EPRS	NV	EPC	AL
EPCS	1										
EPOS	-0.07	1									
RDFC	0.07	0.23	1								
RSEC	0.53	0.27	0.72	1							
GDP	0.15	0.26	0.81	0.78	1						
EU	0.26	0.05	0.03	0.27	0.07	1					
FFEC	0.32	0.20	-0.19	-0.11	-0.30	0.06	1				
EPRS	-0.18	0.03	0.31	0.32	0.45	-0.02	-0.70	1			
NV	0.20	0.13	0.49	0.38	0.41	0.00	-0.12	-0.06	1		
EPC	0.30	0.09	0.52	0.73	0.77	0.14	-0.55	0.69	0.25	1	
AL	0.02	-0.15	-0.06	-0.20	-0.26	0.00	0.48	-0.64	-0.14	-0.64	1
NOx	0.50	0.21	0.56	0.78	0.58	0.25	-0.08	0.33	0.21	0.73	-0.32

Kurzivom su označeni brojevi koji pokazuju jaku korelaciju između ulaznih promjenljivih i jaku korelaciju između ulaznih promjenljivih sa NOx emisijom

Drugi pristup za selekciju ulaznih promjenljivih je zasnovan na VIF analizi. Ovim pristupom se rešava problem multikolinearnosti između ulaznih promjenljivih, i time se određuje koje ulazne promjenljive je potrebno ukloniti. U ovom slučaju su sve one promjenljive koje imaju VIF vrednost veću od 5 sekvencijalno uklonjene. Vrednosti VIF za sve inicijalno odabrane promjenljive su prikazane u Tabeli 23.

Tabela 23. VIF analiza

Ulaz	VIF
EPC	17.03
GDP	8.68
RSEC	7.82
RDFC	4.85
FFEC	4.63
EPCS	4.57
AL	4.51
EPRS	4.22
NV	2.38
EPOS	1.58
EU	1.19

8.3. Rezultati optimizacije GRNN modela

U ovom poglavlju su prikazani rezultati predviđanja osam GRNN modela kreiranih primenom različitih pristupa za odabir najznačajnijih i nekorelisanih ulaznih promenljivih. Vrednosti indikatora performansi svih kreiranih GRNN modela prikazani su u Tabeli 24.

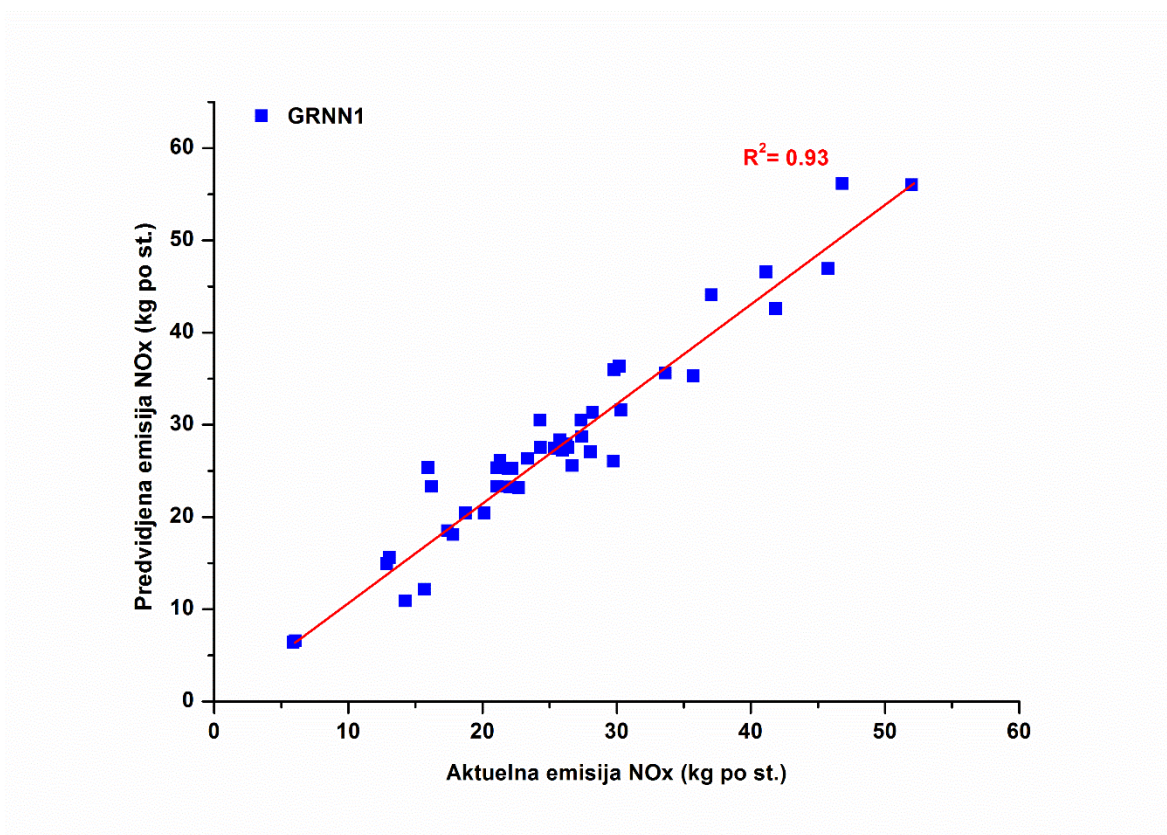
Tabela 24. Vrednosti indikatora performansi kreiranih GRNN modela

Model	Indikatori performansi						
	d_r	NSE	$RMSE$ [kg pc]	MAE [kg pc]	RSR	$FAI.25$ [%]	r
GRNN1	0.92	0.89	3.78	2.97	0.34	89	0.97
GRNN2	0.93	0.91	3.33	2.61	0.30	95	0.97
GRNN3a	0.92	0.86	4.10	2.86	0.38	91	0.95
GRNN3b	0.93	0.91	3.34	2.69	0.30	93	0.97
GRNN3c	0.93	0.91	3.29	2.63	0.30	98	0.98
GRNN3d	0.94	0.93	2.80	2.20	0.26	95	0.98
GRNN4a	0.93	0.92	3.09	2.46	0.28	93	0.97
GRNN4b	0.91	0.87	3.97	3.13	0.37	82	0.96
GRNN4c	0.93	0.87	3.87	2.64	0.36	95	0.95

Kurzivom su označeni brojevi koji ukazuju na najbolje performanse modela

Dobijene vrednosti indikatora performansi ($d_r=0.92$, $FAI.25=89\%$) za GRNN1 model se mogu smatrati dobrim. Poređenje aktuelnih i modelovanih emisija NOx primenjene na validacionom setu podataka je prikazano na Slici 25.

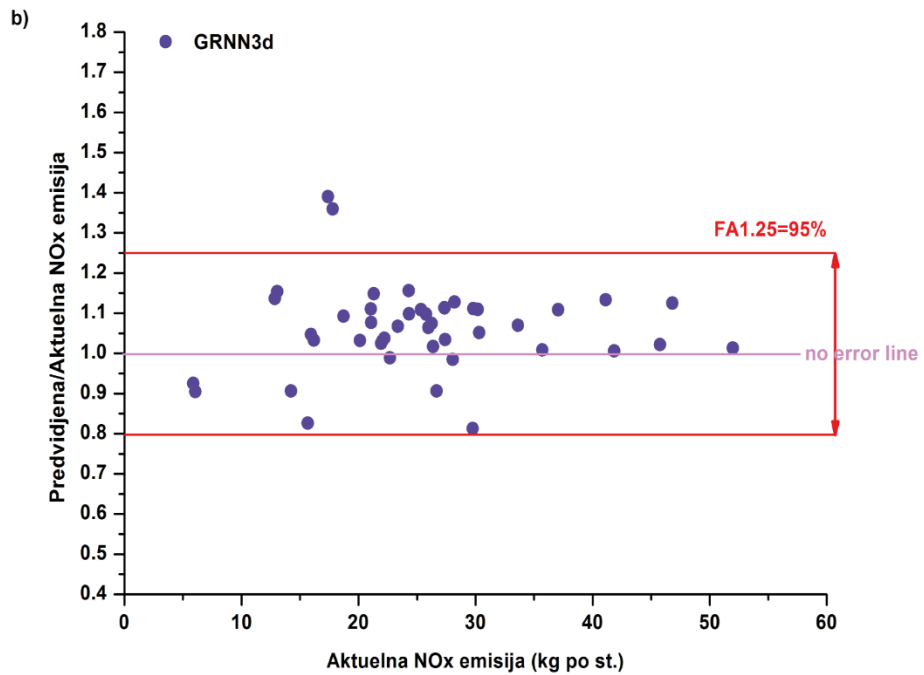
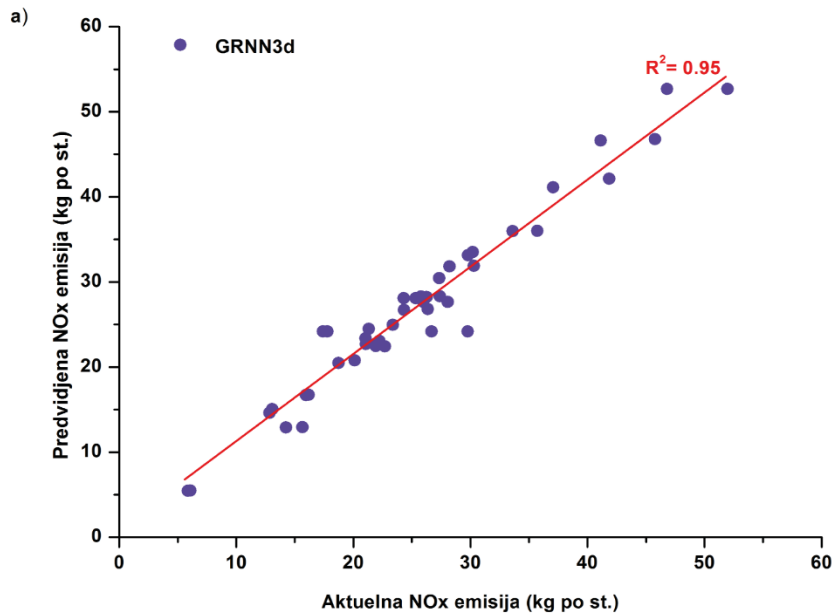
S obzirom na to da kreirani GRNN1 model sa svim inicijalno odabranim ulaznim promenljivama pokazuje dobro slaganje sa aktuelnom NOx emisijom, može se zaključiti da su svih 11 ulaznih promenljivih adekvatne za problem koji se tretira u ovom slučaju.



Slika 25. Poređenje aktuelnih podataka i GRNN1 rezultata (validacioni set podataka)

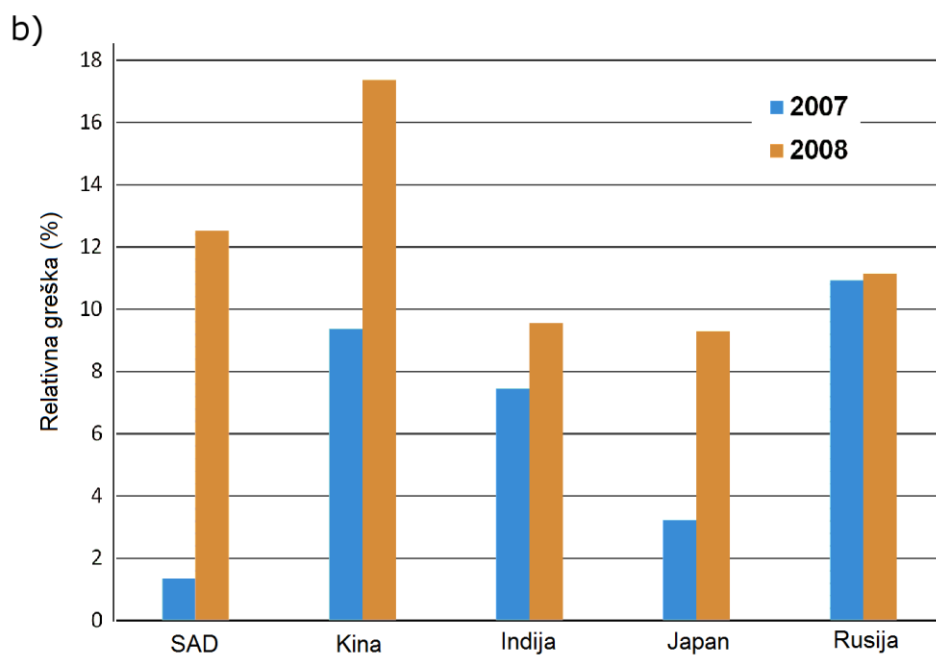
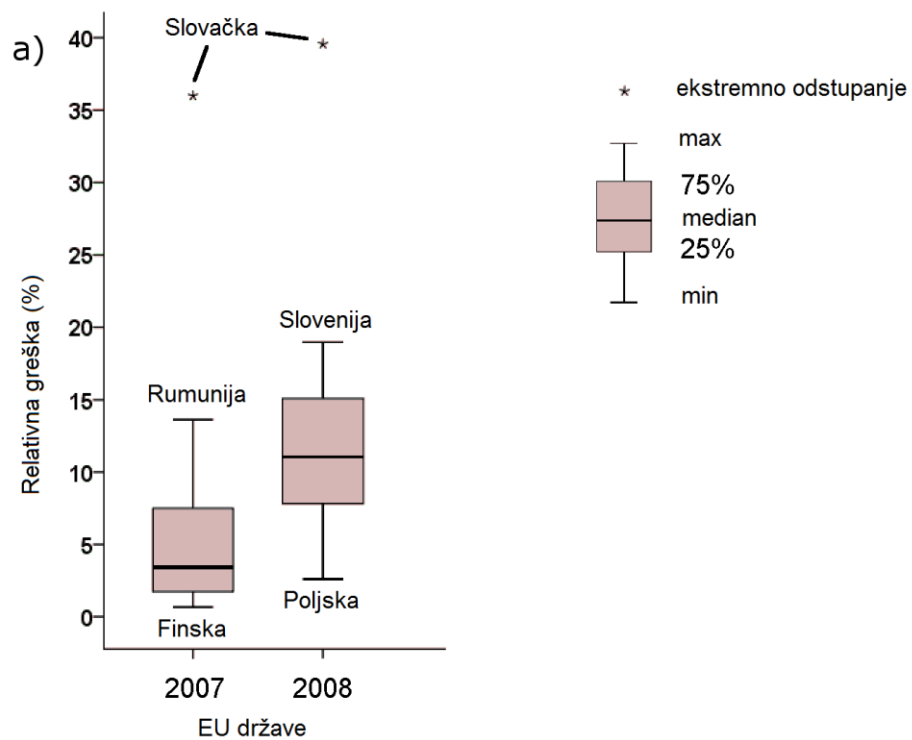
Kao što se može videti iz Tabele 24., svih osam GRNN modela kreiranih primenom različitih tehnika za selekciju ulaznih promenljivih, pokazuju značajno bolje rezultate predviđanja u poređenju sa GRNN1 modelom koji je razvijen sa svih 11 promenljivih. GRNN2 model kreiran na bazi "ulaz - izlaz" korelacije pokazuje zadovoljavajuće rezultate, ali ne i najbolje ($dr=0.93$ and $FAI.25=95\%$); takođe, kod ovog modela je broj ulaznih promenljivih veći u poređenju sa drugim GRNN modelima. Od GRNN modela kreiranih na osnovu VIF analize, GRNN2a model pokazuje bolje performanse ($dr=0.93$ and $FAI.25=93\%$) u odnosu na GRNN2b i GRNN2c modele. Može se zaključiti da dalja redukcija broja ulaznih promenljivih nakon uklanjanja EPC, vodi ka malo lošijim rezultatima modela, a razlog može biti taj što je iz oba modela, GRNN2b i GRNN2c, uklonjen GDP kao ulazni parametar. Naime, svi GRNN modeli kreirani bez GDP kao ulazne promenljive pokazuju nešto slabije rezultate u poređenju sa ostalim modelima (Tabela 24.), što je verovatno posledica činjenice

da je GDP u visokoj korelaciji sa tri ulazne promenljive (Tabela 22.). GRNN3d model je u odnosu na ostale modele, dobijene primenom i korelacione i VIF analize, pokazao najbolje rezultate i takođe, značajno bolje rezultate u odnosu na GRNN1 model. Kao što se može videti, GRNN3d model je kreiran sa najmanjim brojem (7) ulaznih promenljivih (EPCS, EPOS, GDP, EU, FFEC, NV, AL), što predstavlja redukciju od 37% u odnosu na GRNN1 model. Dijagrami aktuelne i predviđene emisije NO_x i *FA1.25* za GRNN3d model su prikazani na Slici 26. Kao što se može videti GRNN3d model pokazuje veoma dobro slaganje između aktuelnih i predviđenih vrednosti ($R^2=0.95$) (Slika 26.a.).



Slika 26. Rezultati GRNN3d modela za validacioni set podataka: a) poredenje aktuelne i predviđene NOx emisije, b) FA1.25 dijagram

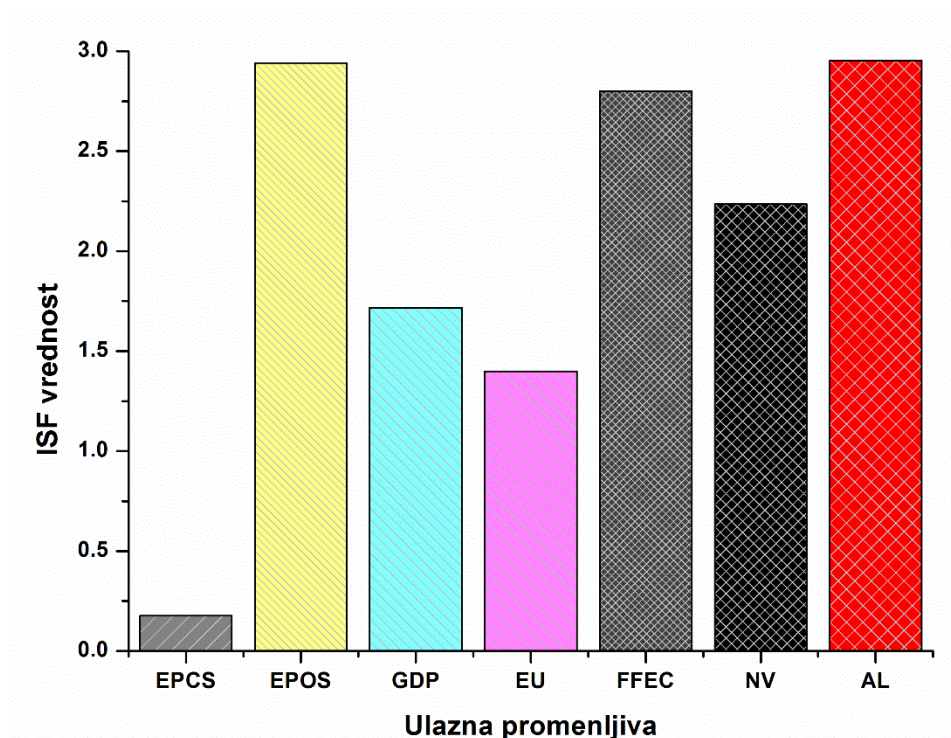
Kao što se može videti na Slici 27., predviđanje NO_x emisije u slučaju Slovačke pokazuje veliko odstupanje, sa vrednošću relativne greške od >35%, dok su maksimalne relativne greške za 2007. i 2008. predviđanja u slučaju Rumunije i Slovenije, vrlo zadovoljavajuća. Generalno se može uočiti da procenjena emisija u inventarima istočnoevropskih zemalja pokazuje visoku nesigurnost, što je verovatno i razlog za značajno odstupanje između predviđenih i aktuelnih emisija NO_x. Takođe, na osnovu dobijenih rezultata se može videti da su jednogodišnja predviđanja modela preciznija u odnosu na dvogodišnja predviđanja: za podatke iz 2007. godine 75% predviđanja imaju grešku do 7%, dok za 2008. godinu za istu frakciju greška iznosi 15% (Slika 27.a.). Sličan trend se može uočiti u slučaju pet država van Evrope (Slika 27.b.), sa opsegom grešaka koje su najveće u slučaju Kine i Rusije.



Slika 27. Relativne greške za a) evropske države b) države van Evrope

8.4. Rezultati analize osetljivosti (značajnosti) ulaznih promenljivih

U cilju određivanja značajnosti svake ulazne promenljive za GRNN3d model, vršena je analiza značajnosti. U slučaju GRNN, analiza značajnosti se vrši određivanjem individualnih faktora ujednačavanja (*ISF - Individual Smoothing Factors*) genetičkim algoritmom za svaku ulaznu promenljivu tokom procesa obučavanja. ISF može imati vrednosti od 0 do 3 i što je ISF veći, to je značajnost te promenljive za model veća [51]. Vrednosti ISF za svaku ulaznu promenljivu GRNN3d modela određene tokom procesa obučavanja mreže su prikazane na Slici 28.



Slika 28. Individualni faktori ujednačavanja (*ISF*) za ulazne promenljive u GRNN3d modelu

Na osnovu vrednosti *ISF* najznačajnije ulazne promenljive za model su EPOS, AL, FFEC i NV, dok GDP i EU imaju nižu, ali ipak značajan uticaj na izlaznu promenljivu modela. Dakle *ISF* vrednosti samo potvrđuju činjenicu da energetski sektor i sektor transporta imaju najznačajniji uticaj na emisiju azotovih oksida.

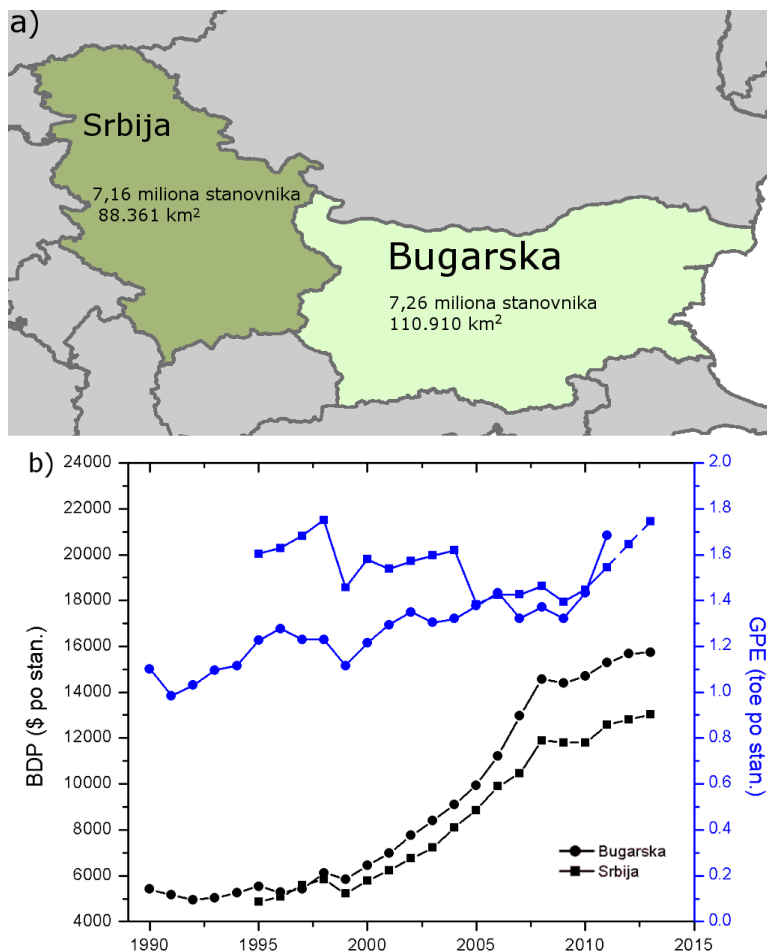
9. Predviđanje emisije gasova staklene bašte u Srbiji

Republika Srbija kao ne-Aneks država i potpisnica Okvirne konvencije UN o promeni klime i Kjoto protokola, ima obavezu da prati emisiju gasova staklene baste (GHG - *Greenhouse gases*) i sprovodi mere sa ciljem ispunjavanja kriterijuma definisanih Konvencijom i Kjoto protokolom [219]. U ovom poglavlju je prikazan razvoj i rezultati modela za predviđanje emisije gasova staklene bašte u Srbiji zasnovanog na neuronskim mrežama i ekonomskim i indikatorima održivog razvoja, kao ulaznih promenljivih modela. Da bi se razvio model zasnovan na ANN za Srbiju, potrebno je poznavanje određenih ekonomskih i industrijskih parametara, kao i indikatora održivog razvoja. Kako je Srbija tek u fazi implementacije i usklađivanja određenih zakonskih propisa sa zakonodavstvom Evropske unije, to je praćenje ovih parametara još u povoju. Kao moguće rešenje za razvoj modela za procenu emisije GHG koji bi bio primenljiv za Srbiju, u ovom radu je razvijen ANN model sa ulaznim podacima za Bugarsku, koja je po nekim karakteristikama vrlo slična Srbiji [220].

9.1. Ulazne i izlazne promenljive i podaci

Da bi se razvio ANN model za predviđanje emisije GHG u Srbiji, u ovom slučaju su zbog nedostatka podataka za Srbiju korišćeni podaci za Bugarsku, koja po svojim osnovnim karakteristikama, kao što su broj stanovnika, ekonomska i privredna razvijenost, proizvodnja i potrošnja energije, itd., najbližnja Srbiji (Slika 29.a.). Pomenute sličnosti ove dve države uslovile su i sličan trend promene najznačajnijih indikatora nacionalnog razvoja tokom vremenskog perioda za koji se model razvija. Na Slici 29.b. je prikazan trend promene tokom ispitivanog perioda za indikatore: bruto domaći proizvod (BDP) i godišnja potrošnja energije (GPE). U najznačajnije sektore za koje se smatra da najviše doprinose emisiji GHG ubrajaju se industrija i energetska sektor. U tom smislu su, u ovom slučaju, kao indikatori održivog razvoja koji su značajni za emisiju gasova staklene bašte odabrani: bruto domaći proizvod

po stanovniku, BDP – kao mera industrijskog i ekonomskog razvoja i godišnja proizvodnja energije po stanovniku, GPE – koja obuhvata sektor energetike.



Slika 29. Poređenje Srbije i Bugarske: a) položaj, broj stanovnika i površina, b) vrednost BDP-a i GPE-a [220]

Dakle, kao ulazne promenljive za razvoj ANN modela su odabrani BDP i GPE. Podaci za odabrane ulazne promenljive za Srbiju i Bugarsku su preuzeti iz baze podataka Svetske banke [206]. Podaci o nacionalnoj emisiji GHG za Bugarsku su preuzeti iz baze podataka Eurostata [221] (Tabela 25.), dok su dostupni podaci (1990. i 1998. godina) emisije GHG za Srbiju

preuzeti iz baze podataka Sekretarijata Ujedinjenih nacija za klimatske promene (UNFCCC) [222].

Tabela 25. Ulazni i izlazni podaci za razvoj ANN modela

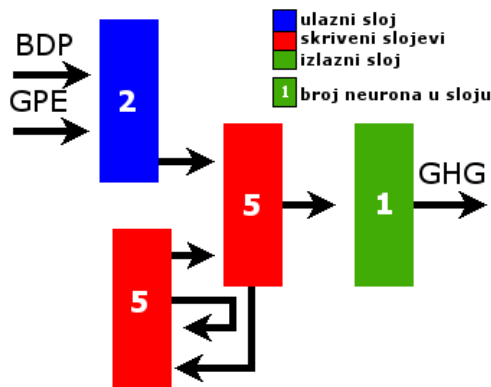
Skup podataka	Godina	Bugarska			Srbija		
		BDP [\$ po st.]	GPE [toe po st.]	GHG [t CO ₂ e po st.]	BDP	GPE	GHG
Obučavanje mreže	1990	5410	1.10	12.6	-	-	10.7
	1991	5168	0.98	10.1	-	-	-
	1992	4954	1.03	9.5	-	-	-
	1993	5037	1.10	9.4	-	-	-
	1994	5256	1.12	9.0	-	-	-
	1995	5543	1.23	9.1	4853	1.60	-
	1996	5281	1.28	9.1	5081	1.63	-
	1997	5409	1.23	8.7	5583	1.68	-
	1998	6123	1.23	8.2	5843	1.75	8.8
	1999	5849	1.12	7.4	5226	1.46	-
	2000	6445	1.21	7.3	5778	1.58	-
	2001	6976	1.29	7.9	6215	1.54	-
	2002	7752	1.35	7.6	6763	1.57	-
	2003	8396	1.30	8.3	7221	1.60	-
	2004	9097	1.32	8.2	8109	1.62	-
	2005	9947	1.38	8.3	8861	1.38	-
	2006	11229	1.43	8.4	9896	1.43	-
Validacija mreže	2007	12985	1.32	9.1	10454	1.43	-
	2008	14566	1.37	9.0	11893	1.46	-
	2009	14410	1.32	7.8	11806	1.39	-
	2010	14690	1.43	8.2	11805	1.45	-
	2011	15278	1.68	9.1	12572	1.54	-
	2012	-	-	-	12806	1.64 ^a	-
	2013	-	-	-	13020	1.74 ^a	-

^a ekstrapolacija rasta GPE

9.2. Razvoj ANN modela i rezultati

Za razvoj ANN modela za predviđanje emisije GHG u Srbiji, korišćena je rekurentna mreža, tip Džordan-Elman, čije su karakteristike objašnjene u poglavlju 4.2.3. Šematski prikaz

rekurentne mreže korišćene u ovom slučaju kao i broj neurona u slojevima je prikazan na Slici 30.



Slika 30. Rekurentna arhitektura (RNN) [220]

RNN model je razvijen sa podacima za Bugarsku za period 1990 - 2006. godine, dok je validacija modela vršena podacima za period 2007 - 2011. godine (Tabela 25.).

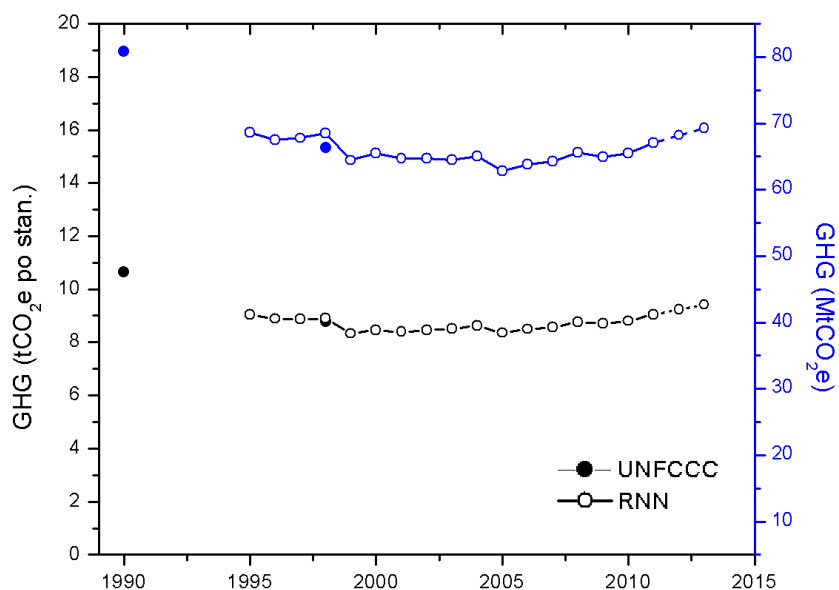
Da bi se procenile mogućnosti razvijenog modela za predviđanje GHG emisije, korišćena je srednja apsolutna procentualna greška (*MAPE*). *MAPE* vrednost dobijena na validacionom setu podataka za Bugarsku je iznosila 16%. Imajući u vidu da se nesigurnost procenjene emisije kod inventarskog pristupa kreće i do 30% [66], dobijena greška od 16% u ovom slučaju se može smatrati zadovoljavajućim rezultatom.

Nakon provere RNN modela na podacima za Bugarsku, pristupilo se predviđanju emisije GHG za Srbiju. Postignuti rezultati predviđanja GHG emisije za Srbiju za period od 1995. do 2013. godine su prikazani na Slici 31. Na osnovu dobijenih rezultata se može zaključiti da RNN model pokazuje dobro slaganje sa aktuelnim dostupnim GHG podacima za Srbiju (UNFCCC). Naime, odstupanje između RNN predviđanja i aktuelne vrednosti GHG emisije u 1998. godini je svega 3%.

Pored toga, i trend promene emisije GHG tokom ispitivanog perioda potvrđuje dobre performanse RNN modela, s obzirom na to da se mogu uočiti promene u emisiji GHG nastale usled određenih društveno-ekonomskih zbivanja, koja su uslovila promene u ekonomskom, industrijskom i energetsom sektoru, što je indirekto uticalo i na emisiju GHG. To se, na

primer, uočava u promeni emisije (opadanju) u 1999. godini. Pad emisije GHG u ovoj godini bi se mogao objasniti smanjenom privrednom i industrijskom aktivnošću usled bombardovanja Srbije. Pored toga, posle 2000. godine se takođe mogu videti izvesna opadanja ili stagnacija u GHG emisiji, što bi se moglo objasniti ekonomskim krizama.

Ukoliko bi trend povećanja proizvodnje energije bio nastavljen i u 2012. i 2013. godini, RNN model pokazuje da će i nivo GHG emisije u ovim godinama porasti i biti viši u odnosu na nivo emisije iz 1998. godine, ali da će i dalje biti znatno ispod nivoa emisije proračunatog za 1990. godinu.



Slika 31. Predviđanje GHG emisije za Srbiju [220]

10. Zaključak

Osnovni cilj istraživanja u okviru ove doktorske disertacije je bio razvoj modela za predviđanje emisija gasovitih zagađujućih materija u vazduh na nacionalnom nivou. Razvijeni modeli su zasnovani na veštačkim neuronskim mrežama (ANN) i primeni dostupnih industrijskih, ekonomskih, poljoprivrednih i indikatora održivog razvoja, kao ulaznih promenljivih za model. Dosadašnja primena veštačkih neuronskih mreža se do sada u najvećoj meri odnosila na modelovanje koncentracija zagađujućih materija u vazduhu, emisiju iz pojedinih sektora, kao i ostalih indikatora životne sredine. U okviru ove doktorske disertacije modelovane su emisije na nacionalnom nivou sledećih zagađujućih materija: amonijaka, nemetanskih isparljivih organskih jedinjenja, metana, azotnih oksida i gasova staklene bašte.

Osnovna karakteristika ANN modela je njihova sposobnost da se mogu primeniti za modelovanje složenih nelinearnih problema. U tom smislu se, u ovom radu, pošlo od pretpostavke da se predviđanje emisije zagađujućih materija u vazduh može posmatrati kao jedan nelinearan regresioni problem između prediktanta (u ovom slučaju emisije gasovitih zagađujućih materija na nacionalnom nivou) i prediktora (u ovom slučaju ekonomski, industrijski i ostali indikatori održivog razvoja).

Tokom razvoja modela zasnovanih na ANN, njihove performanse su proveravane primenom različitih indikatora performansi modela, kao i poređenjem sa rezultatima linearnih regresionih modela.

Za razvoj ANN modela za predviđanje nacionalne emisije amonijaka u SAD i EU primenjen je višeslojni perceptron (MLP), u ovom slučajnu troslojna neuronska mreža. Podaci za period 2000 - 2010. godine su korišćeni za razvoj modela. Odabrana MLP arhitektura je optimizovana primenom "proba – greška" pristupa, pri čemu je određen optimalan broj skrivenih neurona, odgovarajuća aktivaciona funkcija i algoritam za obučavanje mreže. Analiza glavnih komponenti (PCA) je primenjena radi redukcije multikolinearnosti između ulaznih promenljivih. Dobijeni rezultati nakon primene PCA su pokazali da MLP model

razvijen sa glavnim komponentama (PC) pokazuje bolje rezultate predviđanja u odnosu na MLP model koji je razvijen sa inicijalno odabranim (originalnim) ulaznim promenljivama. Rezultati razvijenog PCA – MLP – V modela su zatim upoređeni sa regresionim (PCR) modelom, koji je razvijen takođe sa glavnim komponentama dobijenim nakon primene PCA na originalnim ulaznim promenljivama i pokazali su da optimizovani model zasnovan na neuronskim mrežama ima značajno bolje rezultate u predviđanju u odnosu na konvencionalni PCR model. Kreirani ANN model je pokazao veoma zadovoljavajuće rezultate predviđanja, te se kao takav može koristiti kao jedan alternativni model u predviđanju emisije amonijaka na nacionalnom nivou.

S obzirom na to da je model zasnovan na ANN pristupu pokazao veoma dobre rezultate u predviđanju emisije amonijaka, isti pristup je primenjen i kod razvoja modela za predviđanje emisije nemetanskih isparljivih organskih jedinjenja (NMVOC) u Kini. Za razvoj modela su korišćeni dostupni podaci za devet evropskih država za period 2004 - 2012. godine. Za izbor najznačajnijih ulaznih promenljivih za model, kako za evropske države, tako i za Kinu, od potencijalnih deset, je korišćen *forward* pristup. Na osnovu dobijenih rezultata se može zaključiti da je primenjena tehnika za selekciju najznačajnijih ulaznih promenljivih uspešna, jer su rezultati predviđanja emisije NMVOC za EU države i Kinu slični, tj. vrednosti *MAPE* za EU države iznosi 8%, dok za Kinu ta vrednost iznosi 5%. Dakle, kreirani ANN model je pokazao veoma zadovoljavajuće rezultate predviđanja nacionalne emisije NMVOC za Kinu, koristeći samo pet ulaznih promenljivih, za razliku od inventarskog pristupa koji zahteva poznavanje preko stotinu specifičnih podataka emisionih faktora i stopa aktivnosti za određenu državu, industrijski proces, sektor emisije, itd.

Za predviđanje emisije metana na nacionalnom nivou u EU, pored standardne troslojne neuronske mreže sa propagacijom greške unazad (BPNN) koja je dala dobre rezultate u slučajevima modelovanja emisije amonijaka i NMVOC, primenjena je i generalna regresiona neuronska mreža (GRNN). Šest ulaznih promenljivih koje obuhvataju najznačajnije sektore izvore emisije metana su primenjene za razvoj modela. Nakon toga su njihove performanse upoređene sa odgovarajućim modelom baziranim na višestrukoj linearnoj regresiji (MLR).

Dobijeni rezultati su pokazali da ANN modeli imaju značajno bolje performanse u poređenju sa konvencionalnim MLR modelom.

U ovoj doktorskoj disertaciji je prikazan i razvoj i rezultati ANN modela za predviđanje emisije azotnih oksida (NO_x). Za razvoj modela su korišćeni podaci 17 država Evropske unije, SAD, Kine, Japana, Rusije i Indije, kao i GRNN arhitektura. GRNN model je optimizovan primenom dva pristupa selekcije ulaznih promenljivih (korelaciona i VIF analiza). Rezultati su pokazali da je proces optimizacije GRNN modela doprineo poboljšanju predikcionih mogućnosti u odnosu na GRNN model koji je bio kreiran sa svim inicijalno izabranim promenljivama, te se kao takav može koristiti za predviđanje emisije azotnih oksida na nacionalnom nivou.

Pored navedenih ANN modela, razvijenih za predviđanje emisije različitih zagađujućih materija, korišćenjem podataka, kako za EU države, tako i za druge države van evropskog kontinenta, u okviru ovog rada je razvijen i ANN model za predviđanje emisije gasova staklene bašte (GHG) u Srbiji. Za razvoj modela je primenjena rekurentna arhitektura mreže, dok su kao ulazne promenljive za model korišćeni podaci o bruto domaćem proizvodu i godišnjoj proizvodnji energije. Kako podaci ulaznih i izlazne promenljive za Srbiju nisu dostupni, za kreiranje modela su korišćeni podaci za Bugarsku, koja je po osnovnim indikatorima razvoja, u koje spadaju odabrane ulazne promenljive, najsličnija Srbiji. Model je pokazao dobro slaganje između aktuelne i modelom predviđene emisije GHG za 1998. godinu, sa odstupanjem od svega 3%. Takođe, analiza predviđenog trenda tokom ispitivanog perioda ukazuje na promene u emisiji GHG koje su bile posledica određenih društvenih okolnosti.

Na osnovu dobijenih rezultata i poređenja sa postojećim inventarskim pristupom za procenu emisije, može se zaključiti da kreirani ANN modeli daju dobra i pouzdana predviđanja svih modelovanih gasovitih zagađujućih materija, da se koristi drastično manji broj ulaznih promenljivih za predviđanje emisije – što znatno smanjuje vreme potrebno za njihovo prikupljanje, da su ulazne promenljive za ANN modele široko dostupne i da ne zahtevaju specifične metode i terenska istraživanja za njihovo određivanje. Sve navedeno čini ovaj

pristup ekonomičnijim i pokazuje da je veoma pogodan za primenu u državama u razvoju koje predstavljaju velike emitere zagađujućih materija u vazduh, poput Kine, Indije i Rusije, kod kojih je procena emisije inventarskim pristupom praćena izvesnom nesigurnošću, upravo zbog primene emisionih faktora koji su određeni na međunarodnom nivou i nisu specifični za datu državu i procese koji doprinose emisiji. Pored navedenih prednosti razvijenih ANN modela za predviđanje emisije gasovitih zagađujućih materija u vazduh u odnosu na inventarski pristup i klasične regresione modele, ANN modeli se mogu primeniti i za simulaciju različitih scenarija emisija, koja bi se zasnivala na promeni vrednosti ulaznih promenljivih koje bi bile uslovljene različitim zakonskim regulativama i merama primenjenim za promenu emisije zagađujućih materija. Imajući u vidu pomenute karakteristike ANN modela, moglo bi se zaključiti da ANN modeli mogu biti veoma efikasna sredstva u sprovođenju strategija održivog razvoja i politika zaštite životne sredine.

Buduća istraživanja u ovoj oblasti mogu obuhvatiti ispitivanje mogućnosti primene razvijenog ANN pristupa za modelovanje emisije iz određenih sektora, kao i mogućnost primene za modelovanje ostalih indikatora kvaliteta životne sredine, uz ispitivanje različitih tehnika za odabir najznačajnijih ulaznih promenljivih.

Literatura

- [1] A/RES/42/187 Report of the World Commission on Environment and Development 1987. <http://www.un.org/documents/ga/res/42/ares42-187.htm> (accessed December 11, 2015).
- [2] United Nations Millennium Development Goals. New York: United Nations; 2015.
- [3] UN. Transforming Our World by 2030: A New Agenda For Global Action. UN Summit to adopt the Post-2015 Development Agenda, 2015, p. 1–43.
- [4] UN. SDGs & Topics .:. Sustainable Development Knowledge Platform 2015. <https://sustainabledevelopment.un.org/topics/sustainabledevelopmentgoals> (accessed December 17, 2015).
- [5] Sustainable Development Solutions Network | Indicators and a Monitoring Framework for Sustainable Development Goals: Launching a data revolution for the SDGs n.d. <http://unsdsn.org/resources/publications/indicators/> (accessed December 9, 2015).
- [6] UNECE. Online Guidelines for the Application of Environmental Indicators 2016. <http://www.unece.org/env/indicators.html> (accessed January 28, 2016).
- [7] OECD. OECD Environmental Data and Indicators 2016. <http://www.oecd.org/environment/indicators-modelling-outlooks/data-and-indicators.htm> (accessed January 28, 2016).
- [8] SEPA. Agencija za zaštitu životne sredine - Ministarstvo poljoprivrede i zaštite životne sredine 2016. <http://www.sepa.gov.rs/index.php> (accessed January 28, 2016).
- [9] Edith Smeets and Rob Weterings. Environmental indicators: Typology and overview — European Environment Agency. Copenhagen: EEA (European Environment Agency); 1999.
- [10] Beusen AHW, Bouwman AF, Heuberger PSC, Van Drecht G, Van Der Hoek KW. Bottom-up uncertainty estimates of global ammonia emissions from global agricultural production systems. *Atmospheric Environment* 2008;42:6067–77. doi:10.1016/j.atmosenv.2008.03.044.

- [11] Behera SN, Sharma M, Aneja VP, Balasubramanian R. Ammonia in the atmosphere: A review on emission sources, atmospheric chemistry and deposition on terrestrial bodies. *Environmental Science and Pollution Research* 2013;20:8092–131.
- [12] Asman WAH, Sutton MA, Schjørring JK. Ammonia: Emission, atmospheric transport and deposition. *New Phytologist* 1998;139:27–48. doi:10.1046/j.1469-8137.1998.00180.x.
- [13] Manahan S. Environmental chemistry. *Environmental Chemistry* 2000;8th ed:743.
- [14] Jerrett M, Burnett RT, Ma R, Pope 3rd CA, Krewski D, Newbold KB, et al. Spatial analysis of air pollution and mortality in Los Angeles. *Epidemiology* 2005;16:727–36. doi:00001648-200511000-00004 [pii].
- [15] Anderson N, Strader R, Davidson C. Airborne reduced nitrogen: Ammonia emissions from agriculture and other sources. *Environment International* 2003;29:277–86. doi:10.1016/S0160-4120(02)00186-1.
- [16] Aneja VP, Blunden J, James K, Schlesinger WH, Knighton R, Gilliam W, et al. Ammonia assessment from agriculture: U.S. status and needs. *Journal of Environmental Quality* 2008;37:515–20. doi:10.2134/jeq2007.0002in.
- [17] Gruber N, Galloway JN. An Earth-system perspective of the global nitrogen cycle. *Nature* 2008;451:293–6. doi:10.1038/nature06592.
- [18] Erisman JW, Sutton M a., Galloway J, Klimont Z, Winiwarter W. How a century of ammonia synthesis changed the world. *Nature Geoscience* 2008;1:636–9. doi:10.1038/ngeo325.
- [19] Houlton BZ, Boyer E, Finzi A, Galloway J, Leach A, Liptzin D, et al. Intentional versus unintentional nitrogen use in the United States: Trends, efficiency and implications. *Biogeochemistry* 2013;114:11–23. doi:10.1007/s10533-012-9801-5.
- [20] European Environment Agency (EEA). Ammonia (NH₃) emissions — European Environment Agency 2016.
- [21] Bo Y, Cai H, Xie SD. Spatial and temporal variation of historical anthropogenic NMVOCs emission inventories in China. *Atmospheric Chemistry and Physics* 2008;8:7297–316. doi:10.5194/acp-8-7297-2008.

- [22] Atasoy E, Dögeroglu T, Kara S. The estimation of NMVOC emissions from an urban-scale wastewater treatment plant. *Water Research* 2004;38:3265–74. doi:10.1016/j.watres.2004.04.036.
- [23] Parra R, Jiménez P, Baldasano JM. Development of the high spatial resolution EMICAT2000 emission model for air pollutants from the north-eastern Iberian Peninsula (Catalonia, Spain). *Environmental Pollution* 2006;140:200–19. doi:10.1016/j.envpol.2005.07.021.
- [24] Niedojadlo A, Becker KH, Kurtenbach R, Wiesen P. The contribution of traffic and solvent use to the total NMVOC emission in a German city derived from measurements and CMB modelling. *Atmospheric Environment* 2007;41:7108–26. doi:10.1016/j.atmosenv.2007.04.056.
- [25] Li M, Zhang Q, Streets DG, He KB, Cheng YF, Emmons LK, et al. Mapping Asian anthropogenic emissions of non-methane volatile organic compounds to multiple chemical mechanisms. *Atmospheric Chemistry and Physics* 2014;14:5617–38. doi:10.5194/acp-14-5617-2014.
- [26] Klimont Z, Streets DG, Gupta S, Cofala J, Lixin F, Ichikawa Y. Anthropogenic emissions of non-methane volatile organic compounds in China. *Atmospheric Environment* 2002;36:1309–22. doi:10.1016/S1352-2310(01)00529-5.
- [27] Guo S, Hu M, Zamora ML, Peng J, Shang D, Zheng J, et al. Elucidating severe urban haze formation in China. *Proceedings of the National Academy of Sciences of the United States of America* 2014;111:17373–8. doi:10.1073/pnas.1419604111.
- [28] Qiu K, Yang L, Lin J, Wang P, Yang Y, Ye D, et al. Historical industrial emissions of non-methane volatile organic compounds in China for the period of 1980–2010. *Atmospheric Environment* 2014;86:102–12. doi:10.1016/j.atmosenv.2013.12.026.
- [29] Mauzerall DL, Sultan B, Kim N, Bradford DF. NO_x emissions from large point sources: Variability in ozone production, resulting health damages and economic costs. *Atmospheric Environment* 2005;39:2851–66. doi:10.1016/j.atmosenv.2004.12.041.
- [30] Pope CA, Burnett RT, Thun MJ, Calle EE, Krewski D, Ito K, et al. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA: The Journal of the American Medical Association* 2002;287:1132–41.

doi:10.1016/j.scitotenv.2011.03.017.

- [31] Jerrett M, Burnett RT, Pope CA, Ito K, Thurston G, Krewski D, et al. Long-term ozone exposure and mortality. *The New England Journal of Medicine* 2009;360:1085–95. doi:10.1056/NEJMoa0803894.
- [32] Tong DQ, Lamsal L, Pan L, Ding C, Kim H, Lee P, et al. Long-term NO_x trends over large cities in the United States during the great recession: Comparison of satellite retrievals, ground observations, and emission inventories. *Atmospheric Environment* 2015;107:70–84. doi:10.1016/j.atmosenv.2015.01.035.
- [33] Reşitoğlu İA, Altinişik K, Keskin A. The pollutant emissions from diesel-engine vehicles and exhaust aftertreatment systems. *Clean Technologies and Environmental Policy* 2014;17:15–27. doi:10.1007/s10098-014-0793-9.
- [34] European Environment Agency (EEA). Nitrogen oxides (NO_x) emissions — European Environment Agency 2011. <http://www.eea.europa.eu/data-and-maps/indicators/eea-32-nitrogen-oxides-nox-emissions-1/assessment.2010-08-19.0140149032-3> (accessed February 5, 2016).
- [35] EPA. State and County Emission Summaries | Air Emission Sources | US EPA 2011. http://www3.epa.gov/cgi-bin/broker?polchoice=NOX&_debug=0&_service=data&_program=dataprog.national_1.sas (accessed February 5, 2016).
- [36] Ohara T, Akimoto H, Kurokawa J, Horii N, Yamaji K, Yan X, et al. An Asian emission inventory of anthropogenic emission sources for the period 1980–2020. *Atmospheric Chemistry and Physics* 2007;7:4419–44. doi:10.5194/acp-7-4419-2007.
- [37] Klimont Z, Cofala J, Xing J, Wei W, Zhang C, Wang S, et al. Projections of SO₂, NO_x and carbonaceous aerosols emissions in Asia. *Tellus, Series B: Chemical and Physical Meteorology* 2009;61:602–17. doi:10.1111/j.1600-0889.2009.00428.x.
- [38] You CF, Xu XC. Coal combustion and its pollution control in China. *Energy* 2010;35:4467–72. doi:10.1016/j.energy.2009.04.019.
- [39] Oliveira MLM, Silva CM, Moreno-Tost R, Farias TL, Jiménez-López A, Rodríguez-Castellón E. Modelling of NO_x emission factors from heavy and light-duty vehicles equipped with advanced aftertreatment systems. *Energy Conversion and Management* 2011;52:2945–51. doi:10.1016/j.enconman.2011.02.025.

- [40] US EPA CCD. Methane Emissions 2016.
- [41] IPCC. Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Geneva, Switzerland: 2007.
- [42] Anderson B, Barlett K, Frohling S, Hayhoe K, Jenkins J, Salas W. Methane and Nitrous Oxide Emissions From Natural Sources. 2010. doi:EPA 430-R-10-001.
- [43] Ishii K, Furuichi T. Estimation of methane emission rate changes using age-defined waste in a landfill site. *Waste Management (New York, NY)* 2013;33:1861–9. doi:10.1016/j.wasman.2013.05.011.
- [44] European Union. Council Directive 1999/31/EC of 26 April 1999 on the Landfill Waste. 1999.
- [45] OECD. ENVIRONMENTAL OUTLOOK TO 2050: The consequences of Inaction Key Findings on Health and Environment 2012. <http://www.oecd.org/general/searchresults/?q=deth+cose+air+pollution&cx=012432601748511391518:xzeadub0b0a&cof=FORID:11&ie=UTF-8> (accessed February 9, 2016).
- [46] UNECE. Air Pollution - Air Pollution - Environmental Policy - UNECE 2016. <http://www.unece.org/env/lrtap/welcome.html> (accessed February 10, 2016).
- [47] EEA. EMEP/CORINAIR Emission Inventory Guidebook - 2007 — European Environment Agency 2007.
- [48] Dröge R, Kuenen JJP, Pulles MPJ, Heslinga DC. The revised EMEP/EEA Guidebook compared to the country specific inventory system in the Netherlands. *Atmospheric Environment* 2010;44:3503–10. doi:10.1016/j.atmosenv.2010.06.020.
- [49] UNFCCC. United Nations Framework Convention on Climate Change 2016. <http://unfccc.int/2860.php> (accessed February 10, 2016).
- [50] Antanasijević D, Pocajt V, Popovic I, Redzic N, Ristic M. The forecasting of municipal waste generation using artificial neural networks and sustainability indicators. *Sustainability Science* 2013;8:37–46. doi:10.1007/s11625-012-0161-9.

- [51] Antanasijević DZ, Pocajt V V, Povrenović DS, Ristić MĐ, Perić-Grujić AA. PM(10) emission forecasting using artificial neural networks and genetic algorithm input variable optimization. *The Science of the Total Environment* 2013;443:511–9. doi:10.1016/j.scitotenv.2012.10.110.
- [52] Antanasijević DZ, Ristić MĐ, Perić-Grujić AA, Pocajt V V. Forecasting GHG emissions using an optimized artificial neural network model based on correlation and principal component analysis. *International Journal of Greenhouse Gas Control* 2014;20:244–53. doi:10.1016/j.ijggc.2013.11.011.
- [53] Savić MA, Redžić ND, Jovanović JM, Jovanović MB. Quality assurance of the Serbian national E-PRTR register reported data for large combustion plants. *Hemijška Industrija* 2012;66:95–106.
- [54] Office of Air Quality Planning and Standards UE. Compilation of air pollutant emission factors (AP-42) Volume 1: point and area sources 1995. <http://www3.epa.gov/ttn/chief/ap42/index.html#toc> (accessed February 11, 2016).
- [55] Olivier JGJ, Bouwman AF, Van Der Maas CWM, Berdowski JJM. Emission database for global atmospheric research (EDGAR): Version 2.0. *Studies in Environmental Science*, vol. 65, 1995, p. 651–9. doi:10.1016/S0166-1116(06)80262-1.
- [56] Olivier, J.G.J., Bouwman, A.F., van der Maas, C.W.M. B, J.J.M., Veldt, C., Bloos, J.P.J., Visschedijk, A.J.H. Z, P.Y.J., Haverlag JL. Description of EDGAR Version 2.0: A set of global emission inventories of greenhouse gases and ozone-depleting substances for all anthropogenic and most natural sources on a per country basis and on 1 degree x 1 degree grid - PBL Netherlands Environmental. Bilthoven, the Netherlands: 1996. doi:Report No.771060002.
- [57] EDGAR. EDGAR - Emission Database for Global Atmospheric Research 2016. <http://edgar.jrc.ec.europa.eu/methodology.php#12sou> (accessed February 12, 2016).
- [58] Janssens-Maenhout G., Dentener F., Van Aardenne J., Monni S., Pagliari V., Orlandini L., Klimont Z., Kurokawa J., Akimoto H., Ohara T., Wankmueller R., Battye B., Grano D., Zuber A. KT. EDGAR-HTAP: a Harmonized Gridded Air Pollution Emission Dataset Based on National Inventories, Ispra (Italy): European Commission Publications Office. 2012. doi:ISBN 978-92-79-23122-0, ISSN 1831-9424.

- [59] Amstel A Van, Olivier J, Janssen L. Analysis of differences between national inventories and an Emissions Database for Global Atmospheric Research (EDGAR). *Environmental Science & Policy* 1999;2:275–93. doi:10.1016/S1462-9011(99)00019-2.
- [60] IPCC. 2006 IPCC Guidelines for National Greenhouse Gas Inventories 2006. <http://www.ipcc-nggip.iges.or.jp/public/2006gl/> (accessed February 13, 2016).
- [61] Olivier JGJ, Bouwman AF, Berdowski JJM, Veldt C, Bloos JPJ, Visschedijk AJH, et al. Sectoral emission inventories of greenhouse gases for 1990 on a per country basis as well as on 1°×1°. *Environmental Science & Policy* 1999;2:241–63. doi:10.1016/S1462-9011(99)00027-1.
- [62] PBL. Data sources and methodology - the Netherlands Environmental Assessment Agency (PBL) 2016. <http://themasites.pbl.nl/tridion/en/themasites/edgar/documentation/data-sources-and-methodology/index-2.html> (accessed February 13, 2016).
- [63] Baldasano JM, Güereca LP, López E, Gassó S, Jimenez-Guerrero P. Development of a high-resolution (1km×1km, 1h) emission model for Spain: The High-Resolution Modelling Emission System (HERMES). *Atmospheric Environment* 2008;42:7215–33. doi:10.1016/j.atmosenv.2008.07.026.
- [64] EEA (European Environment Agency). EMEP/EEA air pollutant emission inventory guidebook - 2013 — European Environment Agency 2013.
- [65] Winther M, Boll Illerup J, Fenhann J, Kilde N. The Danish CORINAIR Inventories. 1999.
- [66] EEA (European Environment Agency). EMEP/EEA air pollutant emission inventory guidebook - 2009 — European Environment Agency 2009.
- [67] IPCC. Conceptual Basis for Uncertainty Analysis. IPCC Good Practice Guidance and Uncertainty Management in National Greenhouse Gas Inventories 2000:A1.1–A1.22.
- [68] Schöpp W, Klimont Z, Suutari R, Cofala J. Uncertainty analysis of emission estimates in the RAINS integrated assessment model. *Environmental Science and Policy* 2005;8:601–13. doi:10.1016/j.envsci.2005.06.008.

- [69] Dennis R, Fox T, Fuentes M, Gilliland A, Hanna S, Hogrefe C, et al. A framework for evaluating regional-scale numerical photochemical modeling systems. *Environmental Fluid Mechanics* 2010;10:471–89. doi:10.1007/s10652-009-9163-2.
- [70] L.Ciancarella, G. Briganti, G. Calori, A. Cappelletti, I. Cionni, M. Costa, G. Cremona, I. D’Elia, M. D’Isidoro, S. Finardi, L. Mauri, M. Mircea, G. Pace, A. Piersanti, S. Racalbuto, P. Radice, G. Righini, G. Vialetto, L.Vitali GZ. National Italian integrated atmospheric model on air pollution: Sensitivity to emission inventory. 14th conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Kos, Greece: 2011, p. 277–81.
- [71] Russell A, Dennis R. NARSTO critical review of photochemical models and modeling. *Atmospheric Environment* 2000;34:2283–324. doi:10.1016/S1352-2310(99)00468-9.
- [72] Hanna SR, Lu Z, Christopher Frey H, Wheeler N, Vukovich J, Arunachalam S, et al. Uncertainties in predicted ozone concentrations due to input uncertainties for the UAM-V photochemical grid model applied to the July 1995 OTAG domain. *Atmospheric Environment* 2001;35:891–903. doi:10.1016/S1352-2310(00)00367-8.
- [73] Napelenok SL, Foley KM, Kang DW, Mathur R, Pierce T, Rao ST. Dynamic evaluation of regional air quality model’s response to emission reductions in the presence of uncertain emission inventories. *Atmospheric Environment* 2011;45:4091–8.
- [74] IIASA. The GAINS Model - IIASA 2016. <http://www.iiasa.ac.at/web/home/research/researchPrograms/GAINS.en.html> (accessed February 14, 2016).
- [75] Basheer IA, Hajmeer M. Artificial neural networks: Fundamentals, computing, design, and application. *Journal of Microbiological Methods* 2000;43:3–31. doi:10.1016/S0167-7012(00)00201-3.
- [76] Kalogirou SA. Artificial neural networks in renewable energy systems applications: a review. *Renewable & Sustainable Energy Reviews* 2001;5:373–401. doi:10.1016/S1364-0321(01)00006-5.
- [77] Milenković S. Vestačke neuronske mreže. Beograd: Zadužbina Andrejević; 1997.
- [78] Dayhoff JE, DeLeo JM. Artificial neural networks: opening the black box. *Cancer*

2001;91:1615–35.

- [79] Mareš T, Janouchová E, Kučerová A. Artificial neural networks in the calibration of nonlinear mechanical models. *Advances in Engineering Software* 2016;95:68–81. doi:10.1016/j.advengsoft.2016.01.017.
- [80] Zavrtanik N, Prosen J, Tušar M, Turk G. The use of artificial neural networks for modeling air void content in aggregate mixture. *Automation in Construction* 2016;63:155–61. doi:10.1016/j.autcon.2015.12.009.
- [81] Haykin Simon. *Neural Networks: A Comprehensive Foundation*. New York: Macmillan College Publishing 9780023527616 - Cheshire Book Centre; 1994.
- [82] Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators. *Neural Networks* 1989;2:359–66. doi:10.1016/0893-6080(89)90020-8.
- [83] Hornik K. Approximation capabilities of multilayer feedforward networks. *Neural Networks* 1991;4:251–7. doi:10.1016/0893-6080(91)90009-T.
- [84] Özesmi SL, Tan CO, Özesmi U. Methodological issues in building, training, and testing artificial neural networks in ecological applications. *Ecological Modelling* 2006;195:83–93. doi:10.1016/j.ecolmodel.2005.11.012.
- [85] Razavi S, Tolson B a. A new formulation for feedforward neural networks. *IEEE Transactions on Neural Networks / a Publication of the IEEE Neural Networks Council* 2011;22:1588–98. doi:10.1109/TNN.2011.2163169.
- [86] Kalogirou SA. Artificial intelligence for the modeling and control of combustion processes: A review. *Progress in Energy and Combustion Science* 2003;29:515–66. doi:10.1016/S0360-1285(03)00058-3.
- [87] Russo A, Raischel F, Lind PG. Air quality prediction using optimal neural networks with stochastic variables. *Atmospheric Environment* 2013;79:822–30. doi:10.1016/j.atmosenv.2013.07.072.
- [88] Barr A, Feigenbaum EA. *The Handbook of Artificial Intelligence*. vol. 3. 1981. doi:10.1016/0898-1221(87)90239-2.

- [89] Rumelhart DE, Hinton GE, Williams RJ. Learning internal representations by error propagation 1986:318–62.
- [90] Bishop CM. Neural Networks for Pattern Recognition. Clarendon Press, Oxford; 1995. doi:10.2307/2965437.
- [91] Bishop CM. Pattern Recognition and Machine Learning. Springer; 2006. doi:10.1117/1.2819119.
- [92] Ul-Saufie AZ, Yahaya AS, Ramli NA, Rosaida N, Hamid HA. Future daily PM10 concentrations prediction by combining regression models and feedforward backpropagation models with principle component analysis (PCA). Atmospheric Environment 2013;77:621–30. doi:10.1016/j.atmosenv.2013.05.017.
- [93] Gardner M., Dorling S. Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. Atmospheric Environment 1998;32:2627–36. doi:10.1016/S1352-2310(97)00447-0.
- [94] Pocajt V V. Razvoj modela atmosferske disperzije primenom neuronskih mreža. Doktorska disertacija, Tehnološko-metalurški fakultet, Univerzitet u Beogradu, Beograd, Srbija, 1999.
- [95] Zhang GP, Patuwo EB, Michael Y. H. Forecasting with artificial neural networks: The state of the art. International Journal of Forecasting 1998;14:35–62. doi:10.1016/S0169-2070(97)00044-7.
- [96] Antanasijević DZ. Modelovanje indikatora kvaliteta životne sredine primenom veštačkih neuronskih mreža. Doktorska disertacija, Tehnološko-metalurški fakultet, Univerzitet u Beogradu, Beograd, Srbija, 2014.
- [97] Ramesh J, Vanathi PT, Gunavathi K. Fault Classification in Phase-Locked Loops Using Back Propagation Neural Networks. ETRI Journal 2008;30:546–54. doi:10.4218/etrij.08.0108.0133.
- [98] Bolanča T, Cerjan-Stefanović Š, Ukić Š, Rogošić M, Luša M. Application of different training methodologies for the development of a back propagation artificial neural network retention model in ion chromatography. Journal of Chemometrics 2008;22:106–13. doi:10.1002/cem.1096.

- [99] Jakeman AJ, Letcher RA, Norton JP. Ten iterative steps in development and evaluation of environmental models. *Environmental Modelling and Software* 2006;21:602–14. doi:10.1016/j.envsoft.2006.01.004.
- [100] Cherkassky V, Lari-Najafi H. Data representation for diagnostic neural networks. *IEEE Expert* 1992;7:43–53. doi:10.1109/64.163672.
- [101] Giordano F, Rocca M La, Perna C. Input Variable Selection in Neural Network Models. *Communications in Statistics - Theory and Methods* 2014;43:735–50. doi:10.1080/03610926.2013.804567.
- [102] May RJ, Maier HR, Dandy GC, Fernando TMKG. Non-linear variable selection for artificial neural networks using partial mutual information. *Environmental Modelling & Software* 2008;23:1312–26. doi:10.1016/j.envsoft.2008.03.007.
- [103] Guyon I, Elisseeff A. An introduction to variable and feature selection. *Journal of Machine Learning Research* 2003;3:1157–82. doi:10.1162/153244303322753616.
- [104] Tripathy M. Power transformer differential protection using neural network Principal Component Analysis and Radial Basis Function Neural Network. *Simulation Modelling Practice and Theory* 2010;18:600–11. doi:10.1016/j.simpat.2010.01.003.
- [105] Maier HR, Jain A, Dandy GC, Sudheer KP. Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions. *Environmental Modelling & Software* 2010;25:891–909. doi:10.1016/j.envsoft.2010.02.003.
- [106] Bowden GJ, Dandy GC, Maier HR. Input determination for neural network models in water resources applications. Part 1—background and methodology. *Journal of Hydrology* 2005;301:75–92. doi:10.1016/j.jhydrol.2004.06.021.
- [107] Bowden GJ, Maier HR, Dandy GC. Input determination for neural network models in water resources applications. Part 2. Case study: Forecasting salinity in a river. *Journal of Hydrology* 2005;301:93–107. doi:10.1016/j.jhydrol.2004.06.020.
- [108] May R, Dandy G, Maier H. Artificial Neural Networks - Methodological Advances and Biomedical Applications. Croatia: InTech; 2011. doi:10.5772/16004.
- [109] Castelletti A, Galelli S, Restelli M, Soncini-Sessa R. Data-driven dynamic emulation

modelling for the optimal management of environmental systems. *Environmental Modelling and Software* 2012;34:30–43. doi:10.1016/j.envsoft.2011.09.003.

- [110] Jain AK, Mao J, Mohiuddin KM. Artificial neural networks: a tutorial. *Computer* 1996;29:31–44. doi:10.1109/2.485891.
- [111] Wu W, Dandy GC, Maier HR. Protocol for developing ANN models and its application to the assessment of the quality of the ANN model development process in drinking water quality modelling. *Environmental Modelling & Software* 2014;54:108–27. doi:10.1016/j.envsoft.2013.12.016.
- [112] Karamouz M, Razavi S, Araghinejad S. Long-lead seasonal rainfall forecasting using time-delay recurrent neural networks: a case study. *Hydrological Processes* 2008;22:229–41.
- [113] Ward Systems Group Inc. Advanced Neural Network and Genetic Algorithm Software 2016. <http://www.wardsystems.com/neuroshell2.asp#architectures> (accessed February 28, 2016).
- [114] Viotti P, Liuti G, Di Genova P. Atmospheric urban pollution: Applications of an artificial neural network (ANN) to the city of Perugia. *Ecological Modelling* 2002;148:27–46. doi:10.1016/S0304-3800(01)00434-3.
- [115] Ali Abdoli M, Falah Nezhad M, Salehi Sede R, Behboudian S. Longterm forecasting of solid waste generation by the artificial neural networks. *Environmental Progress & Sustainable Energy* 2012;31:628–36. doi:10.1002/ep.10591.
- [116] Lu WZ, Wang WJ, Wang XK, Xu ZB, Leung AYT. Using improved neural network model to analyze RSP, NOX and NO2 levels in urban air in Mong Kok, Hong Kong. *Environmental Monitoring and Assessment* 2003;87:235–54. doi:10.1023/A:1024819309108.
- [117] Palani S, Liong SY, Tkalich P. An ANN application for water quality forecasting. *Marine Pollution Bulletin* 2008;56:1586–97. doi:10.1016/j.marpolbul.2008.05.021.
- [118] Sharma N, Chaudhry KK, Chalapati-Rao C V. Vehicular pollution modeling using artificial neural network technique: A review. *Journal of Scientific & Industrial Research* 2005;64:637–47.

- [119] Boznar M, Lesjak M, Mlakar P. A neural network-based method for short-term predictions of ambient SO₂ concentrations in highly polluted industrial areas of complex terrain. *Atmospheric Environment Part B, Urban Atmosphere* 1993;27:221–30. doi:10.1016/0957-1272(93)90007-S.
- [120] Yi J, Prybutok VR. A neural network model forecasting for prediction of daily maximum ozone concentration in an industrialized urban area. *Environmental Pollution* 1996;92:349–57. doi:10.1016/0269-7491(95)00078-X.
- [121] Comrie AC. Comparing Neural Networks and Regression Models for Ozone Forecasting. *Journal of the Air & Waste Management Association* 1997;47:653–63. doi:10.1080/10473289.1997.10463925.
- [122] Cai M, Yin Y, Xie M. Prediction of hourly air pollutant concentrations near urban arterials using artificial neural network approach. *Transportation Research Part D: Transport and Environment* 2009;14:32–41. doi:10.1016/j.trd.2008.10.004.
- [123] Grivas G, Chaloulakou A. Artificial neural network models for prediction of PM₁₀ hourly concentrations, in the Greater Area of Athens, Greece. *Atmospheric Environment* 2006;40:1216–29. doi:10.1016/j.atmosenv.2005.10.036.
- [124] Kukkonen J, Partanen L, Karppinen A, Ruuskanen J, Junninen H, Kolehmainen M, et al. 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* 2003;37:4539–50. doi:10.1016/S1352-2310(03)00583-1.
- [125] Fernando HJS, Mammarella MC, Grandoni G, Fedele P, Di Marco R, Dimitrova R, et al. Forecasting PM₁₀ in metropolitan areas: Efficacy of neural networks. *Environmental Pollution (Barking, Essex : 1987)* 2012;163:62–7. doi:10.1016/j.envpol.2011.12.018.
- [126] Singh KP, Gupta S, Kumar A, Shukla SP. Linear and nonlinear modeling approaches for urban air quality prediction. *Science of the Total Environment* 2012;426:244–55. doi:10.1016/j.scitotenv.2012.03.076.
- [127] Hájek P, Olej V. Ozone prediction on the basis of neural networks, support vector regression and methods with uncertainty. *Ecological Informatics* 2012;12:31–42. doi:10.1016/j.ecoinf.2012.09.001.

- [128] Lim Y, Moon Y-S, Kim T-W. Artificial neural network approach for prediction of ammonia emission from field-applied manure and relative significance assessment of ammonia emission factors. *European Journal of Agronomy* 2007;26:425–34. doi:10.1016/j.eja.2007.01.008.
- [129] Tronci S, Baratti R, Servida A. Monitoring pollutant emissions in a power plant through neural network. *Neurocomputing* 2002;43:3–15. doi:10.1016/S0925-2312(01)00617-8.
- [130] Kesgin U. Genetic algorithm and artificial neural network for engine optimisation of efficiency and NOx emission. *Fuel* 2004;83:885–95. doi:10.1016/j.fuel.2003.10.025.
- [131] Zhou H, Cen K, Fan J. Modeling and optimization of the NOx emission characteristics of a tangentially fired boiler with artificial neural networks. *Energy* 2004;29:167–83. doi:10.1016/j.energy.2003.08.004.
- [132] Zheng LZL, Yu SYS, Yu MYM. Monitoring NOx Emissions from Coal Fired Boilers Using Generalized Regression Neural Network. 2nd IEEE International Conference on Bioinformatics and Biomedical Engineering 2008:1916–9. doi:10.1109/ICBBE.2008.808.
- [133] Smrekar J, Potočnik P, Senegačnik a. Multi-step-ahead prediction of NOx emissions for a coal-based boiler. *Applied Energy* 2013;106:89–99. doi:10.1016/j.apenergy.2012.10.056.
- [134] Sözen A, Gülseven Z, Arcaklioğlu E. Estimation of GHG emissions in turkey using energy and economic indicators. *Energy Sources, Part A: Recovery, Utilization and Environmental Effects* 2009;31:1141–59. doi:10.1080/15567030802089086.
- [135] Antanasijević D, Pocajt V, Ristić M, Perić-Grujić A. Modeling of energy consumption and related GHG (greenhouse gas) intensity and emissions in Europe using general regression neural networks. *Energy* 2015;84:816–24. doi:10.1016/j.energy.2015.03.060.
- [136] Agirre-Basurko E, Ibarra-Berastegi G, Madariaga I. Regression and multilayer perceptron-based models to forecast hourly O3 and NO2 levels in the Bilbao area. *Environmental Modelling & Software* 2006;21:430–46. doi:10.1016/j.envsoft.2004.07.008.
- [137] Szoplik J. Forecasting of natural gas consumption with artificial neural networks.

Energy 2015;85:208–20. doi:10.1016/j.energy.2015.03.084.

- [138] Vogl TP, Mangis JK, Rigler AK, Zink WT, Alkon DL. Accelerating the convergence of the back-propagation method. *Biological Cybernetics* 1988;59:257–63. doi:10.1007/BF00332914.
- [139] Howard Demuth MB. *Neural Network Toolbox For Use with MATLAB, User's Guide*. Natick, MA, USA: The MathWorks, Inc.; 2004.
- [140] Patrick van der Smagt P. Minimisation methods for training feedforward neural networks. *Neural Networks* 1994;7:1–11. doi:10.1016/0893-6080(94)90052-3.
- [141] Kollias S, Anastassiou D. An adaptive least squares algorithm for the efficient training of artificial neural networks. *IEEE Transactions on Circuits and Systems* 1989;36:1092–101. doi:10.1109/31.192419.
- [142] Ward Systems Group Inc. *Neuroshell 2*. Ward Systems Group Inc, Frederick 2008.
- [143] Roush WB, Cravener TL. Artificial neural network prediction of amino acid levels in feed ingredients. *Poultry Science* 1997;76:721–7.
- [144] Specht DF. A general regression neural network. *IEEE Transactions on Neural Networks / a Publication of the IEEE Neural Networks Council* 1991;2:568–76. doi:10.1109/72.97934.
- [145] Du C, Tang D, Zhou J, Wang H, Shaviv A. Prediction of nitrate release from polymer-coated fertilizers using an artificial neural network model. *Biosystems Engineering* 2008;99:478–86. doi:10.1016/j.biosystemseng.2007.12.003.
- [146] Zhou Y, Cheng S, Chen D, Lang J, Zhao B, Wei W. A new statistical approach for establishing high-resolution emission inventory of primary gaseous air pollutants. *Atmospheric Environment* 2014;94:392–401. doi:10.1016/j.atmosenv.2014.05.047.
- [147] Zhou Q, Jiang H, Wang J, Zhou J. A hybrid model for PM_{2.5} forecasting based on ensemble empirical mode decomposition and a general regression neural network. *Science of the Total Environment* 2014;496:264–74. doi:10.1016/j.scitotenv.2014.07.051.
- [148] Antanasijević D, Pocajt V, Perić-Grujić A, Ristić M. Modelling of dissolved oxygen

in the Danube River using artificial neural networks and Monte Carlo Simulation uncertainty analysis. *Journal of Hydrology* 2014;519:1895–907. doi:10.1016/j.jhydrol.2014.10.009.

- [149] Chen QCQ, Chai WCW, Qiao JQJ. Modeling of wastewater treatment process using recurrent neural network. *Intelligent Control and Automation (WCICA)*, 2010 8th World Congress on 2010:5872–6. doi:10.1109/WCICA.2010.5554543.
- [150] Qiao JF, Yang WW, Yuan MZ. Recurrent High Order Neural Network Modeling for Wastewater Treatment Process. *Journal of Computers* 2011;6:1570–7. doi:10.4304/jcp.6.8.1570-1577.
- [151] Al-Alawi SM, Abdul-Wahab SA, Bakheit CS. Combining principal component regression and artificial neural networks for more accurate predictions of ground-level ozone. *Environmental Modelling and Software* 2008;23:396–403. doi:10.1016/j.envsoft.2006.08.007.
- [152] Çamdevýren H, Demýr N, Kanik A, Keskýn S. Use of Principal Component Scores in Multiple Linear Regression Models for Prediction of Chlorophyl-a Reservoirs. *Ecological Modelling* 2005;181:581–9. doi:10.1016/j.ecolmodel.2004.06.043.
- [153] Helena B, Pardo R, Vega M, Barrado E, Fernandez JM, Fernandez L. Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis. *Water Research* 2000;34:807–16. doi:10.1016/S0043-1354(99)00225-0.
- [154] Noori R, Karbassi A, Salman Sabahi M. Evaluation of PCA and Gamma test techniques on ANN operation for weekly solid waste prediction. *Journal of Environmental Management* 2010;91:767–71. doi:10.1016/j.jenvman.2009.10.007.
- [155] Jolliffe IT. Principal Component Analysis. *Journal of the American Statistical Association* 2002;98:487. doi:10.1007/b98835.
- [156] Pallant J. *SPSS survival manual: a step by step guide to data analysis using SPSS*. 2013.
- [157] Tabachnick BGLSF. *Using Multivariate Statistics (5th Edition)*: Amazon.com: Books. Boston: University of Chicago Press.; 2007.

- [158] Shaw PJA. *Multivariate statistics for the Environmental Sciences*. New York: 2003.
- [159] Tabachnick BG, Fidell LS. *Tabachnick, Fidell_2001.pdf*. *Using Multivariate Statistics* 2001:1008.
- [160] Bartlett MS. A note on the multiplying factors for various χ^2 approximations. *Journal of the Royal Statistical Society* 1954;16:296–8. doi:10.2307/2984057.
- [161] Kaiser HF. An index of factorial simplicity. *Psychometrika* 1974;39:31–6. doi:10.1007/BF02291575.
- [162] Sousa S, Martins F, Alvimferraz M, Pereira M. Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations. *Environmental Modelling & Software* 2007;22:97–103. doi:10.1016/j.envsoft.2005.12.002.
- [163] Peres-Neto PR, Jackson DA, Somers KM. How many principal components? stopping rules for determining the number of non-trivial axes revisited. *Computational Statistics and Data Analysis* 2005;49:974–97. doi:10.1016/j.csda.2004.06.015.
- [164] Cattell RB. The Scree Test For The Number Of Factors. *Multivariate Behavioral Research* 1966;1:245–76. doi:10.1207/s15327906mbr0102_10.
- [165] Horn JL, Engstrom R. Cattell's scree test in relation to Bartlett's chi-square test and other observations on the number of factors problem. *Multivariate Behavioral Research* 1979;14:283–300. doi:10.1207/s15327906mbr1403_1.
- [166] Horn JL. A rationale and test for the number of factors in factor analysis. *Psychometrika* 1965;30:179–85. doi:10.1007/BF02289447.
- [167] Thurstone LL. *Multiple Factor Analysis*. Chicago: University of Chicago Press; 1947.
- [168] IBM Corp. Released. *IBM SPSS Statistics for Windows, Version 19.0*. NY: Armonk; 2010.
- [169] Zhang Z. Quotient correlation: A sample based alternative to Pearson's correlation. *Annals of Statistics* 2008;36:1007–30. doi:10.1214/009053607000000866.

- [170] Cohen J. Statistical power analysis for the behavioral sciences. *Statistical Power Analysis for the Behavioral Sciences* 1988;2nd:567. doi:10.1234/12345678.
- [171] Zhang W-Y, Wei Z-W, Wang B-H, Han X-P. Measuring mixing patterns in complex networks by Spearman rank correlation coefficient. *Physica A: Statistical Mechanics and Its Applications* 2016. doi:10.1016/j.physa.2016.01.056.
- [172] O'Brien RM. A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality & Quantity* 2007;41:673–90. doi:10.1007/s11135-006-9018-6.
- [173] Alin A. Multicollinearity. *Wiley Interdisciplinary Reviews: Computational Statistics* 2010;2:370–4. doi:10.1002/wics.84.
- [174] Stine RA. Graphical Interpretation of Variance Inflation Factors. *The American Statistician* 1995;49:53–6. doi:10.1080/00031305.1995.10476113.
- [175] Vu DH, Muttaqi KM, Agalgaonkar AP. A variance inflation factor and backward elimination based robust regression model for forecasting monthly electricity demand using climatic variables. *Applied Energy* 2015;140:385–94. doi:10.1016/j.apenergy.2014.12.011.
- [176] Giacomino A, Abollino O, Malandrino M, Mentasti E. The role of chemometrics in single and sequential extraction assays: A Review. Part II. Cluster analysis, multiple linear regression, mixture resolution, experimental design and other techniques. *Analytica Chimica Acta* 2011;688:122–39. doi:10.1016/j.aca.2010.12.028.
- [177] Vlachogianni A, Kassomenos P, Karppinen A, Karakitsios S, Kukkonen J. Evaluation of a multiple regression model for the forecasting of the concentrations of NO_x and PM₁₀ in Athens and Helsinki. *The Science of the Total Environment* 2011;409:1559–71. doi:10.1016/j.scitotenv.2010.12.040.
- [178] Abdul-Wahab SA, Bakheit CS, Al-Alawi SM. Principal component and multiple regression analysis in modelling of ground-level ozone and factors affecting its concentrations. *Environmental Modelling and Software* 2005;20:1263–71. doi:10.1016/j.envsoft.2004.09.001.
- [179] Pires JCM, Martins FG, Sousa SI V, Alvim-Ferraz MCM, Pereira MC. Selection and validation of parameters in multiple linear and principal component regressions. *Environmental Modelling and Software* 2008;23:50–5. doi:10.1016/j.envsoft.2007.04.012.

- [180] Hanna S, Chang J. Acceptance criteria for urban dispersion model evaluation. *Meteorology and Atmospheric Physics* 2012;116:133–46. doi:10.1007/s00703-011-0177-1.
- [181] Chang JC, Hanna SR. Air quality model performance evaluation. *Meteorology and Atmospheric Physics* 2004;87:167–96. doi:10.1007/s00703-003-0070-7.
- [182] Karacan CÖ. Modeling and prediction of ventilation methane emissions of U.S. longwall mines using supervised artificial neural networks. *International Journal of Coal Geology* 2008;73:371–87. doi:10.1016/j.coal.2007.09.003.
- [183] Najafi G, Ghobadian B, Tavakoli T, Buttsworth DR, Yusaf TF, Faizollahnejad M. Performance and exhaust emissions of a gasoline engine with ethanol blended gasoline fuels using artificial neural network. *Applied Energy* 2009;86:630–9. doi:10.1016/j.apenergy.2008.09.017.
- [184] Sayin C, Ertunc HM, Hosoz M, Kilicaslan I, Canakci M. Performance and exhaust emissions of a gasoline engine using artificial neural network. *Applied Thermal Engineering* 2007;27:46–54. doi:10.1016/j.applthermaleng.2006.05.016.
- [185] Nunnari G, Dorling S, Schlink U, Cawley G, Foxall R, Chatterton T. Modelling SO₂ concentration at a point with statistical approaches. *Environmental Modelling and Software* 2004;19:887–905. doi:10.1016/j.envsoft.2003.10.003.
- [186] Willmott CJ, Robeson SM, Matsuura K. A refined index of model performance. *International Journal of Climatology* 2012;32:2088–94. doi:10.1002/joc.2419.
- [187] Willmott CJ. On the validation of models. *Physical Geography* 2013;2:184–94.
- [188] Hanna SR, Hansen OR, Dharmavaram S. FLACS CFD air quality model performance evaluation with Kit Fox, MUST, Prairie Grass, and EMU observations. *Atmospheric Environment* 2004;38:4675–87. doi:10.1016/j.atmosenv.2004.05.041.
- [189] Willmott CJ, Robeson SM, Matsuura K. A refined index of model performance. *International Journal of Climatology* 2012;32:2088–94. doi:10.1002/joc.2419.
- [190] Moriasi DN, Arnold JG, Van Liew MW, Binger RL, Harmel RD, Veith TL. Model evaluation guidelines for systematic quantification of accuracy in watershed

simulations. Transactions of the ASABE 2007;50:885–900. doi:10.13031/2013.23153.

- [191] Stamenković LJ, Antanasijević DZ, Ristić M, Perić-Grujić AA, Pocajt V V. Modeling of ammonia emission in the USA and EU countries using an artificial neural network approach. Environmental Science and Pollution Research 2015;22:18849–58. doi:10.1007/s11356-015-5075-5.
- [192] Radojević D, Pocajt V, Popović I, Perić-Grujić A, Ristić M. Forecasting of greenhouse gas emissions in serbia using artificial neural networks. Energy Sources, Part A: Recovery, Utilization and Environmental Effects 2013;35:733–40. doi:10.1080/15567036.2010.514597.
- [193] Sundvor I, Balaguer NC, Guerreiro C, Viana M, Querol X RC, Amato F MG. Road traffic's contribution to air quality in European cities ETC/ACM Technical Paper 2012/14 2013. http://acm.eionet.europa.eu/reports/ETCACM_TP_2012_14_traffic_contribution_city_aq (accessed March 7, 2016).
- [194] Eurostat. Final energy consumption by sector 2014. <http://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&language=en&pcode=tsdpc320&plugin=1> (accessed March 7, 2016).
- [195] Eurostat. Statistics database 2014. <http://epp.eurostat.ec.europa.eu/web/main> (accessed March 7, 2016).
- [196] USDA. National Agricultural Statistics Service 2014. <http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1017> (accessed March 7, 2016).
- [197] USEIA. Annual Energy Review - Energy Information Administration 2014. <http://www.eia.gov/totalenergy/data/annual/> (accessed March 7, 2016).
- [198] EPA. National Emissions Inventory (NEI) air pollutant emissions trends data 2014. <http://www3.epa.gov/> (accessed March 7, 2016).
- [199] Stamenković LJ, Antanasijević DZ, Ristić MĐ, Perić-Grujić AA, Pocajt V V. Estimation of NMVOC emissions using artificial neural networks and economical and sustainability indicators as inputs. Environmental Science and Pollution Research 2016. doi:10.1007/s11356-016-6279-z.

- [200] Li X, Maier HR, Zecchin AC. Improved PMI-based input variable selection approach for artificial neural network and other data driven environmental and water resource models. *Environmental Modelling & Software* 2015;65:15–29. doi:10.1016/j.envsoft.2014.11.028.
- [201] (EEA) EEA. Non-methane volatile organic compounds (NMVOC) emissions — European Environment Agency 2014.
- [202] (JRC-IES) I for E and S. Global Emissions EDGAR v4.2 2011. <http://edgar.jrc.ec.europa.eu/overview.php?v=42>. (accessed March 8, 2016).
- [203] American Chemistry Council (ACC). Chemical Shipments 2015. <https://www.americanchemistry.com/> (accessed March 8, 2016).
- [204] OICA. The International Organization of Motor Vehicle Manufacturers. Production Statistics | OICA 2015. <http://www.oica.net/category/production-statistics/> (accessed October 20, 2015).
- [205] National Bureau of Statistics of China. National Bureau of Statistics of China 2015. <http://www.stats.gov.cn/english/> (accessed March 8, 2016).
- [206] The World Bank. World Development Indicators | Data 2015. <http://data.worldbank.org/data-catalog/world-development-indicators> (accessed November 7, 2015).
- [207] Zhao B, Wang SX, Dong XY, Wang JD, Duan L, Fu X, et al. Environmental effects of the recent emission changes in China: implications for particulate matter pollution and soil acidification. *Environmental Research Letters* 2013;8:10. doi:10.1088/1748-9326/8/2/024031.
- [208] Theloke J FR. Emissions of Air Pollutants, Measurements, Calculations and Uncertainties. Springer-Verlag Berlin Heidelberg 2004:63–79. <http://download.springer.com/static/pdf/351/bfm%3A978-3-662-07015-4%2F1.pdf?originUrl=http://link.springer.com/book/bfm:978-3-662-07015-4/1&token2=exp=1457448516~acl=/static/pdf/351/bfm%253A978-3-662-07015-4%252F1.pdf?origi> (accessed March 8, 2016).
- [209] Stamenković LJ, Antanasijević DZ, Ristić MĐ, Perić-Grujić AA, Pocajt V V.

Modeling of methane emissions using artificial neural network approach. *Journal of the Serbian Chemical Society* 2015;80:421–33. doi:10.2298/JSC020414110S.

- [210] Arhami M, Kamali N, Rajabi MM. Predicting hourly air pollutant levels using artificial neural networks coupled with uncertainty analysis by Monte Carlo simulations. *Environmental Science and Pollution Research International* 2013;20:4777–89. doi:10.1007/s11356-012-1451-6.
- [211] Couth R, Trois C, Vaughan-Jones S. Modelling of greenhouse gas emissions from municipal solid waste disposal in Africa. *International Journal of Greenhouse Gas Control* 2011;5:1443–53. doi:10.1016/j.ijggc.2011.08.001.
- [212] Schimel J. Rice, microbes and methane. *Nature* 2000;403:375, 377. doi:10.1038/35000325.
- [213] Cheng K, Ogle SM, Parton WJ, Pan G. Predicting methanogenesis from rice paddies using the DAYCENT ecosystem model. *Ecological Modelling* 2013;261-262:19–31. doi:10.1016/j.ecolmodel.2013.04.003.
- [214] Aljaloud AA, Yan T, Abdukader AM. Development of a national methane emission inventory for domestic livestock in Saudi Arabia. *Animal Feed Science and Technology* 2011;166-167:619–27. doi:10.1016/j.anifeedsci.2011.04.044.
- [215] EDGAR. EDGAR - Emission Database for Global Atmospheric Research. Global Emissions EDGAR v42 2015. <http://edgar.jrc.ec.europa.eu/overview.php?v=42> (accessed October 19, 2015).
- [216] Rypdal K, Winiwarter W. Uncertainties in greenhouse gas emission inventories - Evaluation, comparability and implications. *Environmental Science and Policy* 2001;4:107–16. doi:10.1016/S1462-9011(00)00113-1.
- [217] Stamenković LJ, Antanasijević DZ, Ristić MĐ, Perić-Grujić AA, Pocajt V V. Prediction of nitrogen oxides emissions at the national level based on optimized artificial neural network model. *Air Quality, Atmosphere & Health* 2016. doi:10.1007/s11869-016-0403-6.
- [218] NationMaster. NationMaster: Stats by Category 2015. <http://www.nationmaster.com/statistics> (accessed November 7, 2015).

- [219] Cvetinović D, Stefanović P, Marković Z, Bakić V, Turanjanin V, Jovanović M, et al. GHG (Greenhouse Gases) emission inventory and mitigation measures for public district heating plants in the Republic of Serbia. *Energy* 2013;57:788–95. doi:10.1016/j.energy.2012.09.063.
- [220] Stamenković L, Antanasijević D, Ristić M, Perić-Grujić A, Pocajt V. Predviđanje emisije gasova staklene bašte u Srbiji za period 1995-2013. godina primenom rekurentnih neuronskih mreža. *Ecologica* 2015;22:488–92.
- [221] Eurostat. Database - Eurostat 2015. <http://ec.europa.eu/eurostat/data/database> (accessed March 10, 2016).
- [222] UNFCCC. The UNFCCC secretariat. Emissions Summary for Serbia 2015. <http://unfccc.int/secretariat/items/1629.php> (accessed March 10, 2016).

BIOGRAFIJA AUTORA

Lidija Stamenković rođena je 17.11.1977. godine u Leskovcu, gde je završila srednju školu sa odličnim uspehom. Prirodno-matematički fakultet u Nišu upisala je školske 1996/97. godine, na odseku Hemija. Diplomski rad pod nazivom „*Hemijsko-termički aktiviran treset kao adsorbens i filtraciono sredstvo u prečišćavanju fekalnih voda*“, odbranila je 2006. godine, sa ocenom 10. Od 2009. do 2011. godine radila je kao saradnik u nastavi, a od novembra 2011. do jula 2015. godine radila je kao asistent na Fakultetu za ekologiju i zaštitu životne sredine, Univerziteta Union-Nikola Tesla u Beogradu.

Doktorske studije na Tehnološko-metalurškom fakultetu Univerziteta u Beogradu upisala je školske 2010/11. godine na studijskom programu Inženjerstvo zaštite životne sredine. U okviru doktorskih studija položila je sve ispite predviđene studijskim programom, sa prosečnom ocenom 9,42. Završni ispit pod nazivom „*Razvoj modela za predviđanje emisije zagađujućih materija u vazduhu primenom neuronskih mreža*“ je odbranila u septembru 2012. godine sa ocenom 10.

Koautor je 7 radova objavljenih u časopisima međunarodnog i nacionalnog značaja i 7 naučnih saopštenja publikovanih u zbornicima radova sa međunarodnih i nacionalnih skupova.

Izjava o autorstvu

Potpisani-a Lidija J. Stamenković

broj indeksa 4010/2010

Izjavljujem

da je doktorska disertacija pod naslovom

Predviđanje emisije gasovitih zagađujućih materija na nacionalnom nivou primenom modela zasnovanih na veštačkim neuronskim mrežama

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

Potpis doktoranda

U Beogradu, 06. 06. 2016.

Stamenković Lidija

Prilog 2.

Izjava o istovetnosti štampane i elektronske verzije doktorskog rada

Ime i prezime autora Lidija J. Stamenković

Broj indeksa 4010/2010

Studijski program Inženjerstvo zaštite životne sredine

Naslov rada Predviđanje emisije gasovitih zagađujućih materija na nacionalnom nivou primenom modela zasnovanih na veštačkim neuronskim mrežama

Mentor Dr Viktor Pocajt, vanredni profesor

Potpisani/a Lidija J. Stamenković

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, 06. 06. 2016.

Stamenković Lidija

Izjava o korišćenju

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

Predviđanje emisije gasovitih zagađujućih materija na nacionalnom nivou primenom modela zasnovanih na veštačkim neuronskim mrežama

koja je moje autorsko delo.

Disertaciju sa svim prilogima 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 poledini lista).

U Beogradu, 06. 06. 2016.

Potpis doktoranda

Stamenković Joliza

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.