

Univerza v Ljubljani
Fakulteta za elektrotehniko

Boštjan Grašič

**Napovedovanje povišanih koncentracij ozona z uporabo umetnih
nevronske mreže, Gaussovih procesov in mehke logike**

Magistrsko delo

Mentor: prof. dr. Marko Munih
Somentor: prof. dr. Drago Matko

V Ljubljani, marec 2005

Zahvala

Iskreno se zahvaljujem prof. dr. Marku Munihu za vso pomoč v času študija in pri izdelavi magistrske naloge ter prof. dr. Dragu Matku za vso pomoč pri izdelavi seminarske in magistrske naloge.

Še posebej se zahvaljujem tudi sodelavcema dr. Primožu Mlakarju in dr. Mariji Božnar. Skupaj smo predebatirali marsikateri problem s področja okoljskega modeliranja, kar je bilo podlaga za raziskave zajete v tej magistrski nalogi.

Prav tako se zahvaljujem tudi vsem sodelavcem podjetja AMES d.o.o, ki so mi omogočili delo na področju okoljskega monitoringa in mi nudili ustrezne materialne pogoje za delo.

Povzetek

Namen magistrskega dela je izgradnja modela za vnaprejšnje opozarjanje ljudi na pričakovane presežene opozorilne in alarmne vrednosti koncentracije ozona. Splošno problematiko na tem področju prikazuje uvodno poglavje. V nasprotju z višjimi plastmi atmosfere, kjer je ozona premalo, je v spodnji plasti ozračja ozona preveč, zaradi česar je škodljiv. Ozon spada v skupino sekundarnih onesnaževalcev zraka in se v urbanem smogu pojavlja kot ena od glavnih sestavin.

S problemom škodljivega ozona se srečujemo tudi v Sloveniji. Običajno je ukrepanje ob nastopu prekoračitve opozorilne vrednosti prepozno, saj se maksimalne koncentracije pojavljajo v relativno kratkem časovnem intervalu v času okrog poldneva, ko je sončno sevanje najmočnejše. Obveščanje ljudi o trenutno preseženih koncentracijah ozona ne pomaga bistveno, saj je bila škoda, ki je bila storjena za njihovo zdravje, že v glavnem storjena in je nepopravljiva.

Za modeliranje je bila izbrana lokacija Nova Gorica, ki predstavlja tipično urbano okolje z visokimi koncentracijami ozona v poletnih mesecih. Za izgradnjo modela so bili uporabljeni in preizkušeni trije pristopi modeliranja z umetno nevronske mreže, Gaussovimi procesi in mehko logiko.

V magistrskem delu so v drugem poglavju najprej predstavljene kemične lastnosti in fotokemični proces nastajanja in razgradnje ozona v atmosferi. Razumevanje kompleksnosti pri nastajanju in razgradnji ozona tudi razloži, zakaj sem se odločil za uporabo modeliranja s pomočjo modelov kot so nevronske mreže, Gaussovi procesi in mehka logika, namesto kakšnega drugega fizikalnega ali kemijskega modela. Sledi splošen pregled glavnih skupin povzročiteljev ozona kot so industrija, električne naprave, izpuhi motornih vozil, hlapci goriva in kemična topila, ki so glavni viri dušikovih oksidov (NO_x) in hlapnih organskih spojin (VOC). Na kratko je opisan tudi vpliv povišanih koncentracij ozona na ljudi, kot ga opisuje svetovna zdravstvena organizacija.

Predstavitvi problematike ozona sledi pregled dela v tretjem poglavju, ki je bilo narejeno na področju modeliranja onesnaženja ozračja. Pregled se osredotoča predvsem na modeliranje kompleksnih pojavov v ozračju, ki jih ni mogoče opisati s preprostimi fizikalnimi in kemijskimi enačbami. Umetne nevronske mreže in mehka logika so bile že uporabljene v modelih za kratkoročno napovedovanje visokih koncentracij škodljivih snovi v ozračju. V literaturi pa ni bilo zaslediti uporabe Gaussovih procesov za podobno nalogo modeliranja.

Četrto poglavje zajema metode, ki sem jih uporabil za sintezo modela z umetnimi nevronskimi mrežami, Gaussovimi procesi in mehko logiko. Posebna pozornost je namenjena Gaussovemu procesom, ki predstavljajo eno izmed novejših metod za modeliranje. Poleg tega pa so v tem poglavju predstavljene tudi metode za vrednotenje modelov.

Pred začetkom izgradnje modela je v petem poglavju najprej podana analiza izmerjenih vrednosti okoljskih parametrov na avtomatski okoljski merilni postaji v

Novi Gorici, med drugim tudi glede na priporočila iz literature. Opredelil sem tudi geografski položaj Nove Gorice in glavne vire onesnaževanja ozračja. Analizo sem zaključil s predstavitvijo vhodnih spremenljivk modela.

Za napovedovanje povišanih koncentracij ozona sem v naslednjem poglavju glede na izbrane vhodne in izhodne spremenljivke zgradil tri različne modele z uporabo treh različnih metod modeliranja opisanih predhodno: z umetno nevronske mreže, Gaussovimi procesi in mehko logiko. Za izgradnjo vseh treh modelov je bila uporabljena ista učna množica.

Vse tri modele sem po sintezi tudi ovrednotil z uporabo različnih cenilk, kar je podano v sedmem poglavju. Napovedi povišanih koncentracij sem ocenil z uporabo štirih različnih cenilk: korelacijskega koeficienta (r), kvadratnega korena povprečne kvadratne napake (RMSE), cenilko p^6 in parametri sposobnosti (SP, SR in SI).

Rezultati primerjave vrednotenja modelov kažejo, da je mogoče z različnimi pristopi modeliranja zgraditi povsem enakovredne modele. Prvo prednost modela z Gaussovimi procesi v primerjavi modeloma z umetno nevronske mreže in mehko logiko vidim v tem, da je predikcija izhoda modela na osnovi Gaussovih procesov normalna verjetnostna porazdelitev s srednjo vrednostjo in varianco. Drugo prednost pa v optimizaciji precej manjšega števila hiperparametrov, kot v primeru umetne nevronske mreže in mehke logike. Slabost, ki se je izkazala pri optimizaciji modela na osnovi Gaussovih procesov, pa je nastopala v primeru neustreznih začetnih vrednosti hiperparametrov.

Izkušnje pridobljene pri izdelavi naloge kažejo, da je dobra identifikacija modela odvisna predvsem od pravilne izbire vhodnih spremenljivk. Glede na pregled priporočil iz literature se je izkazalo, da imajo nekatere vhodne spremenljivke bistveno manjši vpliv na izhod, kot je bilo omenjeno, ali pa celo onemogočajo dobro identifikacijo.

Nova Gorica se bo kot urbano okolje v prihodnosti verjetno še širila, kar morda pomeni povečevanje števila prekoračenih vrednosti koncentracij ozona. Ker za vpliv na predvidene povišane koncentracije ozona v urbanem okolju zaenkrat še nimamo direktnih vzvodov, se kaže uporaba modela za napovedovanje povišanih koncentracij ozona edino smiselna in upravičena za varovanje zdravja ljudi.

Abstract

Aim of the thesis is to build a predictive model for high ozone concentrations. Predicted high ozone concentrations that exceed warning and alarming limits would be used to inform public in advance. General environmental aspects of ozone are given in introductory chapter. Contrary to highest layers of atmosphere where the concentration is too low the lower layer of atmosphere contains too high levels of ozone concentration, making it harmful for our health. Ozone is declared as secondary air pollutant. It occurs as one of the main components of urban smog.

Slovenia is also confronted with the harmful ozone problem. Usually, it is too late for measures to be taken after ozone concentrations have already exceeded warning (limiting) levels, because high ozone concentrations occur in relative short time intervals around noon time when solar radiation is the highest. Informing the public about exceeded ozone concentrations does not help much, because damage that inflicted people health has already been made and their health cannot be restored.

Nova Gorica town was used as a modelling location because it represents typical urban environment with high ozone concentrations in summer months. In order to build a model three different modelling methods were used and tested: artificial neural networks, Gaussian processes and fuzzy logic.

The thesis in second chapter begins with a description of ozone chemical properties and a photochemical process of ozone composition and decomposition in atmosphere. Understanding this complex process explains my decision to build a model based on modelling techniques such as artificial neural networks, Gaussian processes and fuzzy logic, instead of using some other chemical or physical model. The thesis proceeds with a general overview of main origins of ozone, such as industrial fuel combustion, electrical appliances, motor vehicles and consumer solvents; they represent the main sources of nitrogen oxides (NO_x) and volatile organic compounds (VOC). The influence of high ozone concentrations on human health is also briefly described.

The thesis continues with an overview of air pollution modelling research field. It mainly focuses on modelling complex occurrences in atmosphere that cannot be described by simple physical and chemical equations. Artificial neural networks and fuzzy logic have already been used to model high concentrations of harmful compounds in atmosphere. However, I did not trace any example of using Gaussian processes for this task.

The fourth chapter is covering the methods that were used for model synthesis: artificial neural networks, Gaussian processes and fuzzy logic. The representation in particular focuses attention on Gaussian processes.

Before the model synthesis are in chapter five given all measured environmental data from automatic environmental measuring station located in Nova Gorica. Geographical position of Nova Gorica town and its main sources of ozone were

also defined. The model inputs analysis was conducted.

According to selected input variables and output variables were in chapter five created three models for high ozone concentration prediction based on different modelling approaches previously described: artificial neural networks, Gaussian processes and fuzzy logic. The construction of all three models applied to one identical learning set of data.

In the seventh chapter are models also evaluated with four different evaluation methods: correlation coefficient (r), square root of mean square error (RMSE), p^6 and skill parameters (SP, SR in SI). The comparison of evaluation results shows that is possible to create very similar models by applying different modelling methods.

Output of Gaussian processes is given as Gaussian conditional distribution that consists of mean and standard deviation of prediction. This is one of the advantages of Gaussian processes, namely the output of the other two modelling techniques is only a means of prediction. The second advantage is small number of hyperparameters that are used during optimization process. A disadvantage of Gaussian processes modelling technique showed numeric instability of the method occurring in a case of inappropriate initial hyperparameter values.

Experiences gained in this thesis is confirming that is good model identification basically based on proper selection of input variables. According to the review of recommendations found in literature some of input variables have insignificant influence on output. In the worst case they obstruct good model identification.

In future Nova Gorica town as an urban environment will probably expand furthermore, meaning that the number of high ozone concentrations will also rise. The use of high ozone concentration prediction model seems to me only feasible and justified move for human health protection because, no other alternative seems to be currently available.

Vsebina

1.Uvod.....	1
2.Ozon.....	5
2.1.Kemična struktura ozona.....	5
2.2.Ozon kot oksidant.....	6
2.2.1.Električna redukcija ozona.....	6
2.2.2.Kemična redukcija ozona.....	6
2.3.Reagiranje ozona.....	7
2.4.Fotokemični proces nastajanja in razgradnje ozona v atmosferi.....	7
2.4.1.Nastajanje in razgradnja ozona v naravi.....	8
2.4.2.Nastajanje in razgradnja ozona v urbanem okolju.....	8
2.5.Povzročitelji ozona v urbanem okolju.....	8
2.5.1.NOx.....	9
2.5.2.Hlapne organske spojine (VOC).....	10
2.6.Vpliv ozona na ljudi.....	11
2.7.Meritve ozona v Sloveniji.....	12
3.Modeliranje onesnaženja ozračja.....	13
3.1.Uporaba modelov za modeliranje ozračja.....	16
4.Metode za modeliranje in vrednotenje modelov.....	21
4.1.Umetne nevronske mreže.....	21
4.2.Gaussovi procesi.....	24
4.2.1.Bayesovo modeliranje	24
4.2.2.Modeliranje z Gaussovimi procesi.....	26
4.2.3.Kovariančna funkcija.....	32
4.2.4.Stacionarne kovariančne funkcije.....	32
4.2.5.Nestacionarne kovariančne funkcije.....	35
4.2.5.1.Linearni izraz.....	35
4.2.5.2.Vhodno odvisen model šuma.....	36
4.2.5.3.Prostorsko spremenljiva dolžinska merila.....	36
4.2.6.Določanje hiperparametrov kovariančne funkcije.....	36
4.2.6.1.Metoda največje podobnosti.....	37
4.2.6.2.Metoda z uporabo Monte Carlo pristopa.....	37
4.2.6.3.Primerjava opisanih metod.....	38
4.3.Mehka logika.....	39
4.3.1.Mehka identifikacija sistemov.....	40
4.3.2.Grupiranje mehkih množic.....	42
4.3.3.ANFIS (Adaptive Neuro-Fuzzy Inference System).....	43
4.4.Metode za vrednotenje modelov.....	43
4.4.1.Korelacijski koeficient.....	44
4.4.2.Kvadratni koren povprečne kvadratne napake (RMSE).....	44
4.4.3.Cenilka p6.....	44
4.4.4.Parametri sposobnosti.....	45
5.Analiza izmerjenih podatkov.....	47
5.1.Geografski položaj Nove Gorice.....	47

5.1.1. <i>Vipavska dolina in Goriško polje</i>	47
5.1.2. <i>Nova Gorica</i>	48
5.2. Meritve meteoroloških podatkov in koncentracij snovi v zunanjem zraku v Novi Gorici	51
5.3. Izbira vhodnih spremenljivk modela	56
6. Sinteza modelov za napovedovanje povišanih koncentracij ozona	63
6.1. <i>Sinteza modela z umetno nevronske mreže</i>	63
6.2. <i>Sinteza modela z Gaussovimi procesi</i>	65
6.3. <i>Sinteza modela z mehko logiko</i>	67
7. Vrednotenje modelov	69
7.1. <i>Preizkušanje modela na osnovi umetne nevronske mreže</i>	70
7.2. <i>Preizkušanje modela na osnovi Gaussovih procesov</i>	71
7.3. <i>Preizkušanje modela na osnovi mehke logike</i>	72
7.4. <i>Primerjava rezultatov preizkušanja modelov</i>	73
8. Zaključek	75
9. Viri in literatura	79
Dodatek A: Osnove verjetnostnega modeliranja	85
A.1. <i>Naključna spremenljivka</i>	85
A.2. <i>Analiza porazdelitev naključnih spremenljivk</i>	86
A.3. <i>Normalna (Gaussova) porazdelitev</i>	86
A.4. <i>Naključni vektorji</i>	87
Dodatek B: Preprost primer modeliranja nelinearne funkcije z Gaussovimi procesi	91
B.1. <i>Identifikacija modela</i>	91
B.2. <i>Prikaz vpliva posameznih hiperparametrov modela</i>	98
Dodatek C: Odvisnost negotovosti od kovariančne funkcije	103
C.1. <i>Uporaba stacionarne kovariančne funkcije brez upoštevanja šuma</i>	103
C.2. <i>Uporaba stacionarne kovariančne funkcije z upoštevanjem šuma</i>	104

Seznam slik

Slika 1: Ozon v atmosferi.....	1
Slika 2: Resonačni strukturi ozona.....	5
Slika 3: "Opazovana" struktura ozona.....	5
Slika 4: Redukcija ozona.....	6
Slika 5: Idealizirana reakcija etilena z ozonom.....	7
Slika 6: Viri NO _x	9
Slika 7: Viri VOC.....	9
Slika 8: Izmerjene vrednosti avtomatske merilne postaje v Novi Gorici dne 05.07.2004.....	10
Slika 9: Večnivojska preceptronska mreža.....	22
Slika 10: Nevron.....	22
Slika 11: Log-sigmoidna prenosna funkcija.....	23
Slika 12: Tan-sigmoidna prenosna funkcija.....	23
Slika 13: Linearna prenosna funkcija.....	23
Slika 14: Primer Gaussovega procesa.....	27
Slika 15: Pogojna porazdelitev.....	28
Slika 16: Porazdelitev v novi točki.....	32
Slika 17: Diagram poteka mehkega sklepanja v mehkih identifikacijskih modelih s pomočjo praktičnega primera "Dinner for two" iz programskega orodja Matlab [54].....	42
Slika 18: Izsek iz državne pregledne karte Slovenije prikazuje širšo okolico Vipavske doline in Goriškega polja.....	47
Slika 19: Relief na področju Vipavske doline in Goriškega polja.....	48
Slika 20: Zemljevid Nove Gorice.....	49
Slika 21: Letni potek mesečnih povprečnih temperatur zraka v večjih krajih v Sloveniji za obdobje od leta 1991 do leta 2000 [3].....	50
Slika 22: Letni potek mesečnega povprečnega trajanja sončnega obsevanja v urah v večjih krajih v Sloveniji za obdobje od leta 1991 do leta 2000 [3].....	51
Slika 23: Histogram temperature zraka.....	54
Slika 24: Histogram relativne vlage.....	54
Slika 25: Histogram moči globalnega sončnega sevanja na enoto površine.....	54
Slika 26: Histogram zračnega pritiska.....	54
Slika 27: Roža vetrov.....	54
Slika 28: Statistika rože vetrov.....	54
Slika 29: Histogram koncentracij žveplovega dioksida.....	55
Slika 30: Histogram koncentracij ogljikovega monoksida.....	55
Slika 31: Histogram koncentracij dušikovega monoksida.....	55
Slika 32: Histogram koncentracij dušikovega dioksida.....	55
Slika 33: Histogram koncentracij dušikovih oksidov.....	55
Slika 34: Histogram koncentracij ozona.....	55
Slika 35: Vhod 1, Temperatura zraka.....	58
Slika 36: Vhod 2, Moč globalnega sončnega sevanja na enoto površine.....	58
Slika 37: Vhod 3, Koncentracija dušikovega monoksida (NO).....	58
Slika 38: Vhod 4, Koncentracija dušikovega dioksida (NO ₂).....	59
Slika 39: Vhod 5, Koncentracija ozona (O ₃).....	59
Slika 40: Vhod 6, Napovedana maksimalna temperatura zraka.....	59
Slika 41: Vhod 7, Napovedan veter v smeri sever-jug.....	60
Slika 42: Vhod 8, Napoved vetra v smeri vzhod-zahod.....	60

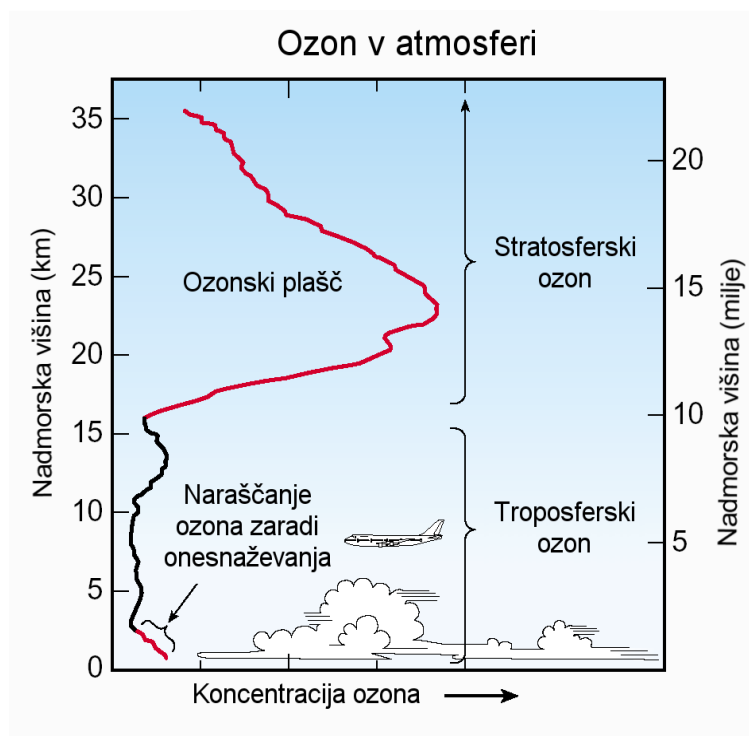
Slika 43: Izhod, Maksimalne dnevne koncentracije ozona.....	60
Slika 44: Vrednosti korelacijskih koeficientov med posameznimi vhodnimi spremenljivkami v model in izhodno spremenljivko modela.....	61
Slika 45: Rezultati preizkusa modela z umetno nevronske mreže s podatki iz učne množice.....	64
Slika 46: Histogram napak preizkusa modela z umetno nevronske mreže s podatki iz učne množice.....	64
Slika 47: Rezultati preizkusa modela z Gaussovimi procesi s podatki iz učne množice.....	66
Slika 48: Histogram napak preizkusa modela z Gaussovimi procesi s podatki iz učne množice.....	66
Slika 49: Rezultati preizkusa modela z mehko logiko s podatki iz učne množice.....	68
Slika 50: Histogram napak preizkusa modela z mehko logiko s podatki iz učne množice.....	68
Slika 51: Rezultat preizkušanja modela na osnovi umetne nevronske mreže.....	70
Slika 52: Odstopanje med izmerjenimi in napovedanimi vrednostmi.....	70
Slika 53: Histogram napak preizkušanja modela na osnovi umetne nevronske mreže.....	70
Slika 54: Rezultat preizkušanja modela na osnovi Gaussovih procesov.....	71
Slika 55: Odstopanje med izmerjenimi in napovedanimi vrednostmi.....	71
Slika 56: Histogram napak preizkušanja modela na osnovi Gaussovih procesov.....	71
Slika 57: Rezultat preizkušanja modela na osnovi mehke logike.....	72
Slika 58: Odstopanje med izmerjenimi in napovedanimi vrednostmi.....	72
Slika 59: Histogram napak preizkušanja modela na osnovi mehke logike.....	72
Slika 60: Primerjava korelacijskega faktorja.....	73
Slika 61: Primerjava RMSE.....	73
Slika 62: Primerjava cenilke p6.....	73
Slika 63: Primerjava merila pravih napovedi SP pri meji 180 mg/m ³	73
Slika 64: Primerjava merila uresničenih napovedi SR pri meji 180 mg/m ³	73
Slika 65: Primerjava indeksa uspeha SI pri meji 180 mg/m ³	73
Slika 66: Primerjava merila pravih napovedi SP pri meji 140 mg/m ³	73
Slika 67: Primerjava merila uresničenih napovedi SR pri meji 140 mg/m ³	73
Slika 68: Primerjava indeksa uspeha SI pri meji 140 mg/m ³	73
Slika 69: Primer dvodimenzionalne Gaussove porazdelitvene funkcije.....	89
Slika 70: $\sigma_1=1.0$, $\sigma_2=1.0$, $r=0.99$	89
Slika 71: $\sigma_1=1.0$, $\sigma_2=1.0$, $r=0.00$	89
Slika 72: $\sigma_1=1.0$, $\sigma_2=2.0$, $r=0.99$	90
Slika 73: $\sigma_1=2.0$, $\sigma_2=1.0$, $r=0.99$	90
Slika 74: $\sigma_1=1.0$, $\sigma_2=2.0$, $r=0.0$	90
Slika 75: $\sigma_1=2.0$, $\sigma_2=1.0$, $r=0.0$	90
Slika 76: Prvi model temperature rosišča.....	93
Slika 77: Napaka prvega modela temperature rosišča.....	93
Slika 78: Drugi model temperature rosišča.....	94
Slika 79: Napaka drugega modela temperature rosišča.....	94
Slika 80: Tretji model temperature rosišča.....	95
Slika 81: Napaka tretjega modela temperature rosišča.....	95
Slika 82: Četrty model temperature rosišča.....	96
Slika 83: Napaka četrtega modela temperature rosišča.....	96
Slika 84: Peti model temperature rosišča.....	97
Slika 85: Napaka petega modela temperature rosišča.....	97
Slika 86: Vpliv prvega hiperparametra na model.....	100
Slika 87: Vpliv drugega hiperparametra na model.....	101
Slika 88: Vpliv tretjega hiperparametra na model.....	102

Seznam tabel

Tabela 4.1: Tabela deležev (contingency table).....	45
Tabela 6.1: Vrednosti hiperparametrov.....	65

1.Uvod

Vrsto let nas okoljevarstveniki opozarjajo, da v višjih plasteh atmosfere primanjkuje ozona, kar povzroča rast ozonske luknje. V spodnji plasti ozračja pa se vedno bolj srečujemo z obratno težavo. V zraku, ki ga dihamo, je ozona preveč, zaradi česar je škodljiv.



Slika 1: Ozon v atmosferi

Ozon ima pomemben vpliv na zdravje ljudi in okolje. Glede na lokacijo pojavljanja (slika 1) in koncentracijo je lahko koristen ali škodljiv. Zelo škodljiv vpliv imajo visoke koncentracije ozona, ki se pojavljajo v najnižji plasti atmosfere ob površini Zemlje, troposferi. Ozon spada v skupino sekundarnih onesnaževalcev zraka in škoduje živim bitjem predvsem ob vdihavanju. Škodljiv pa je tudi za poljske pridelke, drevesa in drugo vegetacijo. V urbanem smogu, ki predstavlja zmes prahu, dima in izpušnih plinov, se pojavlja kot ena od glavnih sestavin.

Občasne višje koncentracije ozona v troposferi pa imajo lahko tudi zelo koristen vpliv, saj deluje kot dezinfekcijsko sredstvo bolj učinkovito kot klor [30]. Med drugim se uporablja višje koncentracije za sterilizacijo različnih predmetov. Dokazano je bilo, da ozon uniči katerokoli bakterijo, virus, gobo ali plesen pri izpostavljenosti koncentraciji $850 \mu\text{g}/\text{m}^3$ v štirih minutah.

Troposfera sega približno do višine 10.000 metrov, kjer se začne druga plast atmosfere, to je stratosfera. Stratosfera ali plast koristnega ozona sega od višine 10.000 metrov do 50.000 metrov. V njej se nahaja ozonska plast, ki nas ščiti pred sončnimi ultravijoličnimi žarki in je najgostejša na višini od 18 do 23 kilometrov.

S problemom škodljivega ozona se srečujemo tudi v Sloveniji. Glede na publikacije Agencije RS za okolje [2,4] so bile najvišje koncentracije ozona izmerjene v Novi Gorici in na Krvavcu, kjer v poletnih mesecih, ko je na voljo dovolj sončne svetlobe, tudi prekoračijo opozorilne vrednosti ali celo alarmne vrednosti [58].

Običajno je ukrepanje ob nastopu prekoračitve opozorilne vrednosti prepozno, saj se maksimalne koncentracije pojavljajo v relativno kratkem časovnem intervalu v času okrog poldneva, ko je sončno sevanje najmočnejše. Obveščanje ljudi o trenutno preseženih koncentracijah ozona ne pomaga bistveno, saj je bila škoda, ki je bila storjena za njihovo zdravje, že v glavnem storjena in je nepopravljiva.

Izgradnja modela za vnaprejšnje opozarjanje ljudi na pričakovane presežene opozorilne in alarmne vrednosti koncentracije ozona je zato postala glavna rdeča nit magistrske naloge. Za modeliranje je bila izbrana lokacija Nova Gorica, ki predstavlja tipično urbano okolje z visokimi koncentracijami ozona v poletnih mesecih. Lokacija na Krvavcu ni enako zanimiva za vnaprejšnje opozarjanje, saj je gostota prebivalstva relativno redka.

Napovedovanje preseženih opozorilnih in alarmnih vrednosti koncentracij ozona omogoča lokalnim oblastem vnaprejšnje opozarjanje ljudi in izvajanje določenih ukrepov za njihovo zaščito. Ljudje lahko ob opozorilu spremenijo svoje načrte za prihodnji dan tako, da ne izvajajo težjih telesnih aktivnosti v času visokih koncentracij in se tudi po možnosti čim manj gibljejo zunaj. Lahko tudi zjutraj dobro prezračijo stanovanja in čez dan ne puščajo odprtih oken, po možnosti pa se celo odmaknejo iz območja onesnaženja.

Kemične lastnosti in fotokemični proces nastajanja in razgradnje ozona v atmosferi sta podrobneje opisana v drugem poglavju. Razumevanje kompleksnosti pri nastajanju in razgradnji ozona tudi razloži, zakaj sem se odločil za uporabo modeliranja s pomočjo nevronske mreže, mehke logike in Gaussovih procesov namesto kakšnega drugega fizikalnega ali kemijskega modela. V tem poglavju so predstavljeni tudi glavni viri ozona v urbanem okolju, vpliv ozona na ljudi ter na kratko meritve ozona

v Sloveniji.

Umetne nevronske mreže in mehka logika so bile že uporabljene v modelih za kratkoročno napovedovanje visokih koncentracij škodljivih snovi v ozračju. V literaturi pa ni bilo zaslediti uporabe Gaussovih procesov za podobno nalogo modeliranja. Pregled dela, ki je bilo narejeno na tem področju, je podano v tretjem poglavju.

Sledi četrto poglavje, kjer so tudi podrobneje predstavljeni in opisani omenjeni modeli. Pri opisu omenjenih modelov sem posebno pozornost namenil predstavitvi Gaussovih procesov.

V petem poglavju sem izdelal analizo izmerjenih vrednosti okoljskih parametrov na avtomatski okoljski merilni postaji v Novi Gorici, med drugim tudi glede na priporočila iz literature. V poglavju je tudi opredeljen geografski položaj Nove Gorice in glavni viri oneznaževanja ozračja. Poglavje sem zaključil s predstavitvijo vhodnih spremenljivk modela.

Sledi šesto poglavje, kjer je predstavljena sinteza modela napovedovanja povišanih koncentracij iz izmerjenih podatkov s pomočjo treh različnih metod modeliranja: umetne nevronske mreže, mehke logike in Gaussovih procesov.

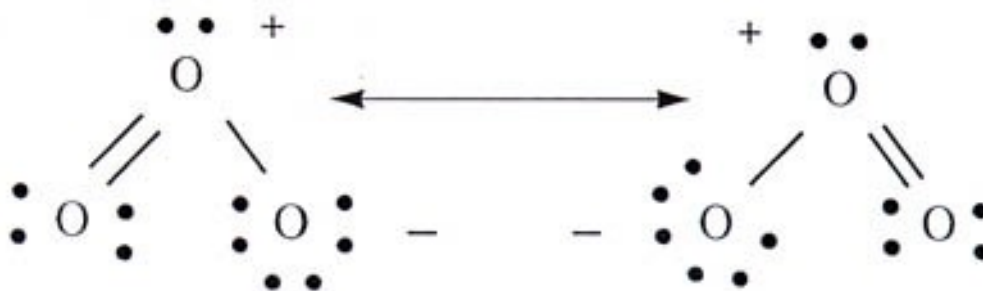
Vrednotenje izgrajenih modelov je podana v sedmem poglavju. Napovedi povišanih koncentracij sem ocenil z uporabo štirih različnih cenilk, ki so tudi podrobneje predstavljene: korelacijski koeficient (r), kvadratni koren povprečne kvadratne napake (RMSE), cenilko p^6 in parametri sposobnosti (SP, SR in SI).

V zaključku je podana splošna primerjava med uporabljenimi tremi načini modeliranja, izpostavljene so glavne ugotovitve ter postavljene smernice za nadaljnje raziskovanje.

2. Ozon

2.1. Kemična struktura ozona

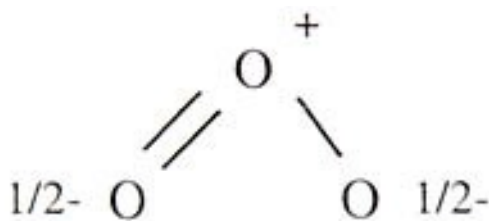
Ozon se pojavlja kot naravna oblika kisika [32], [3]. Ime ozon izhaja iz grške besede “ozein”, kar prevedeno pomeni duh ali vonj. Prvič je bil odkrit v laboratorijskih eksperimentih v srednini 19. stoletja. Njegovo kemično strukturo predstavlja slika 2.



Slika 2: Resonačni strukturi ozona

Iz ilustracije je razvidno, da je vsaka resonančna struktura sestavljena iz ene enojne vezi in ene dvojne vezi. Enojna vez je analogna peroksidnim vezem, ki so relativno šibke in vodijo k tvorbi prostih radikalov. Dvojna vez je analogna molekularnemu kisiku (O_2), ki je močno povezan in relativno nereaktiven.

Pretvarjanje med ilustriranimi resonančnima strukturama ozona je tako hitro, da je struktura “opazovanega” ozona mešanica obeh resonančnih struktur. Posledično temu je moč vezi med osrednjim atomom kisika in preostalima atomoma kisika enaka in znaša eno in pol vezi (1.5), kar je prikazano na sliki 3.



Slika 3: "Opazovana" struktura ozona

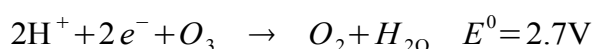
Čisti ozon v obliki plina je modrikaste barve, utekočinjen ($-112\text{ }^{\circ}\text{C}$) je modrikasto-črn, v trdni obliki ($-193\text{ }^{\circ}\text{C}$) pa vijolično-črne barve.

2.2. Ozon kot oksidant

Razumevanje koncepta oksidacijskih stanj nam omogoča razumevanje zakaj je ozon oksidant. Oksidacijsko stanje določa, koliko elektronov je izgubil (v primeru redukcije pa pridobil) atom v molekuli. Oksidacijsko stanje obeh vodikovih atomov v molekuli vode je +1 (ker vsak atom vodika formalno preda en elektron atomu kisika), oksidacijsko stanje atoma kisika v molekuli vode pa je -2 (ker je atom kisika formalno pridobil po en elektron od vsakega atoma vodika). Oksidacijsko stanje kisikovih atomov je ponavadi -2. V molekuli ozona in molekuli kisika pa imajo vsi kisikovi atomi oksidacijsko stanje 0. Iz tega sledi, da sta tako ozon kot kisik oksidanta, ki znata dobro pridobivati manjkajoče elektrone iz različnih virov, s čimer se poveča oksidacijsko stanje vsaj enega od kisikovih atomov v procesu. V praksi je ozon dejansko močnejši oksidant kot kisik. Delno tudi zato, ker lahko ozon direktno reagira z nekaterimi spojinami, medtem, ko kisik za oksidacijo običajno potrebuje še kakšen katalizator za začetek reakcije.

2.2.1. Električna redukcija ozona

Produkt električno pospešene razgradnje ozona je nastanek molekularnega kiska (O_2) in atoma kisika z oksidacijskim stanjem -2 (običajno opisan v obliki molekule vode na sliki 4). Enosmerni proces vsebuje tudi redukcijski električni potencial vrednosti 2.07V (relativno na vodikovo elektrodo). Ta vrednost je večja od redukcijskih potencialov večine drugih snovi. To pomeni, da je zmožnost ozona, da oksidira večino drugih snovi, termodinamično mogoča.



Slika 4: Redukcija ozona

2.2.2. Kemična redukcija ozona

Kemično vzpodbujena razgradnja ozona ni povsem preprosta zaradi nastopanja velikega števila različnih reakcij. Eden od bolj preprostih primerov uporabe ozona kot oksidanta v kemičnem procesu predstavlja reakcija ozona z etilenom na sliki 5. V tem procesu nastane pri reakciji ozona z etilenom molekulska struktura A, ki vsebuje dve povezavi kisik-ogljik. V strukturi A izgine tudi dvojna vez ogljik-ogljik. Ogljikovi atomi postanejo oksidirani, kisikovi atomi pa pri povezavi z ogljikom spremenijo oksidacijsko stanje iz 0 na -1. Reakcija se nadaljuje, njen rezultat pa je molekula kisika O_2 in molekulska struktura B. Kisikov atom in struktura B sta sedaj v oksidacijskem stanju -2. Substrat (etilen)

$$\text{H}_2\text{C}=\text{CH}_2 + \text{O}_3 \longrightarrow \begin{array}{c} \text{O} \\ \diagup \quad \diagdown \\ \text{O} \quad \text{O} \\ | \quad | \\ \text{H}_2\text{C} - \text{CH}_2 \\ \text{A} \end{array} \longrightarrow \begin{array}{c} \text{O} \\ || \\ \text{H} - \text{C} - \text{H} \\ \text{B} \end{array} + \text{O}_2$$

2.3.Reaqiranje ozona

Glede na predstavljeno je včasih boljše gledati na ozon kot na zelo reaktivno snov sposobno reagirati z velikim številom drugih snovi, kot pa na snov z zelo visokim redukcijskim potencialom. Razlog je v tem, da so praktične akcije ozona pogosto bolj odvisne od tega kako ozon reagira z ostalimi onesnaževalnimi snovmi, kot pa od njegove zmožnosti, da preprosto pridobiva elektrone (razen v primeru preproste oksidacije ionov). Tako razmišljanje poudarja, da obstaja neka logična reakcijska pot, da lahko ozon reagira z drugimi snovmi. To pomeni, da čeprav ima termodinamična lastnost ozona, da hitro oksidira, prednost, lahko v veliko primerih kinetični faktorji določajo ali bo ozon lahko oksidiral določeno onesnaževalno snov v določenem času.

Nastajanje ozona v ozračju je zelo zapleten proces katerega osnova je fotokemična reakcija, ki potrebuje svetlobo [41]. Brez svetlobe poteka proces nastajanja ozona zelo počasi.

2.4.1. Nastajanje in razgradnja ozona v naravi

V naravi nastajo ozon kot posledica obsevanja segretega kisika O_2 z ultravijolično svetlobo. Reakcijo prikazuje enačba



Ob ohlajanju ozračja in prenehanju obsevanja z ultravijolično svetlobo v nočnih urah se začne razgradnja ozona, kot ga opisuje enačba



2.4.2. Nastajanje in razgradnja ozona v urbanem okolju

Fotokemični oksidanti spadajo v skupino sekundarnih osnaževalcev ozračja. Nastajajo kot posledica sončnega obsevanja atmosfere onesnažene s hlapnimi organskimi spojinami (VOC – Volatile Organic Compounds) in dušikovimi oksidi (NO_x). S kemičnega vidika je sinteza fotokemičnega smoga zelo kompleksna in je še v fazi raziskovanja. Splošno znano je, da ogljikovodiki reagirajo z atomi kisika, z ozonom in z oksidacijskimi radikali zaradi povečanja prostega ravnotežja radikalov. Ti prosti radikali reagirajo z ostalimi ogljikovodiki in drugimi oksidnimi vrstami ter dušikovimi oksidi, kar povzroča smog. To je skupek zmesi kot so dušikovi oksidi, ogljikov monoksid, ogljikovodiki, aldehidi, ozon in različne organske komponente. Dušikovi oksidi se štejejo kot ena poglavitnih komponent, ki vplivajo na nastanek fotokemičnega smoga. Pri fotolizni reakciji z valovno dolžno svetlobe, ki je manjša kot 280 nm, se sproščata kisikov radikal in dušikov oksid. Prosti radikali omogočajo nastajanje sekundarnih oksidantov v atmosferi. V takem smislu hidrooksidni in hidroperoksidni radikali povzročajo oksidacijo ogljikovodikov in s tem pomembno verigo reakcij v atmosferi. Čeprav v fotokemični oksidaciji nastopa več različnih komponent, ki povzročajo reakcijo, v splošnem ni mogoče glavnih reakcijskih procesov opisati s preprostim modelom. Iz tega razloga tudi ni mogoče visokih koncentracij ozona v urbanih predelih opisati z reakcijskim procesom, ki bi upošteval samo reakcijo med NO in NO_2 .

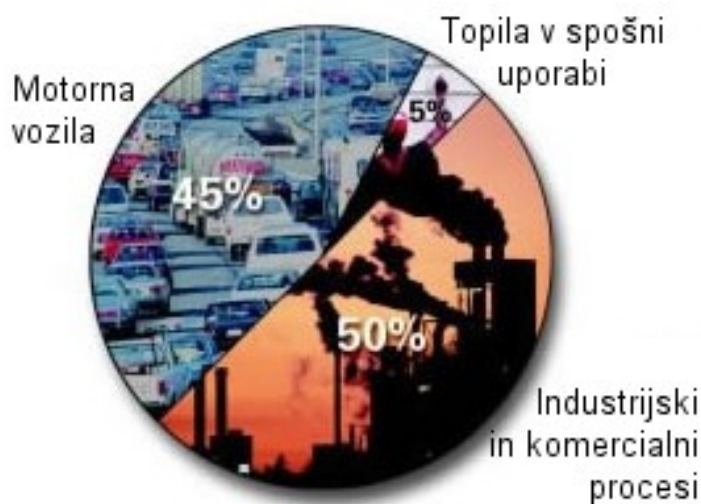
2.5. Povzročitelji ozona v urbanem okolju

Glavne skupine povzročiteljev predstavljajo industrija, električne naprave, izpuhi motornih vozil, hlapi goriva in kemična topila, ki so glavni viri dušikovih oksidov (NO_x) in hlapnih organskih spojin (VOC) [20]. Približen odstotek posameznega onesnaževalca prikazujeta slika 6 in slika 7. Koncentracije ozona se povečajo predvsem v poletnih

mesecih, ko je veliko sončne svetlobe in so temperature zraka visoke.



Slika 6: Viri NO_x



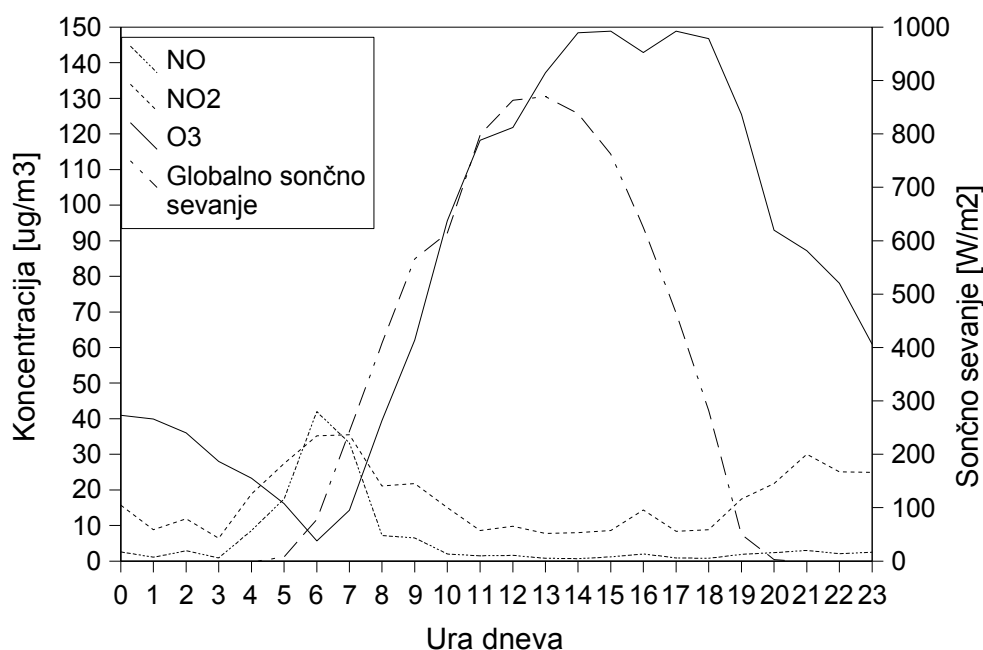
Slika 7: Viri VOC

2.5.1. NO_x

Izmed vseh znanih dušikovih spojin (N_2O , NO , NO_2 , NO_3 , N_2O_2 , N_2O_4 , N_2O_5) se pojavljata kot pomembna onesnaževalca zraka samo dušikov monoksid (NO) in dušikov dioksid (NO_2). Izmerjena koncentracija obeh plinov se pogosto tudi sešteje in označi z NO_x (dušikovi oksidi). Dušikov monoksid je stabilna naravna dušikova snov, ki je zelo pomembna v kemiji stratosfere. Preostali dušikovi oksidi pa igrajo predvsem pomembno vlogo posrednika v mehanizmih fotokemičnega smoga [34].

Dušikov monoksid (NO) je brezbarvna zmes brez vonja, ni gorljiv in je slabo topen v vodi. Spada v skupino zelo toksičnih plinov. V naravi nastaja NO kot produkt bakterijskega delovanja in procesov izgorevanja, ki se pojavljajo naravno.

Dušikov dioksid (NO₂) je plin rdečkasto-oranžno-rjave barve z ostrim dražljivim vonjem. Spada v skupino zelo toksičnih plinov in je zelo jedek in koroziven. Molekule plina absorbirajo večino svetlobe v vidnem spektru, zato se lahko v atmosferi pojavljajo rumenkaste ali oranžne meglice dušikovega dioksida. V zraku nastaja predvsem kot posledica kemične reakcije pretvorbe dušikovega monoksida v dušikov dioksid in ozon. Pretvorba poteka počasneje v čistem zraku in zelo hitro v onesnaženem zraku.



Slika 8: Izmerjene vrednosti avtomatske merilne postaje v Novi Gorici dne 05.07.2004

NO in NO₂ koncentracije v urbanih okoljih se pojavljajo v ponavljajočem vzorcu. V jutranjih urah se najprej pojavijo visoke koncentracije dušikovega monoksida NO kot posledica zgodnjega jutranjega prometa. Običajen primer opisanega procesa v poletnih dneh prikazuje slika 8. Na ilustraciji so prikazane izmerjene vrednosti avtomatske merilne postaje v Novi Gorici dne 05.07.2004.

2.5.2.Hlapne organske spojine (VOC)

Hlapne organske spojine (VOC) [34,62] so kemične zmesi, ki imajo visok hlapni pritisk in nizko vodno topljivost. V to skupino spada širok

spekter molekul, ki temeljijo na ogljiku, kot so na primer aldehid, ciklične zmesi in dolgo-verižni ogljik. Splošno znano je, da so drevesa naravni učinkoviti viri hlapnih organskih spojin. V vsakdanjem življenju se hlapne organske spojine pojavljajo tudi v obliki različnih topil za barve, čistilnih raztopil in kot primesi v gorivu. Pri onesnaževanju zraka nastajajo kot posledica nepopolnega izgorevanja v motornih vozilih ter kot posledica izhlapevanja pri transportu in pretakanju goriva.

V primerih onesnaževanja zraka se hlapne organske spojine mnogokrat razdelijo v dve podskupini, metan (CH_4) in nemetanske hlapne organske spojine (NMVOC). Metan spada v skupino izredno učinkovitih plinov, ki povzročajo globalno pregrevanje ozračja. Med nemetanskimi hlapnimi spojinami pa sta najbolj znana benzen, ki lahko povzroča levkemijo pri predolgem izpostavljanju, in 1,3-butadien, ki je pogosto uporabljen v industrijskih procesih.

Podobno kot dušikovi oksidi tudi hlapne organske spojine kažejo podoben vsakodnevni vzorec v urbanem okolju [34]. V zgodnjih jutranjih urah ob povečanem prometu koncentracije narastejo. Ob pojavu sončne svetlobe pa se začne koncentracija zmanjševati, ker se porabljajo pri nastajanju fotokemičnih oksidantov. Ob poznem popoldnevu se pojavi drugi manjši maksimum, v noči pa so koncentracije relativno nizke.

2.6. Vpliv ozona na ljudi

Količina ozona, ki se nahaja v troposferi, najnižji plasti ozračja, je bistveno manjša od količine stratosferskega ozona, ki nas ščiti pred ultravijoličnimi žarki [19,63]. Na povišanje koncentracije ozona so najbolj občutljivi ljudje z boleznimi dihal ter boleznimi srca in ožilja. Ozon prodre globoko v pljuča, kjer draži sluznico in pljučno tkivo, tako da ovira dihanje. Pri višjih koncentracijah pride do bolečin pri globljem dihanju in do draženja na kašelj. Prav tako draži sluznico v grlu in povzroča pekoč občutek v očeh. Svetovna zdravstvena organizacija priporoča, naj ljudje ne bodo izpostavljeni osemurnim koncentracijam, ki presegajo $120 \mu\text{g}/\text{m}^3$. Za kratkotrajno enourno koncentracijo ozona ugotavljajo, da je izpostavljenost na več kot $180 \mu\text{g}/\text{m}^3$ že neprijetna za občutljive skupine prebivalcev, med katere spadajo predvsem bolniki in starejši ljudje. Pri $240 \mu\text{g}/\text{m}^3$ pa je koncentracija že tako visoka, da lahko nastanejo škodljive posledice. To je tudi predpisana alarmna vrednost pri kateri je potrebno sprejeti nujne ukrepe.

2.7.Meritve ozona v Sloveniji

V Sloveniji se količina ozona v zraku meri že od leta 1992, število merilnih mest pa se je povečalo v letu 1997 [5]. S pomočja sredstev iz projekta Phare se je mreža avtomatskih merilnih okoljskih postaj leta 2001 povečala za osem merilnih mest. Trenutno se v okviru merilne mreže Agencije Republike Slovenije za okolje (ARSO) meri ozon na lokacijah: Ljubljana, Maribor, Nova Gorica, Murska Sobota, Trbovlje, Hrastnik, Zagorje, Krvavec, Iskrba (v bližini Kočevja) in Celje. V začetku leta 2005 pa bo zgrajena tudi postaja v Kopru. ARSO pa sprejema podatke o koncentracijah ozona tudi iz merilne mreže TE Šoštanj, TE Trbovlje in TE Brestanica ter iz avtomatskih mobilnih postaj EIMV (Elektro inštitut Milan Vidmar). Pri meritvah ozona izstopata predvsem postaji Nova Gorica in Krvavec, kjer so bile izmerjene večje vrednosti koncentracije ozona kot drugod [2]. Iz tega razloga so bili za izgradnjo modela za napovedovanje povišane vrednosti ozona izbrani meteorološki in imisijski podatki iz avtomatske merilne postaje v Novi Gorici.

3. Modeliranje onesnaženja ozračja

Glavni namen pregleda je ugotoviti kaj nam moderna znanost do danes ponuja na področju modeliranja onesnaženja ozračja. Pregled se osredotoča predvsem na modeliranje kompleksnih pojavov v ozračju, ki jih ni mogoče opisati s preprostimi fizikalnimi in kemijskimi enačbami. Ob pregledu sem spoznal kakšne modele so uporabljali raziskovalci, s kakšnimi problemi so se srečevali, kakšna so njihova priporočila in nasveti za obvladovanje problemov ter kakšne izboljšave bi bilo še možno uvesti. Uporaba obstoječega znanja in izkušenj mi je omogočila izgradnjo izboljšanega modela za srednjeročno napovedovanje povišanih koncentracij ozona.

Za modeliranje procesov, ki se odvijajo v okolju, je na voljo mnogo različnih pristopov, ki lahko nastopajo tudi v kombinaciji. Izbira najbolj primerne pristopa običajno temelji na kompleksnosti problema, ki ga obravnavamo, in na stopnji razumevanja problema. Uporaba numeričnega modela je najbrž najbolj zaželena rešitev ob predpostavki, da je na voljo zadostno velika in kvalitetna množica podatkov, ustrezna računalniška zmogljivost in dobro teoretično poznavanje problema. V splošnem pa velja, da se s povečevanjem kompleksnosti problema zmanjšuje teoretično razumevanje zaradi slabo definiranih interakcij med sistemi, kar pomeni, da je potrebno do problema pristopiti s statističnimi metodami. Tudi modeliranje ozona v ozračju spada zaradi svoje kompleksnosti v slednjo skupino. Uporaba modernejših pristopov z umetnimi nevronske mrežami, mehko logiko ali Gaussovimi procesi, se zdi boljše izbira in predstavlja učinkovito alternativo tradicionalnim statističnim metodam.

Razširjanje snovi oziroma primesi v ozračju označujemo s skupno besedo disperzija, ki se uporablja kot skupna oznaka za dva mehanizma [40,15]. Prvi mehanizem je transport snovi v ozračju, ki se izvaja s premikom zračnih mas (z vetrom). Drugi mehanizem pa predstavlja difuzija, ki povzroča, da se primesi v ozračju razširjajo tudi prečno na smer vetra in se tako razredčujejo. Difuzija je lahko na ravni molekul, ko primesi prehajajo iz mesta z višjo koncentracijo na mesto z nižjo koncentracijo. To dogajanje je v ozračju praktično zanemarljivo v primerjavi s poglavitno drugo obliko – prisilno difuzijo. Poglavitni mehanizem prisilne difuzije so vrtinci oziroma turbulence v ozračju, ki poskrbijo, da se dim iz vira močno razredči. Kako se bodo škodljivi plini razredčili in kam se bodo razširili določajo emisijski parametri in meteorološke razmere. Pri enakih pogojih emisije je torej predvsem od vremenskih razmer odvisno, kako visoke bodo koncentracije. Razgiban relief močno vpliva predvsem na temperaturno in vetrovno polje, ter s tem posredno tudi na koncentracije škodljivih plinov.

Modeliranje širjenja snovi v ozračju v grobem lahko razdelimo na dve skupini [40,15]. V prvo skupino sodijo modeli, ki skušajo ponazoriti (rekonstruirati) fizikalno dogajanje. S pomočjo podatkov o emisiji in o meteorološkem dogajanju želijo v obravnavanem območju rekonstruirati koncentracije primesi. V to skupino sodijo Gaussovi modeli [31,27,44, 46], različni “puff” modeli [25] in zapleteni numerični modeli delcev [55]. Omenjeni fizikalni modeli temeljijo na matematičnem opisu procesov v atmosferi pri katerem se rezultati računajo kot posledica vzrokov [15]. Takšni modeli skušajo razrešiti pripadajoče kemične in fizikalne enačbe, ki določajo koncentracije onesnaževalnih snovi, zaradi česar pa potrebujejo tudi natančne podatke o emisiji in meteorologiji na opazovanem geografskem področju. V splošnem pa obstajajo resne omejitve v prostorski in časovni natančnosti podatkov. Dodatno pa nekaterih podatkov in rezultatov ni možno pridobiti od okoljskih agencij ali lokalne industrije. To pa pomeni resen problem pri uporabi fizikalnih modelov, saj so nekateri vhodi v model neznani. Iz tega razloga so v praksi uporabljajo statistični modeli.

V drugo skupino pa sodijo modeli, ki za določena značilna mesta v opazovanem območju skušajo napovedati koncentracije za določeno časovno obdobje vnaprej. Pri tem pa ni nujno da je v model eksplicitno vpleten zapleten mehanizem disperzije. V to skupino sodijo predvsem različni statistični modeli [11,13,22], modeli na osnovi umetnih nevronske mreže [17,64,16,23,24,14,1], mehke logike [21] in Gaussovi

procesu.

Glede na časovni interval napovedi se lahko napovedovalni modeli delijo v štiri skupine [7]:

- dolgoročni (dva do tri dni vnaprej),
- srednjeročni (do 24 ur vnaprej),
- kratkoročni (od 6 do 12 ur vnaprej),
- super kratkoročni (od 1 do 2 uri vnaprej).

Glede na cilj, ki sem si ga zadal v tej magistrski nalogi, sem v nadaljevanju posvetil večjo pozornost drugi skupini modelov: umetnim nevronskim mrežam, mehki logiki in Gaussovemu procesu. Omenjeni modeli omogočajo razvoj relativno dobrega modela tudi v primeru izredno kompleksnega procesa, kot je nastajanje ozona [6]. Mešanice NO_x in hlapnih organskih snovi (VOC) v prisotnosti sončne svetlobe igrajo najpomembnejšo vlogo v procesu nastajanja ozona. To je bilo prvič dokazano leta 1940 v dolini Los Angeles [28]. Od takrat naprej je bilo veliko naporov vloženih v razumevanje troposferske fotokemije. Ta prizadevanja dodatno otežuje ogromna množica hlapnih organskih spojin v onesnaženem okolju in izredno veliko načinov, kako lahko posamezne spojine reagirajo. Ocenjeno je bilo, da bi eksplicitna obravnava nastajanja ozona vsebovala več kot 20 000 reakcij z upoštevanjem nekaj tisoč reaktantov in produktov [18]. Za omejitev števila reaktantov in produktov je bilo razvitih kar nekaj kemičnih modelov, ki v svojih mehanizmih uporabljajo povprečno od 30 do 50 spojin ter med 80 in 200 reakcij [18].

3.1. Uporaba modelov za modeliranje ozračja

M. Z. Božnar in P. Mlakar v svojem članku [14] opisujeta dve vrsti nevronske mreže, ki sta uporabni na področju modeliranja onesnaženja ozračja: večnivojska perceptronska nevronska mreža in Kohonenova nevronska mreža. Obe mreži sta na kratko opisani, poleg pa so podani tudi praktični primeri uporabe obeh.

Prvi primer praktične uporabe predstavlja napovedovanje koncentracij žveplovega dioksida (SO_2) na kompleksnem geografskem področju s pomočjo večnivojske perceptronske nevronske mreže [16]. Omenjena nevronska mreža je uporabljena tudi v dveh drugih praktičnih primerih: za napovedovanje hitrosti in smeri vetra na kompleksnem geografskem področju in za rekonstrukcijo meritev SODAR-ja¹. Meritve SODAR-ja v različnih plasteh nad kotlino so bile rekonstruirane iz talnih meritev vetra v kotlini in na sosednjih hribih.

Kohonenova nevronska mreža je uporabno orodje za samo-organizirajoče razvrščanje. Še posebno je uporabna za analizo gruč. V prvem praktičnem primeru uporabe je razloženo, kako je bila Kohonenova nevronska mreža uporabna za iskanje skupin vetrovnih polj na majhnem kompleksnem geografskem področju. Vetrovno polje je bilo predstavljeno s šestimi talnimi meritvami smeri in hitrosti vetra v kotlini in na sosednjih hribih. V študiji je bila za preiskovanje uporabljena velika množica meritev. Kohonenova nevronska mreža je razvrstila množico podatkov v približno trideset skupin. Za ocenitev kvalitete razvrščanja so izrisane tudi rože vetrov za vsako merilno točko po skupinah. Grafična predstavitev izbranih skupin je dokazala izredno učinkovitost razvrščanja. Drugi primer učinkovite uporabe Kohonenove nevronske mreže pa prikazuje uporabo v pripravljalni fazi izgradnje napovedovalnega modela onesnaženja ozračja. V ta namen je bila razvita

¹ SODAR (Sonic Detection And Ranging) sistemi se uporabljajo za merjenje na daljavo strukture vertikalne turbulence in profila vetra v nižjih plasteh atmosfere. SODAR sistemi so podobni radarskim sistemom, (radio detection and ranging), razlikujejo se samo v tem, da so namesto radijskih valov uporabljeni za zaznavanje zvočni valovi. Za SODAR sisteme se uporabljajo tudi druga imena kot na primer "sounder", "echosounder" ali akustični radar. Večina SODAR sistemov deluje tako, da najprej odda zvočni pulz (pisk) in zatem kratek čas posluša vračajoč se signal. V splošnem se v vračajočem se signalu analizirata tako jakost kot Dopplerjev (frekvenčni) premik. S tem je se ugotovi hitrost vetra, smer vetra in turbulenten značaj atmosfere. Profil atmosfere, ki je funkcija višine, se pridobi z analizo vračajočega signala v zaporednih časovnih intervalih, ki sledijo oddaji vsakega zvočnega impulza. Vračajoč se signal, ki je zabeležen ob določenem časovnem zamiku, vsebuje podatke o atmosferi na višini, ki jo je mogoče izračunati iz hitrosti zvoka. Maksimalni doseg običajnih SODAR sistemov se giblje med nekaj sto metrov do nekaj tisoč metrov. Maksimalni doseg se običajno doseže na lokacijah, ki imajo nizek okoliški šum in zmerno do visoko relativno vlažnost. Na puščavskih lokacijah imajo SODAR sistemi omejeno višinsko zmogljivost, ker se zvok razredčuje hitreje v suhem zraku [8].

tehnika razvrščanja vzorcev s pomočjo Kohonenove nevronske mreže [15].

Obsežen pregled modelov na področju atmosferskih znanosti nam ponujata avtorja M. W. Gardner in S. R. Dorling v svojem članku [23], kjer se osredotočata predvsem na modeliranje s pomočjo umetnih nevronske mreže. V članku je prikazana uporabnost umetnih nevronske mreže, ki postajajo vse bolj uporabna alternativa tradicionalnim statističnim metodam. Predstavljeno je kar nekaj praktičnih primerov uporabe nevronske mreže za aproksimacijo, napovedovanje in klasifikacijo v atmosferskih znanostih.

Omenjena avtorja sta v drugem članku [24] tudi predstavila uporabo večnivojske preceptronske nevronske mreže za aproksimacijo funkcije za boljše modeliranje relacij med podatki. Prikazan je bil primer modeliranja relacij med urnimi koncentracijami ozona in različnimi meteorološkimi spremenljivkami na lokaciji ob obali. Glavni cilj tega projekta je bilo oceniti pomembnost meteorologije pri ugotavljanju urne koncentracije ozona. Nazorno je prikazano, da prejšnji poskusi reševanja tega problema niso bili primerni zaradi nelinearnega značaja problema. Linearna regresija je podcenjevala pomembnost meteorologije.

J. Yi in R. Prybutok sta se v svojem projektu [64] lotila napovedovanja koncentracij ozona v industrijskih predelih severne Amerike. Vhod v njun model predstavlja 9 spremenljivk: jutranja koncentracija ozona, maksimalna dnevna temperatura, koncentracije CO_2 , NO , NO_2 , NO_x , ter hitrost in smer vetra. Rezultati večnivojske nevronske mreže so bili boljši od rezultatov iz regresijske analize, pri čemer so bili uporabljeni isti podatki. Avtorja omenjata tudi, da je uporabljena nevronska mreža prekosila ARIMA (autoregressive integrated moving average) pristop modeliranja, čeprav bi morale biti take primerjave različnih tehnik modeliranja narejene bolj previdno. Za pošteno primerjavo obeh modelov bi bilo potrebno zgraditi oba modela z istimi podatki. V konkretnem primeru pa je bila nevronska mreža učena z vsemi devetimi meteorološkimi spremenljivkami, medtem, ko je bil ARIMA model zgrajen zgolj s časovnim potekom ozona.

Comrie je v svojem prispevku [17] naredil primerjavo napovedovanja ozona s pomočjo večnivojske preceptronske mreže in z regresijskimi modeli. Napovedovanje maksimalnih dnevnih urnih koncentracij ozona v različnih ameriških urbanih okoljih je bilo narejeno z uporabo dnevnih povprečnih meteoroloških podatkov, ki so v bližnji relaciji z ozonom:

maksimalna dnevna temperatura, povprečna dnevna temperatura, povprečna dnevna hitrost vetra in dnevna doza sončnega sevanja. Comrie je v svojem zaključku navedel, da pristop z umetnimi nevronskimi mrežami ni doprinesel dramatičnih izboljšav v primerjavi z linearnimi regresijskimi modeli za napovedovanje ozona s pomočjo meteorološkega stanja okolja. Vsi uprabljeni modeli pa so zelo slabo napovedovali epizode visokih koncentracij. Avtor zaključuje, da bi se dalo model še dodatno izboljšati z ustrezno prilagoditvijo ustrezni geografski lokaciji in z uporabo dodatnih vhodnih spremenljivk.

Raziskovanja in napovedovanja troposferskega ozona s pomočjo nevronskih mrež sta se lotila tudi S. A. Abdul-Wahab in S. M. Al-Alawi [1]. Model za napovedovanje ozona je zgrajen na osnovi predpostavke, da je koncentracija ozona funkcija meteoroloških pogojev in različnih parametrov ozračja. V svojem delu raziskovalca ugotavljata, da predstavljajo nevronske mreže uporabno orodje zaradi sposobnosti učenja iz historičnih podatkov in zmožnosti modeliranja nelinearnosti. Mrežo sta učila na podlagi meteoroloških podatkov in podatkov o koncentracijah² snovi v zunanjem zraku v poletnem času, ko so bile izmerjene najvišje koncentracije ozona. Vsi podatki so bili izmerjeni v urbanem okolju. Končni produkt njunega raziskovanja so bili trije modeli nevronskih mrež. Glavni namen prvega modela je bilo ugotavljanje parametrov, ki imajo največji vpliv na koncentracije ozona v časovnem intervalu 24 ur (vključeni so bili dnevni in nočni podatki). Drugi model je bil razvit za opazovanje vpliva posameznih parametrov na koncentracije ozona v dnevnem času, ko so bile izmerjene višje koncentracije ozona. Tretji, zadnji model, pa je bil razvit za napovedovanje maksimalnih dnevnih koncentracij ozona. V svojem delu sta ugotovila, da je znašal prispevek meteoroloških parametrov na variacijo ozona približno od 33% do 40%. Ugotovila sta tudi, da imajo dušikov oksid, žveplov dioksid, relativna vlažnost, nemetanski ogljikovodiki in dušikov dioksid najpomembnejši vpliv na napovedane koncentracije ozona. Dodatno je pomembno vlogo igrala tudi temperatura zraka, medtem, ko je imelo globalno sončno sevanje bistveno manjši vpliv, kot je bilo pričakovano.

G. Finzi in G. Nunnari sta se v svojem delu [21] lotila raziskave različnih pristopov modeliranja z uporabo nevronskih mrež in mehkih modelov. Poleg tega sta tudi ovrednotila rezultate realnih primerov modeliranja z

2 Koncentracija snovi v zunanjem zraku se izraža v masnih enotah (miligramih, mikrogramih) na enoto volumna zraka (m^3) pri temperaturi 293 K (20°C) in zračnem tlaku 101.3 kPa [59]

omenjenimi modeli na praktičnih primerih v urbanih in industrijskih področjih v Italiji.ocene rezultatov so predstavljene in razložene v obliki urnih in dnevnih indeksov učinkovitosti napovedovanja in statističnih indikatorjev. Z umetnimi nevronske mreže sta bila zgrajena dva modela za napovedovanje koncentracij ozona od 1 do 6 ur vnaprej v dveh mestih: Brescia (v severni Italiji) in Catania (na Siciliji). Za vhodne parametre modela so bile izbrane izmerjene urne vrednosti ozona in njegovih pozviročiteljev (CO , NO in NO_2 koncentracije), medtem ko avtorja poročata, da meteoroloških podatkov ni bilo na voljo. Pri oceni uporabnosti se je izkazalo, da imata modela slabši učinek pri napovedovanju izrednih situacij v primerjavi z modelom *sive škatle*, ki pa obratno povzroča preveliko število lažnih alarmov. V drugem primeru pa sta bila zgrajena dva modela na osnovi mehke logike za napovedovanje preseženih vrednosti koncentracij ozona v dveh mestih: Brescia in Siracusa (na Siciliji). Za vhodne podatke modela so bile tokrat na voljo izmerjene koncentracije snovi v zunanjem zraku in meteorološki parametri. Končna izbira dnevnih povprečnih vrednosti vhodov je bila sledeča: O_3 , NO_2 , NO_x , temperatura zraka, sončno sevanje, zračni pritisk in smer vetra. Končna analiza napovedi je pokazala relativno zadovoljivo učinkovitost napovedovanja izrednih dogodkov ter dobro izogibanje lažnim alarmom. Avtorja v svojem delu zaključujeta, da so pristopi modeliranja z uporabo nevronske mreže in mehke logike zelo obetajoči za napovedovanje izrednih dogodkov, potrebne pa so dodatne raziskave.

Ballester in ostali so se v svoji raziskavi [11] lotili napovedovanja urnih koncentracij za 24 ur vnaprej v treh mestih v Španiji: Paterna, Alcoy in Carcagente. Za napovedovanje so uporabili tri vrste modelov: ARMAX (Autoregressive-moving average with exogenous inputs), večnivojsko perceptronsko nevronske mreže in FIR nevronske mreže [54]. Posebnost večnivojskega perceptronskega modela je v tem, da uporablja za napovedovanje koncentracije ozona za eno uro vnaprej pretekle izmerjene vrednosti in že izračunane napovedi. Za takšno modeliranje dinamike procesa so bili za vhode v model izbrani sledeči regresorji: predhodne vrednosti ozona ($\text{O}_3(t-1)$, $\text{O}_3(t-2)$, ..., $\text{O}_3(t-24)$), meteoroloških podatkov (hitrost vetra($t-24$), smer vetra($t-24$), temperatura zraka ($t-24$), ...) in dušikovega dioksida ($\text{NO}_2(t-24)$). Učinkovitost opisanega modela je primerjana z ostalima dvema modeloma, za katera avtorji trdijo, da sta se izkazala slabše, še posebej FIR nevronske mreže, katere glavna omejitev je, da je bila za vhod uporabljena samo koncentracija

ozona brez ostalih meteoroloških podatkov. Pri analizi slabših napovedi so uporabili metodo gruč (cluster analysis). Izkazalo se je da, so bile slabše napovedi povezane z gručami, ki so bile manj razpršene. Za izboljšavo v takih primerih avtorji predlagajo uporabo dodatnih spremenljivk (na primer posebne vremenske elemente).

4. Metode za modeliranje in vrednotenje modelov

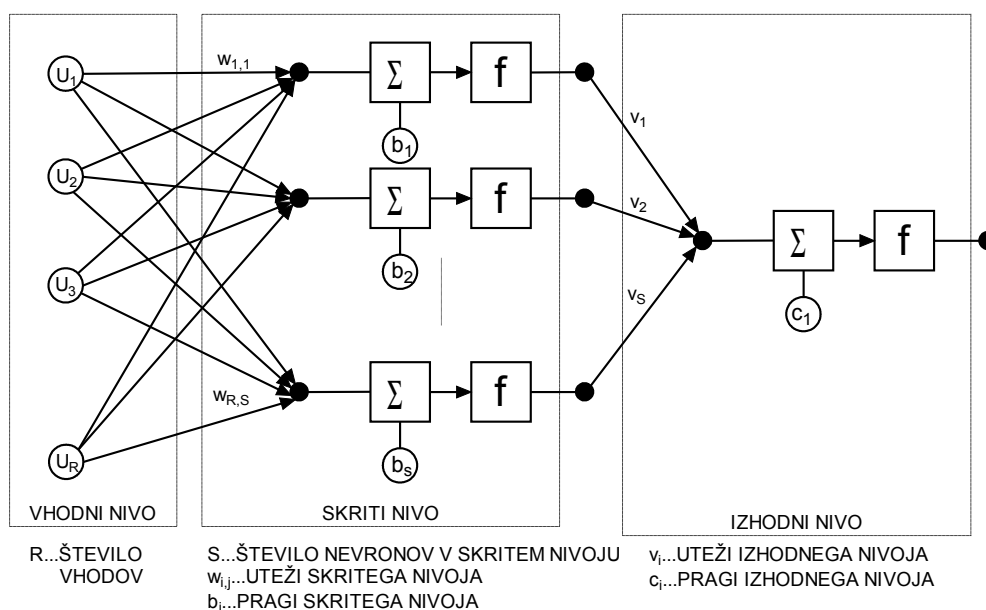
Za izgradnjo modela za napovedovanje povišanih koncentracij ozona se mi zdi uporaba modelov na osnovi umetne nevronske mreže, Gaussovih procesov in mehke logike glede na opisano kompleksnost ozona in predstavljene izkušnje ostalih raziskovalcev ustrezna. V nadaljevanju sem se najprej posvetil predstavitvi metod, ki sem jih uporabil za sintezo modela: umetnim nevronskim mrežam, Gaussovim procesom in mehki logiki. Posebno pozornost sem namenil Gaussovim procesom, ki predstavljajo eno izmed novejših metod za modeliranje.

Za primerjavo rezultatov modelov zgrajenih z različnimi metodami pa sem uporabil štiri različne metode določanje uspešnosti modelov na preizkusnih podatkih.

4.1. Umetne nevronske mreže

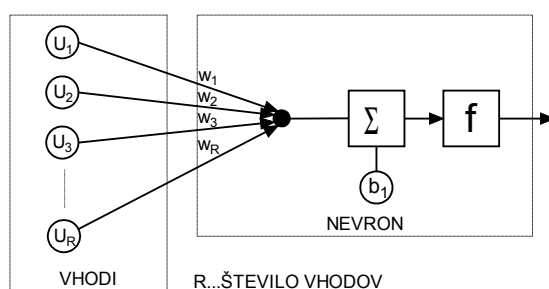
Kot je bilo že omenjeno, predstavljajo umetne nevronske mreže, natančneje večnivojske perceptronske mreže, učinkovito alternativo tradicionalnim statističnim metodam. Dokazano je bilo [29], da je možno uporabiti in naučiti večnivojsko perceptronsko mrežo za aproksimacijo kakršnekoli gladke omejene funkcije. V nasprotju z ostalimi statističnimi metodami ni potrebno pri učenju večnivojske perceptronske mreže upoštevati nobenih predpostavk o verjetnostni porazdelitvi podatkov. Mreža je sposobna modelirati izredno nelinearne funkcije in natančno posploševati pri učenju z novimi podatki. Predhodno je bilo opisanih kar nekaj praktičnih primerov uporabe nevronskih mrež v atmosfarski znanosti.

Večnivojska perceptronska mreža je sestavljena iz sistema povezanih preprostih nevronov (vozlišč), kot je prikazano na sliki 9 [54]. Vozlišča so razdeljena v več nivojev, običajno na vhodni nivo, poljubno število skritih nivojev in izhodni nivo. V vhodnem nivoju je število vozlišč enako številu vhodov, v izhodnem nivoju pa številu izhodov. V skritih nivojih pa je število nevronov poljubno in se ga običajno določi glede na izkušnje, ki se pridobijo pri procesu učenja nevronske mreže. V posameznih nivojih vozlišča niso med seboj povezana. Vsi izhodi posameznega nivoja so povezani preko uteži z vsemi vhodi naslednjega nivoja. Smer toka podatkov tudi določa ime naprej podajajoča nevronska mreža (feedforward neural network).

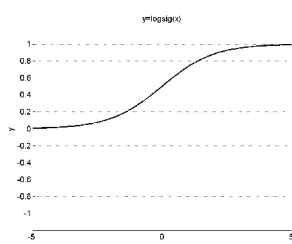


Slika 9: Večnivojska preceptronska mreža

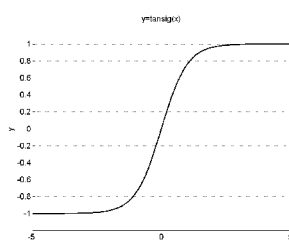
Nevron (vozlišče) predstavlja osnoven gradnik večnivojske vnaprejšne perceptronske nevronske mreže. Prikazuje ga slika 10. Vsak vhod v nevron je utežen z ustrezno utežjo $w_{i,R}$. Vsota vseh zmnožkov vhodov z utežmi in vrednostjo praga b tvorijo vhod v prenosno (aktivacijsko) funkcijo f , ki je lahko poljubna odvedljiva funkcija. Uporaba mnogo preprostih nelinearnih prenosnih funkcij v večnivojski vnaprejšni mreži lahko po predpostavki omogoča aproksimacijo izredno nelinearnih funkcij. V primeru uporabe linearnih prenosnih funkcij pa je možno modeliranje samo linearnih funkcij [29]. Običajno se uprabljata za modeliranje log-sigmoidna (slika 11) in tan-sigmoidna (slika 12) prenosni funkciji, ki imata preprosto izračunljiv odvod, občasno pa tudi linearna prenosna funkcija (slika 13).



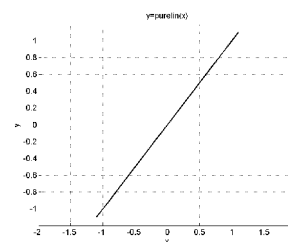
Slika 10: Nevron



Slika 11: Log-sigmoidna prenosna funkcija



Slika 12: Tan-sigmoidna prenosna funkcija



Slika 13: Linearna prenosna funkcija

Eno od ključnih vprašanj za uspešno modeliranje je določitev števila nevronov v skritih plasteh. To vprašanje je v literaturi [35] sicer obdelano, vendar ne daje zadovoljivih odgovorov na splošne primere. Število nevronov sem določal eksperimentalno. Osnova pa je bilo število nevronov, ki jih je predlagal programski paket *NeuroShell2* [43], izračunanih po naslednji formuli:

$$m = \frac{n+q}{2} + \sqrt{s} \quad (4.1)$$

Pri čemer je pomen oznak sledeč:

m ... število skritih nevronov,

n ... število vhodov,

q ... število izhodov in

s ... število učnih vzorcev.

Število skritih nevronov, ki ga izračunamo po enačbi (4.1) je odvisno od števila vzorcev v učni množici s katero učimo nevronske mreže. Ker se zavedam, da je obširna množica učnih vzorcev zagotovo redundantna, je torej tudi tako ocenjeno število skritih nevronov precenjeno. Zato sem to število tudi zmanjševal dokler nisem dobil zadovoljivih rezultatov. Navedena formula je sicer dober pripomoček za doseg cilja, se pa zavedam, da mora biti število skritih nevronov odvisno od zapletenosti procesa in ne od števila učnih vzorcev, ki so na voljo, saj se le-ti lahko ponavljajo. V primeru prevelikega števila skritih nevronov se lahko nevronska mreža predobro nauči učne podatke s čimer se zmanjša njena zmožnost posploševanja. Obratno pa lahko pride do prevelikega posploševanja v primeru premajhnega števila skritih nevronov.

Po določitvi števila skritih plasti, števila nevronov in začetnih vrednosti uteži in pragov, je nevronska mreža pripravljena za učenje. Učenje

zahteva množico učnih podatkov, ki so sestavljeni iz dvojic {vhodni vektor-izhodni vektor}, kar pomeni, da učenje nevronske mreže spada v obliko nadzorovanega učenja (supervised learning). Med učenjem so nevronske mreže večkrat predstavljene učni podatki. Ob vsaki predstavitvi (iteraciji) se uteži spreminja (prilagaja), dokler ni dosežena ustrezna vhodno-izhodna relacija. Proces učenja dejansko predstavlja iterativni postopek, pri katerem nastavimo vrednosti uteži tako, da minimiziramo določeno kriterijsko funkcijo (običajno je to funkcija povprečne kvadratne napake).

Za učenje perceptronskih nevronske mreže je na voljo mnogo algoritmov, najbolj pogosto uporabljen je algoritem vzvratnega širjenja (backpropagation) [40,15,54,50]. O kvaliteti učenja odločata predvsem dva nastavljava parametra, to sta hitrost učenja in moment. Hitrost učenja pove v kolikšni meri se pri tekočem popravku uteži upoštevajo izračunane napake, moment pa določa, koliko se pri tekočem popravku uteži upoštevajo predhodni popravki uteži. Premajhna hitrost učenja ima za posledico veliko potratu časa, povečana pa je tudi možnost, da bo mreža ostala v lokalnem minimumu, ki ga bo našla. Nasprotno pa prevelika hitrost učenja lahko povzroči oscilacije. Nizke vrednosti momenta pomenijo nevarnost oscilacij med učenjem in počasno učenje. Do oscilacij pa lahko pride tudi v primeru previsoke vrednosti momenta.

4.2. Gaussovi procesi

Teorija Gaussovih procesov (GP) izhaja iz področja verjetnosti in statistike, njihova uporaba za identifikacijo nelinearnih dinamičnih sistemov pa predstavlja zelo mlado vejo raziskovanja. Kot osnovo za identifikacijo z GP se uporablja njihova podobnost z umetnimi nevronske mrežami. Rezultat predikcije z modelom GP je normalno porazdeljena gostota verjetnosti modelirane izhodne vrednosti procesa s srednjo vrednostjo in varianco. To je tudi njihova prednost, saj v nasprotju z drugimi nelinearnimi modeli, dobimo tudi informacijo o zaupanju v napovedano vrednost izhoda. Osnove verjetnostnega modeliranja povzete po [36], ki so potrebne za razumevanje modeliranja z Gaussovimi procesi, so podane v dodatku A.

4.2.1. Bayesovo modeliranje

Najprej definirajmo problem, ki ga želimo rešiti [26,47]. Podane imamo izmerjene podatke $\mathbf{D} = \{\mathbf{X}, \mathbf{t}\} = \{(\mathbf{x}_n, t_n), n=1..N\}$ (izmerjene vrednosti vsebujejo tudi šum), ki so sestavljeni iz N parov vhodnih vektorjev $\{\mathbf{x}_n\}$

(dimenzija vhodnega vektorja je L) in skalarnih izhodov $\{t_n\}$. Iz podanih podatkov želimo sestaviti model s pomočjo katerega bomo lahko izdelovali napovedi. Glede na novi vhodni vektor x_{N+1} želimo poiskati verjetnostno porazdelitev ustreznega izhoda t_{N+1} . Splošen model za opisan problem predstavlja enačba

$$t_n = y(\mathbf{x}_n) + v_n \quad (4.2)$$

kjer označuje $y(\mathbf{x}_n)$ funkcijo modela in v_n šum. V prostoru možnih funkcij, ki bi bile lahko uporabljene za model, definirajmo predpostavljeno verjetnostno porazdelitev $p(y|\mathbf{A})$, kjer je z $\mathbf{A}$ označena množica hiperparametrov. Poleg tega definirajmo tudi prepostavko o verjetnostni porazdelitvi šuma $p(v|\mathbf{B})$, kjer v označuje vektor šumnih vrednosti $v=(v_1, v_2, \dots, v_n)$ in $\mathbf{B}$ označuje množico hiperparametrov modela šuma. Glede na podane hiperparametre lahko z enačbo

$$p(\mathbf{t}_N | \{\mathbf{x}_n\}, \mathbf{A}, \mathbf{B}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(\mathbf{t}_N | \{\mathbf{x}_n\}, y, v) p(y|\mathbf{A}) p(v|\mathbf{B}) dy dv \quad (4.3)$$

zapišemo verjetnostno porazdelitev podatkov, kjer je $\mathbf{t}_N = (t_1, t_2, \dots, t_N)$.

Če sedaj definiramo še vektor $\mathbf{t} = (t_1, t_2, \dots, t_N, t_{N+1})$, lahko zapišemo pogojno verjetnostno porazdelitev vrednosti t_{N+1} z enačbo (4.4) in s tem tudi uporabimo pogojno verjetnostno porazdelitev za izdelavo napovedi o vrednosti t_{N+1} .

$$P(t_{N+1} | \mathbf{D}, \mathbf{A}, \mathbf{B}, \mathbf{x}_{N+1}) = \frac{P(\mathbf{t}_{N+1} | \{\mathbf{x}_n\}, \mathbf{A}, \mathbf{B}, \mathbf{x}_{N+1})}{P(\mathbf{t}_N | \{\mathbf{x}_n\}, \mathbf{A}, \mathbf{B})} \quad (4.4)$$

Na tej točki je potrebno poudariti, da je potrebno ločiti t_{N+1} (neznana izhodna vrednost pri vhodnem vektorju $\mathbf{x}_{N+1}$) od $\mathbf{t}_{N+1}$ (vektor, ki je sestavljen iz $\mathbf{t}_N$ in t_{N+1}).

Z vidika Bayesovega verjetnostnega modeliranja, modeliranje statičnega procesa z nelinearno funkcijo $f=f(\mathbf{x}, \boldsymbol{\omega})$, katere parametre $\boldsymbol{\omega}=\{\mathbf{A}, \mathbf{B}\}$ določimo glede na podatke $\mathbf{D}$, lahko predstavimo v obliki Bayesovega teorema [36,26,47]:

$$p(\boldsymbol{\omega} | \mathbf{D}) = \frac{p(\mathbf{D} | \boldsymbol{\omega}) p(\boldsymbol{\omega})}{p(\mathbf{D})} \quad (4.5)$$

kjer je:

$p(\boldsymbol{\omega})$...“*a-priori*” verjetnostna porazdelitev, ki vsebuje predhodno znanje o parametrih funkcije (ang. prior), ki ponavadi predpostavljajo zveznost, frekvenčno razporeditev moči itd,

$p(\mathbf{D} | \boldsymbol{\omega})$...verjetnostna porazdelitev učne množice pri danih parametrih

funkcije (ang. likelihood),

$p(\mathbf{D})$...verjetnostna porazdelitev učne množice (ang. evidence), v primeru Bayesovega modeliranja služi kot normalizacijska konstanta,

$p(\boldsymbol{\omega}|\mathbf{D})$...“a-posteriori” verjetnostna porazdelitev parametrov $\boldsymbol{\omega}$ pri dani učni množici $\mathbf{D}$ (ang. posterior).

Bayesov teorem združuje predhodno znanje o parametrih $p(\boldsymbol{\omega})$ z znanjem, dobljenim v obliki učne množice $\mathbf{D}$. Z upoštevanjem zapisa $\mathbf{D}=\{\mathbf{X},\mathbf{t}\}$ lahko Bayesov teorem zapišemo v obliki:

$$p(\boldsymbol{\omega}|\mathbf{X},\mathbf{t})=\frac{p(\mathbf{t}|\mathbf{X},\boldsymbol{\omega})p(\boldsymbol{\omega})}{p(\mathbf{t}|\mathbf{X})} \quad (4.6)$$

4.2.2. Modeliranje z Gaussovimi procesi

Z neparametričnimi metodami določamo predikcijo izhoda brez eksplicitne parametrizacije funkcije f . Tak primer je tudi modeliranje z Gaussovimi procesi [26,47,36].

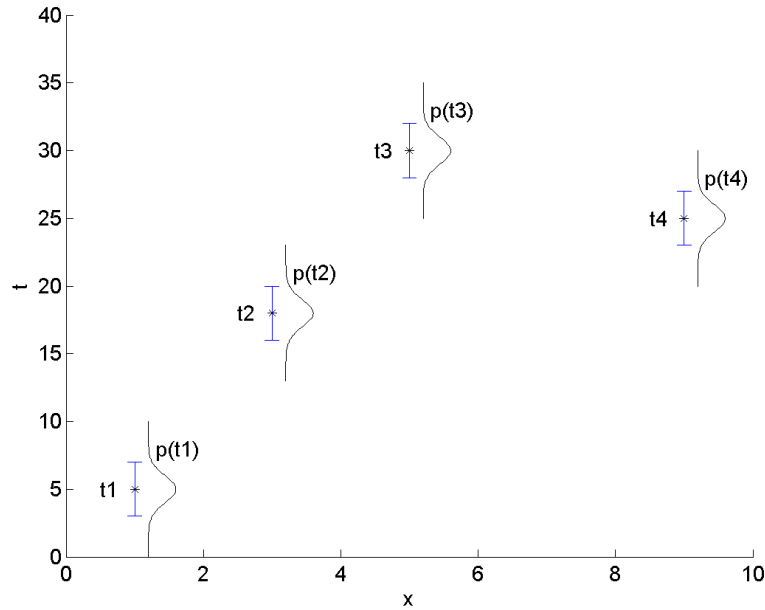
Gaussov proces predstavlja množica naključnih spremenljivk $\mathbf{t}=(t_1(x_1),t_2(x_2),\dots)$, ki imajo skupno verjetnostno porazdelitev opisano z enačbo

$$P(\mathbf{t}_N|\mathbf{C},\{\mathbf{x}_n\})=\frac{1}{Z}e^{-\frac{1}{2}(\mathbf{t}-\boldsymbol{\mu})^T\mathbf{C}^{-1}(\mathbf{t}-\boldsymbol{\mu})} \quad (4.7)$$

za poljubno množico vhodov $\{\mathbf{x}_n\}$, kjer Z predstavlja ustrezno normirno konstanto. V primerjavi z enačbo (A.20) iz dodatka A je kovariančna matrika v tem primeru označena z $\mathbf{C}$, konstanta Z pa definira izraz

$$Z=(\sqrt{2\cdot\pi})^n\sqrt{|\mathbf{C}|} \quad (4.8)$$

Na sliki 14 je predstavljen primer Gaussovega procesa z množico štirih naključnih spremenljivk $\mathbf{t}=(t_1(x_1),t_2(x_2),t_3(x_3),t_4(x_4))$. Vsaka naključna spremenljivka t_n predstavlja eno učno točko. Porazdelitev naključne spremenljivke $p(t_n)$ v učni točki t_n določa kovariančna matrika $\mathbf{C}$, katere dimenzija je v prikazanem praktičnem primeru enaka 4×4 . Preprost in nazoren praktičen primer modeliranja nelinearne funkcije je podan v dodatku B.

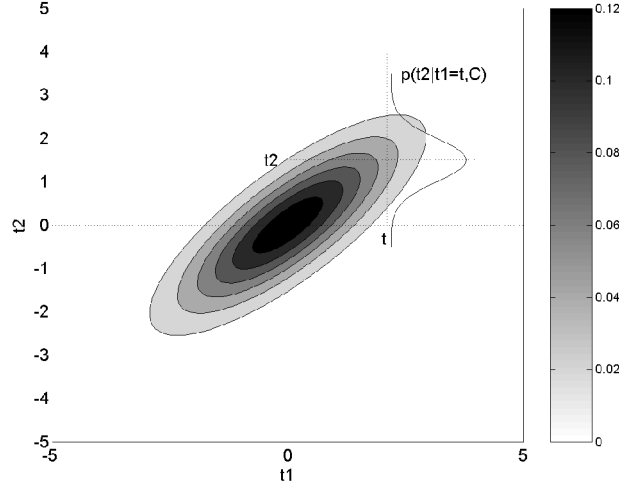


Slika 14: Primer Gaussovega procesa

Kovariančna matrika podatkov $\mathbf{C}$ je definirana s parametrično kovariančno funkcijo $C^{mn}(\mathbf{x}_m, \mathbf{x}_n; \boldsymbol{\theta})$, kjer se označuje $C_{mn} = C^{mn}(\mathbf{x}_m, \mathbf{x}_n; \boldsymbol{\theta})$, $\boldsymbol{\theta}$ pa označuje množico hiperparametrov. Določanje vrednosti hiperparametrov in dimenzije množice hiperparametrov bo predstavljeno v nadaljevanju. Zaenkrat pa za naš praktičen primer predpostavimo, da imamo že določeno množico hiperparametrov in kovariančno funkcijo. Indeks mn zgoraj označuje, da kovariančna funkcija razlikuje med $\mathbf{x}_m$ in $\mathbf{x}_n$ tudi, ko velja $\mathbf{x}_m = \mathbf{x}_n$. To je še posebej pomembno pri obravnavi modelov šuma v okviru kovariančne funkcije. V nadaljevanju bo indeks zaradi preglednosti opuščen, kljub temu pa se nanaša na vse kovariančne funkcije v nadaljevanju. Z $\boldsymbol{\mu}$ je označen vektor srednjih vrednosti. Glavni razlog, da je množica hiperparametrov označena s $\boldsymbol{\theta}$ je v tem, da imajo parametri kovariančne funkcije podobno vlogo kot parametri nevronske mreže.

Sedaj predpostavimo, da je Gaussov proces s kovariančno matriko $\mathbf{C}_N$ in vektorjem srednjih vrednosti $\boldsymbol{\mu}$ ustvaril vektor podatkov $\mathbf{t}_N$ tako kot opisuje enačba (4.7). Očitno je, da ni možno vseh podatkov modelirati s procesom, ki bi imel srednjo vrednost enako 0. V nadaljevanju bomo predpostavljali, da so bili podatki ustrezno skalirani, tako da zagotavljajo domnevno o srednji vrednosti enaki 0. Z eksplicitno definicijo verjetnostne porazdelitve podatkov smo sedaj preskočili korak izražanja

individualnih predpostavk o šumu in modelirni funkciji. Obe predpostavki sta združeni v kovariančni matriki $\mathbf{C}$. Iz predstavljenega je razvidno, da je narava kovariančne matrike $\mathbf{C}$ in posledično kovariančne funkcije bistvena pri modeliranju z Gaussovimi procesi. Kovariančna matrika bo podrobneje predstavljena v naslednjem podpoglavju.



Slika 15: Pogojna porazdelitev

Funkcija zapisana z enačbo (4.7) izraža korelacijo med različni točkami v vhodnem prostoru, kar lahko najbolje ilustriramo s preprostim primerom. Upoštevajmo sliko 15. Slika prikazuje gostoto skupne Gaussove verjetnostne porazdelitve dveh točk (naključnih spremenljivk) t_1 in t_2 , ki ju določa kovariančna matrika $\mathbf{C}$. Pri podani vrednosti $t_1=t$ lahko najdemo Gaussovo pogojno verjetnostno porazdelitev $P(t_2|t_1=t, \mathbf{C})$ in iz tega tudi najbolj verjetno napoved vrednosti t_2 pri podani vrednosti $t_1=t$. Skala in orientacija gostote skupne Gaussove verjetnostne porazdelitve je popolnoma določena z kovariančno matriko, ki zopet kaže na glavno vlogo $\mathbf{C}$ pri modeliranju z Gaussovimi procesi.

V splošnem primeru napovedovanja t_{N+1} s podanimi podatki $D = (\{\mathbf{x}_n\}, \mathbf{t}_n)$ lahko uporabimo enačbo (4.7) ter dejstvo, da je skupna verjetnostna porazdelitev množice podatkov Gaussova, za izračun Gaussove pogojne verjetnostne porazdelitve t_{N+1} .

$$P(t_{N+1}|D, C(\mathbf{x}_m, \mathbf{x}_n; \Theta), \mathbf{x}_{N+1}) = \frac{P(\mathbf{t}_{N+1}|C(\mathbf{x}_m, \mathbf{x}_n; \Theta), \mathbf{x}_{N+1}, \{\mathbf{x}_n\})}{P(\mathbf{t}_N|C(\mathbf{x}_m, \mathbf{x}_n; \Theta), \{\mathbf{x}_n\})} \quad (4.9)$$

$$P(t_{N+1}|D, C(\mathbf{x}_m, \mathbf{x}_n; \Theta), \mathbf{x}_{N+1}) = \frac{Z_N}{Z_{N+1}} e^{-\frac{1}{2}(\mathbf{t}_{N+1}^T \mathbf{C}_{N+1}^{-1} \mathbf{t}_{N+1} - \mathbf{t}_N^T \mathbf{C}_N^{-1} \mathbf{t}_N)} \quad (4.10)$$

Z_N in Z_{N+1} označujeta primerni normirni konstanti. Enačba

$$\mathbf{C}_{N+1} = \begin{bmatrix} \mathbf{C}_N & \mathbf{k}_{N+1} \\ \mathbf{k}_{N+1}^T & \kappa \end{bmatrix} \quad (4.11)$$

prikazuje relacije med $\mathbf{C}_{N+1}$ ter $\mathbf{C}_N$, kjer je pomen oznake

$$\kappa = C(\mathbf{x}_{N+1}, \mathbf{x}_{N+1}; \Theta) \quad (4.12)$$

ter oznake

$$\mathbf{k}_{N+1} = (C(\mathbf{x}_1, \mathbf{x}_{N+1}; \Theta), \dots, C(\mathbf{x}_N, \mathbf{x}_{N+1}; \Theta)) \quad (4.13)$$

Gaussovo pogojno verjetnostno porazdelitev t_{N+1} pri $\mathbf{x}_{N+1}$ lahko izpeljemo iz enačbe (4.10) po postopku opisanem v nadaljevanju. Če določimo inverz matrike $\mathbf{C}_{N+1}$ z enačbo

$$\mathbf{C}_{N+1}^{-1} = \begin{bmatrix} \mathbf{M}_N & \mathbf{m}_{N+1} \\ \mathbf{m}_{N+1}^T & \mu \end{bmatrix} \quad (4.14)$$

in uporabimo dejstvo, da je

$$\mathbf{C}_{N+1} \mathbf{C}_{N+1}^{-1} = \mathbf{I}_{N+1} \quad (4.15)$$

lahko zapišemo

$$\mathbf{C}_N \mathbf{M}_N + \mathbf{k}_{N+1} \mathbf{m}_{N+1}^T = \mathbf{I}_N \quad (4.16)$$

$$\mathbf{C}_N \mathbf{m}_{N+1} + \mu \mathbf{k}_{N+1} = \mathbf{0} \quad (4.17)$$

$$\mathbf{k}_{N+1}^T \mathbf{M}_N + \kappa \mathbf{m}_{N+1}^T = \mathbf{0}^T \quad (4.18)$$

$$\mathbf{k}_{N+1}^T \mathbf{m}_{N+1} + \kappa \mu = 1 \quad (4.19)$$

S premultiplikacijo enačbe (4.16) z C_N^{-1} in nadomestitvijo dobljenega izraza za M_N v enačbi (4.18) ter upoštevanjem $C_N^{-T} = C_N^{-1}$ dobimo

$$m_{N+1} = \frac{-C_N^{-1} k_{N+1}}{(\kappa - k_{N+1}^T C_N^{-1} k_{N+1})} . \quad (4.20)$$

Če dobljeni izraz vstavimo v enačbo (4.17), dobimo

$$\mu = \frac{1}{\kappa - k_{N+1}^T C_N^{-1} k_{N+1}} . \quad (4.21)$$

Izraz (4.20) lahko zapišemo z upoštevanjem enačbe (4.21) tudi v drugačni obliki

$$m_{N+1} = -\mu C_N^{-1} k_{N+1} \Rightarrow m_{N+1}^T = -\mu k_{N+1}^T C_N^{-T} = -\mu k_{N+1}^T C_N^{-1} . \quad (4.22)$$

S pomočjo izrazov (4.16) in (4.22) pa lahko določimo tudi izraz za

$$M_N = C_N^{-1} - C_N^{-1} k_{N+1} m_{N+1}^T = C_N^{-1} + \mu C_N^{-1} k_{N+1} (C_N^{-1} k_{N+1})^T . \quad (4.23)$$

Če se vrnemo nazaj k enačbi (4.10) lahko zapišemo odvisnost eksponenta enačbe (4.10) z elementi matrike C_{N+1}^{-1} definirane v enačbi 4.14

$$\begin{aligned} t_{N+1}^T C_{N+1}^{-1} t_{N+1} - t_N^T C_N^{-1} t_N &= \begin{bmatrix} t_N^T & t_{N+1} \end{bmatrix} \begin{bmatrix} M_N & m_{N+1} \\ m_{N+1}^T & \mu \end{bmatrix} \begin{bmatrix} t_N \\ t_{N+1} \end{bmatrix} - t_N^T C_N^{-1} t_N = \\ &= \mu t_{N+1}^2 + t_{N+1} (m_{N+1}^T t_N + t_N^T m_{N+1}) + t_N^T M_N t_N - t_N^T C_N^{-1} t_N = \\ &= \mu t_{N+1}^2 + 2 t_{N+1} m_{N+1}^T t_N + t_N^T M_N t_N - t_N^T C_N^{-1} t_N \end{aligned} \quad (4.24)$$

Z uporabo izrazov (4.22) in (4.23) je nadaljevanje izpeljave (4.24) sledeče

$$\begin{aligned} t_{N+1}^T C_{N+1}^{-1} t_{N+1} - t_N^T C_N^{-1} t_N &= \\ &= \mu t_{N+1}^2 - \mu 2 t_{N+1} k_{N+1}^T C_N^{-1} t_N + t_N^T C_N^{-1} t_N + \mu t_N^T C_N^{-1} k_{N+1} (C_N^{-1} k_{N+1})^T t_N - t_N^T C_N^{-1} t_N = \\ &= \mu (t_{N+1}^2 - 2 t_{N+1} (k_{N+1}^T C_N^{-1} t_N) + ((C_N^{-1} k_{N+1})^T t_N)^T (k_{N+1}^T C_N^{-1} t_N)) = \\ &= \mu (t_{N+1}^2 - 2 t_{N+1} (k_{N+1}^T C_N^{-1} t_N) + (k_{N+1}^T C_N^{-1} t_N)^T (k_{N+1}^T C_N^{-1} t_N)) = \\ &= \mu (t_{N+1} - (k_{N+1}^T C_N^{-1} t_N))^2 = \\ &= \frac{(t_{N+1} - (k_{N+1}^T C_N^{-1} t_N))^2}{(\kappa - k_{N+1}^T C_N^{-1} t_N)} \end{aligned} \quad (4.25)$$

Če v rezultatu izpeljave (4.25) uporabimo dve novi oznaki

$$\hat{t}_{N+1} = \mathbf{k}_{N+1}^T \mathbf{C}_N^{-1} \mathbf{t}_N \quad \text{in} \quad (4.26)$$

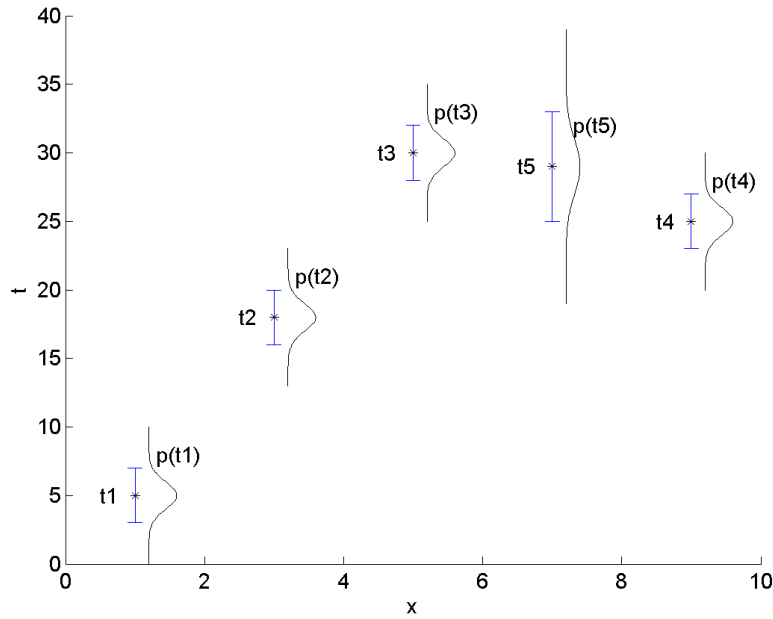
$$\sigma_{\hat{t}_{N+1}}^2 = \kappa - \mathbf{k}_{N+1}^T \mathbf{C}_N^{-1} \mathbf{k}_{N+1} \quad , \quad (4.27)$$

lahko enačbo za izračun Gaussove pogojne verjetnostne porazdelitve t_{N+1} preoblikujemo v novo obliko

$$P(t_{N+1} | D, C(\mathbf{x}_m, \mathbf{x}_n; \Theta), \mathbf{x}_{N+1}) = \frac{1}{Z} e^{\left(-\frac{(t_{N+1} - \hat{t}_{N+1})^2}{2\sigma_{\hat{t}_{N+1}}^2}\right)} \quad , \quad (4.28)$$

kjer $\hat{t}_{N+1}$ označuje srednjo vrednost napovedi v novi točki, $\sigma_{\hat{t}_{N+1}}$ pa standardno deviacijo napovedi. Pomembno je tudi poudariti, da obe enačbi vsebujeta samo $\mathbf{C}_N$ in ne $\mathbf{C}_{N+1}$. Posledično potrebujemo za napovedovanje pri konstatni učni množici ter pri konstantnih hiperparametrih kovariančne funkcije samo en izračun inverzne matrike reda $N \times N$.

Glede na primer modeliranja s štirimi učnimi točkami, ki je bil podan na sliki 14 na začetku podpoglavja, je prikazan na sliki 16 primer napovedi v novi točki t_5 . Kot je nazorno ilustrirano v dodatku C, sledi iz enačbe (4.27), da je negotovost, ki jo določa kovariančna funkcija, v novi napovedani točki večja od negotovosti v ostalih učnih točkah. Negotovost se z oddaljevanjem od učnih točk povečuje, z bližanjem učnim točkam pa zmanjšuje. V nobenem primeru pa negotovost ne more biti manjša, kot je v najbližji učni točki.



Slika 16: Porazdelitev v novi točki

4.2.3. Kovariančna funkcija

Kot je bilo že predhodno omenjeno, je vloga kovariančne funkcije najbolj pomembna pri modeliranju z Gaussovimi procesi [26]. Napovedane verjetnostne porazdelitve, ki nastopajo pri podanih podatkih, so v glavnem odvisne od kovariančne funkcije in njenih hiperparametrov. V nadaljevanju sledi obravnava različnih omejitev pri obliki kovariančnih funkcij in nekateri primeri kovariančnih funkcij za različne probleme.

4.2.4. Stacionarne kovariančne funkcije

Za kovariančno matriko $\mathbf{C}_N$ obstaja samo ena omejitev. Kovariančna matrika $\mathbf{C}_N$, ki jo določa kovariančna funkcija $C(\mathbf{x}_m, \mathbf{x}_n; \Theta)$ in množica vhodov $\{\mathbf{x}_n\}_{n=1}^N$, mora biti pozitivno definitna za katerokoli rang N in katerokoli množico vhodov.

Najprej preučimo samo stacionarne kovariančne funkcije, ki so odvisne samo od relativne pozicije dveh vhodnih vektorjev in zadoščajo enačbi

$$C(\mathbf{x}_i, \mathbf{x}_i + \mathbf{h}; \Theta) \equiv C(\mathbf{h}; \Theta) \quad . \quad (4.29)$$

Taka kovariančna funkcija $C(\mathbf{h}; \Theta)$ zadošča pogoju pozitivne definitnosti, če velja zanjo pozitivna omejena simetrična metrika $G(\cdot)$, kot je predstavljena v enačbi

$$C(\mathbf{x}_m - \mathbf{x}_n; \Theta)_{stat} = C(\mathbf{h}; \Theta)_{stat} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{i\mathbf{w}^T \mathbf{h}} G(d\mathbf{w}) \quad (4.30)$$

Enačba je znana pod imenom Bochnerjev teorem [12]. Enostavno se ga da prikazati z upoštevanjem izraza (4.30) o pozitivni definitnosti za katerikoli realni vektor $\mathbf{r}$.

$$\mathbf{r}^T \mathbf{C} \mathbf{r} > 0 \quad (4.31)$$

Glede na obliko stacionarne kovariančne funkcije lahko to zapišemo z enačbo

$$\sum_{m,n} r_m r_n C_{stat}(\mathbf{x}_m - \mathbf{x}_n; \Theta) > 0 \quad (4.32)$$

Če sedaj uporabimo v enačbi (4.32) izraz za $C_{stat}(\mathbf{h})$ iz enačbe (4.30) dobimo enačbo

$$\begin{aligned} \sum_{m,n} r_m r_n C_{stat}(\mathbf{x}_m - \mathbf{x}_n; \Theta) &= \sum_{m,n} r_m r_n \int e^{i\mathbf{w}^T (\mathbf{x}_m - \mathbf{x}_n)} G(d\mathbf{w}) \\ &= \int \left(\sum_m r_m e^{i\mathbf{w}^T \mathbf{x}_m} \right) \left(\sum_n r_n e^{-i\mathbf{w}^T \mathbf{x}_n} \right) G(d\mathbf{w}) \\ &= \int \left| \sum_m r_m e^{i\mathbf{w}^T \mathbf{x}_m} \right|^2 G(d\mathbf{w}) \end{aligned} \quad (4.33)$$

Iz tega sledi, da je $C_{stat}(\mathbf{h}; \Theta)$ pozitivno definitna za katerokoli množico vhodov $\{\mathbf{x}_m\}$, če je glede na enačbo (4.30) $G(\cdot)$ pozitivna metrika.

Dejansko enačba (4.30) ne omejuje drastično možnosti izbire kovariančnih funkcij. Primer je katerakoli Fourierjeva transformacija $f(\mathbf{h})$, kjer $f(\mathbf{h}) \geq 0$ for all $\mathbf{h}$ zadošča zahtevi o pozitivni definitnosti. Za mnogo primerov modeliranja pa želimo, da vsebuje kovariančna funkcija dve lastnosti.

1. Za večino primerov modeliranja pričakujemo, da je korelacija med dvema točkama, ki sta si blizu v vhodnem prostoru (glede na določeno metriko), večja od korelacije med dvema bolj oddaljenima točkama.
2. Pogosto ugotavljamo, da se v podatkih nahaja tudi šum, zato želimo vpeljati v kovariančno funkcijo tudi model šuma.

Ker je vsota dveh pozitivno definitnih funkcij zopet pozitivno definitna funkcija, bomo razdelili definicijo kovariančne funkcije na dva dela. Prvi del bo predstavljal obliko funkcijo s katero želimo modelirati naše podatke, drugi del pa bo povezan z modelom šuma

$$C_{stat}(\mathbf{x}_m - \mathbf{x}_n; \Theta) = C_f(\mathbf{x}_m - \mathbf{x}_n; \Theta) + \delta_{mn} \Theta_3 \quad (4.34)$$

Z δ_{mn} je označena Kronecker-jeva delta funkcija

$$\delta_{mn} = \begin{cases} 0; m \neq n \\ 1; m = n \end{cases} \quad (4.35)$$

Eno izmed možnih oblik C_f prikazuje enačba (4.36), kjer $x_n^{(l)}$ označuje l -to komponento vektorja $\mathbf{x}_n$.

$$C_f(\mathbf{x}_m, \mathbf{x}_n; \Theta) = \Theta_1 e^{-\frac{1}{2} \sum_{l=1}^L \frac{(x_m^{(l)} - x_n^{(l)})^2}{2r_l^2}} + \Theta_2 \quad (4.36)$$

Zgornji izraz vsebuje prvo lastnost, ki jo zahtevamo od kovariančne funkcije. To pomeni, da so točke, ki so si blizu v vhodnem prostoru, močno korelirane in zato omogočajo močan odziv pri podobnih vrednostih t -jev. Definicija bližine je izražena v obliki izrazov množice dolžinskih meril $\{r_l\}$. Za vsak vhod obstaja ustrezna dolžinsko merilo, ki označuje razdaljo v določeni smeri preko katere pričakujemo pomembne spremembe t . Zelo veliko dolžinsko merilo označuje, da imajo napovedi izdelane z uporabo Gaussovih procesov, zelo majhno ali celo ničelno odvisno razmerje z ustreznim vhodom. Za tak vhod lahko rečemo, da je nepomemben. Takšna interpretacija je v veliki povezavi z avtomatskim določanjem značilnk [37], [42]. Hiperparameter Θ_1 določa skupno vertikalno skalo relativno glede na srednjo vrednost Gaussovega procesa v izhodnem prostoru. Hiperparameter Θ_2 pa določa vertikalno negotovost in s tem označuje kako daleč lahko pričakujemo, da bo srednja vrednost funkcije odstopala od srednje vrednosti Gaussovega procesa. Pomen posameznih hiperparametrov je najbolje predstavljen v preprostem primeru modeliranja nelinearne funkcije v dodatku B.

Drugi izraz v kovariančni funkciji $\delta_{mn}\Theta_3$ predstavlja model šuma. Predpostavljamo, da je šum naključen. Iz tega sledi, da ne pričakujemo nobene korelacije med šumom in določenimi izhodi ter, da drugi izraz vpliva samo na diagonalne elemente kovariančne matrike. Varianco šuma podaja Θ_3 , za katero se predpostavlja, da je v primeru stacionarne kovariančne funkcije neodvisna od vhodov (modeli šuma, ki so odvisni tudi od vhodov bodo predstavljeni v nadaljevanju).

Predstavljena kovariančna funkcija omogoča dobro temeljno obliko za delo. Obstajajo pa tudi druge uporabne oblike stacionarnih kovariančnih funkcij. Za primer poskušajmo izdelati model z uporabo periodične funkcije, ki ima znano valovno dolžino λ_l v vhodni smeri l . Kovariančna funkcija, ki vsebuje takšno periodičnost, je predstavljena v enačbi

$$C_f(\mathbf{x}_m, \mathbf{x}_n) = \Theta_1 e^{-\frac{1}{2} \sum_{l=1}^L \left(\frac{\sin\left(\frac{\pi}{\lambda_l}(x_m^{(l)} - x_n^{(l)})\right)}{r_l} \right)^2} \quad (4.37)$$

Funkcije, ki jih opisuje taka kovariančna funkcija nimajo močne korelacije samo med točkami, ki so si blizu, ampak tudi med točkami, ki so med seboj oddaljene za dolžinov λ_l v smeri l .

4.2.5. Nestacionarne kovariančne funkcije

Medtem, ko se da veliko večino množic podatkov efektivno modelirati z uporabo stacionarnih kovariančnih funkcij, obstajajo določeni primeri v katerih se mora kovariančna funkcija izražati tudi v obliki nestacionarnosti. V nadaljevanju je podan opis treh posebno uporabnih nestacionarnih oblik.

4.2.5.1. Linearni izraz

Pri daljšem pomiku stran od znanih podatkov, napovedi z uporabno kovariančne funkcije, kot je opisana z enačbo (4.36), težijo h konstatni vrednosti. V nasprotju s tem pa lahko predvidevamo, da obstaja neka linearna težnja v množici podatkov.

Predstavljajmo si ploskev, ki jo opisuje enačba $y(\mathbf{x}) = \sum_l a_l x^{(l)} + c$. Če imata množica $\{a_l\}$ in c obliko Gaussove verjetnostne porazdelitve s srednjo vrednostjo enako 0 in variancama σ_a in σ_c , potem je kovariančna funkcija za ploskev zapisana z enačbo

$$C_{lin}(\mathbf{x}_m, \mathbf{x}_n; \{\alpha\}) = \sum_{l=1}^L \sigma_a^2 x_m^{(l)} x_n^{(l)} + \sigma_c^2. \quad (4.38)$$

Podani izraz $\sum_l \alpha_l^2 x_m^{(l)} x_n^{(l)}$ doda h kovariančni funkciji v enačbi (4.36) še linearno komponento k napovedim. Napovedi, ki so bolj oddaljene od množice podatkov, ležijo na ploskvi opisani z enačbo

$$y(\mathbf{x}) = \sum_{l=1}^L \left(\sum_{n=1}^N \alpha_l x_n^{(l)} (\mathbf{C}_N^{-1} \mathbf{t}_n)_n \right) x^{(l)} + \left(\sum_{n=1}^N \theta_2 (\mathbf{C}_N^{-1} \mathbf{t}_n)_n \right). \quad (4.39)$$

Na žalost bo dolžinsko merilo r_l korelirano z hiperparametrom α_l , ker imata oba vpliv na variabilnost napovedi v l -ti smeri. To lahko povzroči sorodnost Θ z mnogimi oblikami, ki lahko povzročajo probleme pri pristopih h Gaussovimi procesom, ki temeljijo na optimizaciji hiperparametrov. Možno rešitev predstavlja prilagoditev podatkov z uporabo samo linearnega izraza na začetku modeliranja. Potem pa sledi spreminjanje hiperparametrov z uporabo drugih elementov kovariančne funkcije.

4.2.5.2. Vhodno odvisen model šuma

Predhodno smo obravnavali samo modele šuma, ki so bili neodvisni od vhodov. Povsem smiselno pa je pričakovati, da se stopnja šuma spreminja v odvisnosti od vhodnih podatkov. Vgradnja take odvisnosti v model z Gaussovim procesom je povsem naravna in je lahko definirana z izrazom (4.40), kjer $\{\phi_j(\mathbf{x})\}$ predstavlja množico baznih funkcij in $\{\beta_j\} \in \Theta$ primerne koeficiente.

$$\theta_3(\mathbf{x}_m; \Theta) = e^{(\sum_{j=1}^J \beta_j \phi_j(\mathbf{x}_m))} \quad (4.40)$$

Tak model šuma ni stacionaren, je pa pozitivno definiten, ker prispeva samo k diagonalnim elementom kovariančne matrike.

4.2.5.3. Prostorsko spremenljiva dolžinska merila

V običajni stacionarni funkciji, ki je bila opisana predhodno, predpostavljamo, da je v posamezni smeri vhodnega vektorja dolžinsko merilo konstantno. Z lahkoto pa si predstavljamo primer slabega modela glede na podatke. Zaradi tega pa lahko določimo, da postanejo $\{r_l\}$ funkcije $\mathbf{x}$. Ne moremo pa preprosto nadomestiti parametrizirano obliko za $r_l(\mathbf{x})$ v enačbi (4.36), ker s tem ne dobimo splošne pozitivno definitne kovariančne funkcije. Primer takšne kovariančne funkcije je podrobneje obdelan v [26] in ga opisuje enačba

$$C_f^{ns}(\mathbf{x}_m, \mathbf{x}_n; \Theta) = \Theta_1 \prod_l \sqrt{\frac{2 r_l(\mathbf{x}_m; \Theta) r_l(\mathbf{x}_n; \Theta)}{r_l^2(\mathbf{x}_m; \Theta) + r_l^2(\mathbf{x}_n; \Theta)}} e^{(-\sum_l \frac{(x_m^{(l)} - x_n^{(l)})^2}{r_l^2(\mathbf{x}_m; \Theta) + r_l^2(\mathbf{x}_n; \Theta)})} \quad (4.41)$$

4.2.6. Določanje hiperparametrov kovariančne funkcije

Za učne točke smo zgradili model. Sedaj pa moramo najti ustrezen način, da bomo lahko dosledno določili neznane vrednosti hiperparametrov Θ zgrajenega modela. V idealnem primeru bi želeli za tvorjenje napovedi izračunati integral preko vseh hiperparametrov, oziroma najti

$$P(t_{N+1} | \mathbf{x}_{N+1}, D, C(\cdot)) = \int P(t_{N+1} | \mathbf{x}_{N+1}, D, C(\cdot), \Theta) P(\Theta | D, C(\cdot)) d\Theta, \quad (4.42)$$

kjer $C(\cdot)$ predstavlja kovariančno funkcijo. Za poljubno obliko kovariančne funkcije je analitičen pristop nemogoč. Lahko pa pridobimo približek končne ("a-posteriori") verjetnostne porazdelitve s približevanjem povprečja napovedi preko vseh možnih vrednosti hiperparametrov z uporabo najbolj verjetnih vrednosti hiperparametrov (metoda največje podobnosti) ali z izvajanjem numerične integracije preko Θ z uporabo Monte Carlo metode.

4.2.6.1. Metoda največje podobnosti

Metoda določanja najbolj verjetnih vrednosti hiperparametrov uporablja za določitev vrednosti integrala (4.42) približek

$$P(t_{N+1} | \mathbf{x}_{N+1}, D, C(\cdot)) \simeq P(t_{N+1} | \mathbf{x}_{N+1}, D, C(\cdot), \boldsymbol{\Theta}_{MP}) \quad (4.43)$$

ki temelji na najbolj verjetni množici hiperparametrov $\boldsymbol{\Theta}_{MP}$.

Približek temelji na predpostavki, da ima končna verjetnostna porazdelitev $\boldsymbol{\Theta}$, $P(\boldsymbol{\Theta} | D, C(\cdot))$ obliko ostre konice okrog $\boldsymbol{\Theta}_{MP}$. V splošnem je opisani približek zelo dober in tako izdelane napovedi so pogosto zelo blizu pravim [37].

Najbolj verjetne vrednosti hiperparametrov pri določeni kovariančni funkciji določimo prek "a-posteriori" zapisa verjetnostne porazdelitve parametrov [36,47,26]

$$p(\boldsymbol{\Theta} | X, \mathbf{t}) = \frac{p(\mathbf{t} | X, \boldsymbol{\Theta}) p(\boldsymbol{\Theta})}{p(\mathbf{t} | X)} \quad (4.44)$$

Optimalne vrednosti parametrov določimo z iskanjem največje vrednosti logaritma porazdelitve, ki je logaritem Gaussovega procesa [26]:

$$\log(p(\boldsymbol{\Theta} | X, \mathbf{t})) = -\frac{1}{2} \log(|\mathbf{C}_N|) - \frac{1}{2} \mathbf{t}_N^T \mathbf{C}_N^{-1} \mathbf{t}_N - \frac{N}{2} \log(2\pi) \quad (4.45)$$

Parcialni odvod logaritma porazdelitve po hiperparametrih se glasi:

$$\frac{\partial L}{\partial \Theta_i} = -\frac{1}{2} \text{tr}(\mathbf{C}_N^{-1} \frac{\partial \mathbf{C}_N}{\partial \Theta_i}) + \frac{1}{2} \mathbf{t}_N^T \mathbf{C}_N^{-1} \frac{\partial \mathbf{C}_N}{\partial \Theta_i} \mathbf{C}_N^{-1} \mathbf{t}_N \quad (4.46)$$

V literaturi se ta metoda imenuje metoda največje podobnosti (ang. maximum likelihood - ML). Za iskanje minimuma se lahko uporablja katerakoli optimizacijska metoda. Ena izmed možnih je uporaba metode konjugiranih gradientov zaradi enostavnih analitičnih izračunov parcialnih odvodov (4.46). Predstavljena metoda je občutljiva na začetno izbiro hiperparametrov (padec v lokalni minimum), poleg tega pa je računsko zahtevna, saj vsak korak optimizacije potrebuje izračun inverzne kovariančne matrike dimenzije $n \times n$, kjer je n število podatkov v učni množici.

4.2.6.2. Metoda z uporabo Monte Carlo pristopa

Monte Carlo pristop uporablja metodo vzorčenja za izračun približka verjetnostne porazdelitve napovedi. Z aproksimacijo integrala v enačbi (4.42) z uporabo Markove verige (Markov chain), lahko zapišemo

$$P(t_{N+1} | \mathbf{x}_{N+1}, D, C(\cdot)) \simeq \frac{1}{T} \sum_{t=1}^T P(t_{N+1} | \mathbf{x}_{N+1}, D, C(\cdot), \boldsymbol{\Theta}_t) \quad , \quad (4.47)$$

kjer predstavljajo $\boldsymbol{\Theta}_t$ vzorce vzete iz končne ("a-posteriori") verjetnostne porazdelitve $P(\boldsymbol{\Theta} | D, C(\cdot))$ hiperparametrov $\boldsymbol{\Theta}$. Vsak izraz v vsoti zgornje enačbe ima Gaussovo obliko, zaradi česar je tudi Monte Carlo približek napovedane verjetnostne porazdelitve mešanice Gaussovih funkcij. Natančnost približka narašča z večanjem uporabljenega števila vzorcev.

Metoda, ki jo uporabljamo za vzorčenje funkcije končne verjetnostne porazdelitve hiperparametrov $\boldsymbol{\Theta}$, ima bistven vpliv na učinkovitost pristopa. Želimo si, da uporabljeni vzorci dobro predstavljajo verjetnostno porazdelitev iz katere so bili vzeti. Če pri procesu vzorčenja izpustimo določeno območje $\boldsymbol{\Theta}$ prostora, ki je močno povezano z verjetnostno porazdelitvijo, bo približek integrala enačbe (4.42) zelo slab.

4.2.6.3. Primerjava opisanih metod

V dodatku B je prikazan preprost primer modeliranja nelinearne funkcije z Gaussovimi procesi z uporabo metode največje podobnosti. Pri optimizaciji hiperparametrov včasih pride do primera, ko glede na različne začetne vrednosti hiperparametrov, dobimo različne končne optimirane vrednosti hiperparametrov. Glede na izkušnje avtorja iz [26], ta problem ne predstavlja bistvene ovire za uporabo omenjene metode pri modeliranju z Gaussovimi procesi. Omenjeni problem lahko rešimo z uporabo lastnih predpostavk o hiperparametrih. Pridobivanje najbolj verjetnih hiperparametrov tudi predstavlja najbolj privlačno odliko metode največje podobnosti. Pri interpretaciji modela z Gaussovimi procesi to pomeni, da nam takšna množica hiperparametrov lahko veliko pove o podatkih, ki jih modeliramo. V primeru samo ene končne množice hiperparametrov to pomeni, da je za izračunavanje napovedi potrebno inverzno vrednosti kovariančne matrike izračunati in shraniti samo enkrat.

Pri uporabi Monte Carlo pristopa se izračuna približek končne verjetnostne porazdelitve z uporabo izračuna povprečja iz množice vzorcev. Po tem, ko pridobimo vse informacije, ki jih potrebujemo za začetek Monte Carlo simulacije, lahko začnemo izračunavati povprečja korak za korakom. Pri tem moramo vsakič izračunati inverz kovariančne matrike, ni pa inverza potrebno shranjevati. Če pa se želimo vrniti na začetek in izračunati novo statistiko o podatkih ali dodati nekaj novih učnih podatkov, je potreben ponoven zahteven izračun vseh inverznih

kovariančnih matrik. Iz tega je razdvidno, da nam Monte Carlo pristop ne omogoča bistveno boljše prilagodljivosti kot metoda največje podobnosti. V Gaussovih procesih, ki uporabljajo Monte Carlo pristop, je možna tudi uporaba modelov šuma, ki nimajo Gaussove porazdelitve. Metoda uporabe Monte Carlo pristopa pa se izkaže tudi za bolj učinkovito v primeru uporabe kompleksnejših kovariančnih funkcij (konkretno v primeru uporabe modela šuma, ki je odvisen od vhodov).

Uporaba metode največje podobnosti v nadaljevanju pa na noben način ne pomeni, da je metoda boljša kot Monte Carlo pristop za vse probleme. Za manjše [26] množice podatkov, kjer shranjevanje matrik ne predstavlja bistvenega problema, se je izkazalo, da daje Monte Carlo pristop boljše rezultate pri nespremenljivi velikosti CPU časa. Za večje množice podatkov pa je metoda največje podobnosti hitrejša in učinkovitejša.

4.3. Mehka logika

Teorija mehke logike se je začela razvijati v šestdesetih letih, ko je Lofti A. Zadeh v članku [66] prvič obravnaval tako imenovane mehke ali zamegljene veličine, ki jih ni možno opisati s formalnim matematičnim zapisom, vključno s teorijo verjetnosti. Potrebo za njihovo uvedbo, kot tudi za razvoj ustreznih matematičnih postopkov, je utemeljil z dejstvom, da je človeško zaznavanje in opisovanje bioloških in mehanskih sistemov daleč od natančno definiranega matematičnega izrazoslovja in brez natančno poznanega zakona porazdelitve. Svoje delo je nadaljeval [65] z uvedbo pojma lingvistične spremenljivke, to je spremenljivke katere vrednost ni določena s številčnimi vrednostmi, pač pa z besedami iz naravnega jezika. Ta članek predstavlja delo, ki je bilo osnova nadaljnjemu teoretičnemu razvoju teorije mehke logike in njeni aplikaciji pri reševanju številnih problemov.

Danes se je mehka logika uveljavila na mnogih področjih kot so razpoznavanje vzorcev, procesiranje signalov, elektronika, zlasti pa se je uveljavila na področju vodenja in identifikacije sistemov. Metode mehke identifikacije sta prva predlagala Takagi in Sugeno [52], podrobnejša predstavitev mehkih indentifikacijskih metod pa je na voljo v literaturi [49], [51], [54], kjer se tudi natančno podane osnovne definicije in metode, potrebne za razumevanje in izgradnjo sistema mehke logike. Predstavljen je tudi izrek, ki uvršča sisteme mehke logike med univerzalne aproksimatorje in s tem postavlja teoretično osnovo za njihovo aplikacijo pri aproksimaciji neznanih funkcij.

4.3.1. Mehka identifikacija sistemov

Pri mehki identifikaciji sistemov lahko podobno kot pri klasičnih pristopih identifikacije problem razdelimo v dva glavna dela: določitev strukture in določitev parametrov sistema mehke logike.

Določitev strukture sistema mehke logike poteka po sledečem postopku:

- določitev vhodov,
- določitev izhodov,
- določitev tipa pravil (Mamdani, Takagi-Sugeno),
- določitev načina sklepanja in mehkih operatorjev,
- in določitev metode mehčanja vhodov in ostrenja izhodov.

Mehke identifikacijske modele lahko razvrstimo v dve skupini glede na to, kakšne tipe pravil uporabljajo.

Mamdani identifikacijski modeli so sestavljeni iz pravil oblike (4.48) kjer so vhodi in izhodi mehke vrednosti, torej je uporabljeno znanje v popolnosti lingvistično interpretirano. Adaptacijskih postopkov, ki bi vodili do dobrega modela, za takšne sisteme skoraj ni na voljo. Izjema predstavljajo genetski algoritmi, ki v aplikacijah v realnem času zaradi visoke računske zahtevnosti niso preveč uporabni.

$$R^m: \text{if } x_1 \wedge A_1^m \wedge x_2 \text{ je } A_2^m \wedge \dots \wedge x_n \text{ is } A_n^m \text{ then } y = Y^m \quad (4.48)$$

Takagi-Sugeno identifikacijski modeli imajo pravila oblike (4.49). Vsako pravilo predstavlja lineariziran lokalni model sistema, identifikacijski model pa izvaja interpolacijo med njimi. Izhod Takagi-Sugeno mehkega sistema pri uporabljeni težiščni metodi mehčanja je reduciran na uteženo povprečje.

$$R^m: \text{if } x_1 \text{ is } A_1^m \wedge x_2 \text{ is } A_2^m \wedge \dots \wedge x_n \text{ is } A_n^m \text{ then } y = f(x_1, x_2, \dots, x_n) \quad (4.49)$$

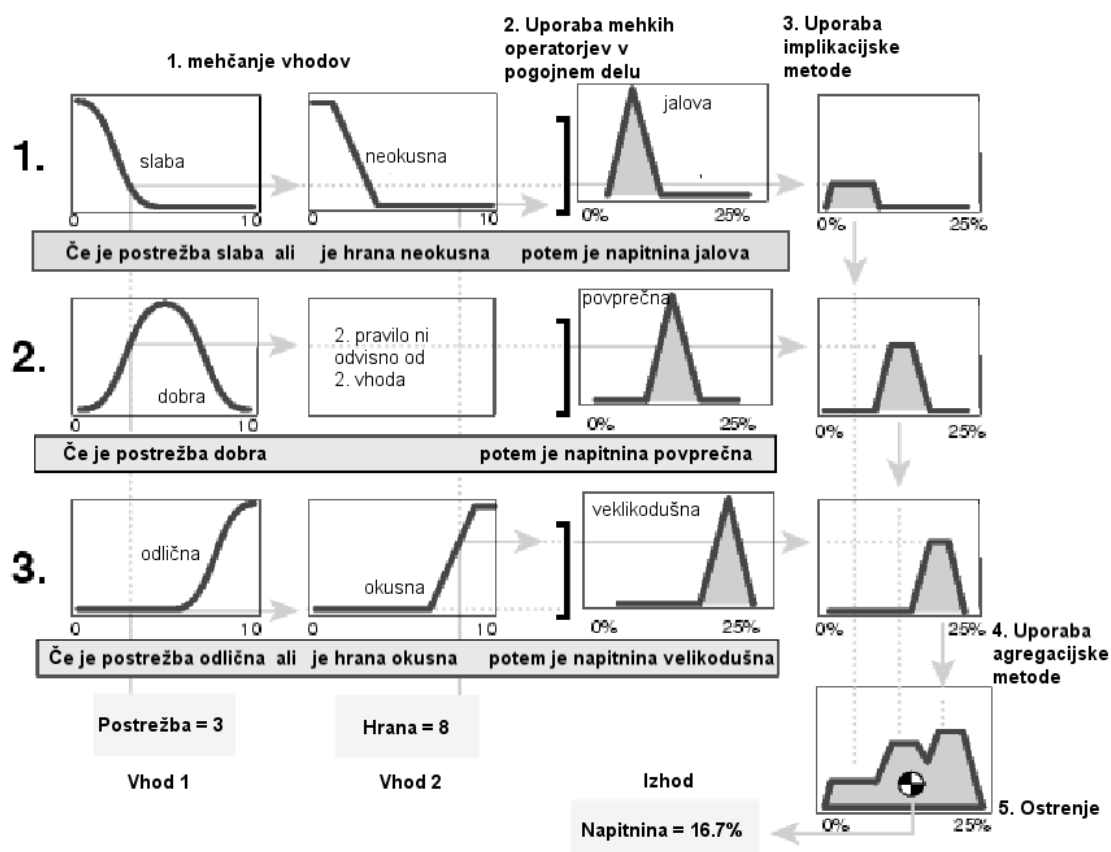
Določitev parametrov sistema mehke logike poteka po sledečem postopku:

- določitev števila, oblike in položaja mehkih množic
- in adaptacija vseh ali samo nekaterih parametrov.

Diagram poteka mehkega sklepanja v mehkih identifikacijskih modelih je prikazan s pomočjo preprostega praktičnega primera na sliki 17 [54]. Informacije v diagramu potekajo vzporedno od treh vhodov do enega izhoda. Vzporedna obravnava pravil predstavlja tudi eno izmed pomembnejših lastnosti mehkih sistemov. Proces mehkega sklepanja je

sestavljen iz petih postopkov.

1. *Mehčanje vhodov* je postopek pri katerem se ugotovi stopnja pripadnosti posameznega vhoda posamezni mehki množici s pomočjo pripadnostnih funkcij.
2. *Mehki operatorji v pogojnem delu (premisi) pravil* se uporabijo zatem, ko se zaključi mehčanje vhodov, v primeru, da je pogojni del (premisa) sestavljen iz večih delov.
3. *Uporaba implikacijske metode* sledi zatem, ko je s pogojnim delom določen tudi posledični del (konsekvenca). V praksi se uporabljajo različne vrste implikacijskih metod [54]: Mamdani implikacija, Lukasiewicz implikacija in produkt.
4. *Agregacija vseh izhodov* je proces pri katerem se mehke množice, ki jih določajo konsekvence pravil z uporabo implikacijske metode, združijo v združeno mehko množico. Vhod v agregacijski proces predstavlja skupina vseh funkcij iz posledičnih delov vseh pravil, ki so bile dodatno okrnjene z uporabo implikacijske metode. Izhod agregacijskega procesa pa je po ena mehka množica za vsak izhod mehkega modela.
5. *Ostrenje*, ki sledi agregaciji, je proces, ki priredi vhodni mehki množici realno število. Vhod v proces ostrenja je združena mehka množica, izhod pa realno število, ki tudi predstavlja izhodno vrednost mehkega modela.



Slika 17: Diagram poteka mehkega sklepanja v mehkih identifikacijskih modelih s pomočjo praktičnega primera “Dinner for two” iz programskega orodja Matlab [54]

4.3.2. Grupiranje mehkih množic

Mehko zbiranje podatkov v gruče predstavlja učinkovit pristop k identifikaciji kompleksnih nelinearnih sistemov s pomočjo mehke logike. Mnogi razvrščevalni postopki in algoritmi za modeliranje sistemov uporabljajo kot temeljeno metodo zbiranje numeričnih podatkov v gruče (angl. clustering). Glavni smisel zbiranja podatkov v gruče je določitev števila naravnih gruč v veliki množici podatkov, s katerimi je možno jedrnato predstaviti značilnosti sistema. Informacije o gručah omogočajo izgradnjo Takagi-Sugeno mehkega sistema sklepanja z uporabo relativno majhnega števila pravil. To v bistvu pomeni, da se za vsako gručo podatkov določi svoje pravilo. Za zbiranje podatkov v gruče obstaja veliko metod, ki so podrobneje opisane v literaturi [49], [51], [54].

4.3.3.ANFIS (Adaptive Neuro-Fuzzy Inference System)

Med sistemi mehke logike in umetnimi nevronske mrežami obstaja podobnost, ki je najočitnejša pri strukturi. Takagi-Sugeno sistem mehke množice je na primer možno prikazati tudi v obliki trinojske nevronske mreže, kar je podrobneje predstavljeno v [51]. Sistemi mehke logike so torej primerljivi po strukturi z umetnimi nevronske mrežami, možna pa je tudi uporaba enakih adaptacijskih algoritmov. Oba identifikacijska modela sta tudi univerzalna aproksimatorja za zvezne, nelinearne funkcije [51,29].

Za večnivojske nevronske mreže, ki imajo prav takšno strukturo, se običajno uporabljajo gradientno osnovani algoritmi za adaptacijo parametrov. Takagi-Sugeno mehke sisteme z gradientnim adaptacijskim postopkom imenujemo tudi ANFIS (adaptive neuro-fuzzy) sisteme.

4.4.Metode za vrednotenje modelov

Za primerjavo rezultatov modelov zgrajenih z različnimi metodami potrebujemo primerno kriterijsko funkcijo s katero se lahko oceni uspešnost modela na preizkusnih podatkih.

Kriterijska funkcija mora biti prilagojena področju uporabe. Oceniti mora predvsem tisto sposobnost modela, ki je poglobitni cilj razvoja [15], [40]. Pri modeliranju napovedovanja povišanih koncentracij ozona predstavlja poglobitni cilj zagotovitev čim boljšega napovedovanja visokih koncentracij. Nižje koncentracije za ocenjevanje niso zanimive, razen v primeru, če jih model napove s tako napako, da ta napoved povzroči tako imenovani lažni alarm.

Za oceno uspešnosti modeliranja se na področju modeliranja onesnaženja ozračja uporablja precej cenilk. Vsaka od njih pa osvetli določen vidik uspešnosti ali neuspešnosti modela [15], [40]. V nadaljevanju bom predstavil cenilke, ki sem jih uporabil za ocenjevanje zgrajenih modelov. Pred tem pa podajam še opis pomena posameznih oznak:

- Cm_i ... i-ta izmerjena vrednost koncentracije,
- Cr_i ... i-ta izračunana vrednost koncentracije,
- $\hat{C}$... povprečna vrednost koncentracije,
- σ_c ... standardna deviacija koncentracije,
- N ... skupno število preizkusnih vzorcev,
- f ... skupno število napovedanih prekoračitev,

- m ... skupno število izmerjenih prekoračitev,
 a ... število pravilno napovedanih prekoračitev.

4.4.1. Korelacijski koeficient

Korelacijski koeficient predstavlja normalizirano mero za ugotavljanje linearne odvisnosti dveh spremenljivk. Za povsem linearno neodvisni spremenljivki znaša vrednost koeficienta 0.0, medtem ko za linearno odvisni znaša 1.0 ali -1.0. Izračun vrednosti korelacijskega faktor med dvema spremenljivkama X in Y predstavlja enačba (4.50).

$$r(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}; \quad \text{cov}(X, Y) = E((X - E(X))(Y - E(Y))) \quad (4.50).$$

Izračun korelacijskega koeficienta za vrednotenje modelov glede na enačbo (4.50) opisuje enačba (4.51).

$$r(C_m, C_r) = \frac{\frac{1}{N} \cdot \sum_{i=1}^N (C_{m_i} - \hat{C}_m) \cdot (C_{r_i} - \hat{C}_r)}{\sigma_{C_m} \sigma_{C_r}} \quad (4.51).$$

4.4.2. Kvadratni koren povprečne kvadratne napake (RMSE)

Povprečno kvadratno napako (RMSE..Root Mean Square Error) med izračunanimi vrednostmi koncentracij iz modela in izmerjenimi vrednostmi določa enačba enačba (4.52).

$$RMSE = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^N (C_{r_i} - C_{m_i})^2} \quad (4.52).$$

4.4.3. Cenilka p^6

S stališča uporabe rezultatov napovedovanja v praksi je pomembno predvsem, da zna model dobro napovedati visoke koncentracije in da se pri nizkih koncentracijah ne zmoti toliko, da bi lahko sprožil lažni alarm. Ker nobena od doslej navedenih cenilk ne pokaže zadosti uspešnosti oziroma neuspešnosti napovedovanja prekoračenih koncentracij, sem za vrednotenje uporabil tudi cenilko p^6 [40], ki sem jo prilagodil velikostnemu razredu koncentracij ozona.

Cenilka p^6 pomeni verjetnost uspešne napovedi visokih koncentracij. Postopek izračuna je sledeč:

1. Določitev vzorcev, ki bodo upoštevani v izračunu. Za izračun se uporabijo vsi vzorci pri katerih je izmerjena visoka koncentracija in

vsi vzorci pri katerih je nastopil lažni alarm. Pogoja opisuje enačba (4.53). Vzorcev, ki ne ustrezajo danemu pogoju, se ne upošteva.

$$Cm_i > 120 \mu g/m^3 \vee Cr_i > 180 \mu g/m^3 \quad (4.53)$$

2. Določitev uspešnosti napovedi J_i^6 za vsak vzorec iz izbrane množice posebej. Posamezna napoved J_i^6 je uspešna (enaka 1), če je absolutna napaka napovedi manjša od $20 \mu g/m^3$ ali če je relativna napaka manjša od 0.2 in pri tem ni nastopil lažni alarm. Opisani pogoji so predstavljeni z enačbo (4.54).

$$J_i^6 = \begin{cases} 1, & \text{če velja: } (|Cm_i - Cr_i| < 20 \mu g/m^3) \vee ((Cm_i - Cr_i)/Cm_i < 0.2)) \wedge \\ & ((Cm_i > 150 \mu g/m^3) \vee (Cr_i > 180 \mu g/m^3)) \\ 0, & \text{v ostalih primerih} \end{cases} \quad (4.54)$$

3. Določitev cenilke verjetnosti uspešne napovedi visokih koncentracij p^6 tako, da preštejemo vse uspešne napovedi (napovedi z oceno 1) in vsoto delimo s številom vseh izbranih vzorcev, kot to opisuje enačba (4.55).

$$p^6 = \frac{1}{N} \cdot \sum_{i=1}^M J_i^6 \quad (4.55)$$

4.4.4. Parametri sposobnosti

Za vrednotenje in ocenjevanje sposobnosti modelov, ki napovedujejo prekoračene koncentracije ozona, je EEA (European Environment Agency – Evropska agencija za okolje) z dokumentom [61] določila standardno *tabelo deležev (contingency table)*, ki je prikazana v tabeli 1. Pomen oznak v tabeli je podan na začetku podpoglavja 4.4.

<i>Alarmi</i>	<i>Izmerjeni</i>		
<i>Napovedani</i>	<i>Da</i>	<i>Ne</i>	<i>Skupaj</i>
<i>Da</i>	<i>a</i>	<i>f-a</i>	<i>f</i>
<i>Ne</i>	<i>m-a</i>	<i>N+a-m-f</i>	<i>N-f</i>
<i>Skupaj</i>	<i>m</i>	<i>N-m</i>	<i>N</i>

Tabela 4.1: Tabela deležev (contingency table)

Z uporabo navedenih definicij so definirani *trije parametri sposobnosti*.

- *Merilo pravih napovedi SP* kritičnih dogodkov (verjetnost detekcije) določa enačba (4.56), ki predstavlja razmerje med številom pravilno napovedanih prekoračitev in skupnim številom izmerjenih

prekoračitev v odstotkih. Obseg merila je od 0% do 100%, kjer 100% označuje najboljšo vrednost.

$$SP = \frac{a}{m} \cdot 100\% \quad (4.56)$$

- *Merilo uresničenih napovedi SR* kritičnih dogodkov določa enačba (4.57), ki predstavlja razmerje med številom pravilno napovedanih prekoračitev in številom vseh napovedanih prekoračitev v odstotkih. Obseg merila je od 0% do 100%, kjer 100% označuje najboljšo vrednost.

$$SR = \frac{a}{f} \cdot 100\% \quad (4.57)$$

- *Indeks uspeha*, ki združuje parametra sposobnosti SP in SR , določa enačba (4.58) Obseg merila je od -100% do 100%, kjer 100% označuje najboljšo vrednost.

$$SI = \left(\frac{a}{m} + \frac{N + a - m - f}{N - m} - 1 \right) \cdot 100\% \quad (4.58)$$

Za vrednotenje modelov sem izdelal dva izračuna parametrov sposobnosti. V prvem primeru sem nastavljal kritično mejo koncentracije na $180 \mu\text{g}/\text{m}^3$, kot je to predpisano v uredbi. Parametre sposobnosti sem označil z SP_{180} , SR_{180} in SI_{180} . Ker pa je bilo število vzorcev v množici za preizkušanje, ki so presegli predpisano kritično vrednost, relativno majhno, sem se odločil v namene primerjave modelov napraviti še izračun pri kritični meji $140 \mu\text{g}/\text{m}^3$. V tem primeru sem parametre sposobnosti označil z SP_{140} , SR_{140} in SI_{140} .

5. Analiza izmerjenih podatkov

5.1. Geografski položaj Nove Gorice

5.1.1. Vipavska dolina in Goriško polje

Območje Vipavske doline in Goriškega polja leži na meji med submediteransko Slovenijo in dinarskim delom Slovenije [33,38,39], kot je prikazano na sliki 18. Vipavska dolina in Goriško polje ležita med nizko planoto Krasa in visokimi kraškimi planotami Nanosa in Trnovskega gozda. Pri Šempasu pride Vipavska dolina v Goriško ravan, zahodno od obeh Goric pa se razprostira Furlanska nižina. Relief na področju Vipavske doline in Goriškega polja prikazuje slika 19.



Slika 18: Izsek iz državne pregledne karte Slovenije prikazuje širšo okolico Vipavske doline in Goriškega polja

Območje Vipavske doline in Goriškega polja je nižje flišno ozemlje, katerega približno tri četrtine zavzema gričevje. Preostali del ozemlja, to je slaba četrtina, pa je ravnina. Te je več pod Trnovskim gozdom na prodnih vršajih potokov in vzdolž reke Vipave. Največja ravnina je na stiku gričevnate regije s Soško ravnino. Soča je v ledeni dobi s prodom nasula na tej strani državne meje Solkansko in Šempetrsko polje. Zaježila je Vipavo, da je nižje od Prvačine odlagala ilovico - surovino mnogih nekdanjih in sedaj redkih opekarn.

Med reko Vipavo in Krasom se Vipavsko gričevje dviga v hribovita Vipavska brda. Dno Vipavske doline ter gričevja na njenem jugozahodnem obrobju gradijo flišne plasti. Ta mehka kamnina, ki jo sestavljajo pole laporja, peščenjaka, apnene breče in glinastih skrilavcev, se je usedla na dnu terciarnega morja, ki je takrat še pokrivalo jugozahodni del Slovenije.

V Vipavski dolini in na Goriškem polju prevladuje submediteransko podnebje. Ker loči Kras Vipavsko dolino od obale, so njene temperature nižje kot ob obali, posebno ponoči in pozimi. Zaradi goratega zaledja je količina letnih padavin relativno velika. Poletna suša je omiljena. V zahodnem delu je pod vplivom sredozemskega podnebja. Pozimi je gornji del doline pogosto izpostavljen burji, ki piha povprečno 42 dni v letu od tega 30 dni precej močno.

Po Vipavski dolini teče reka Vipava ter po spodnjem robu Vipavske doline reka Branica. V ti dve reki se zlivajo številni manjši potoki. Večji potoki so na primer Hubelj, Skrivšek, Malenšček, Košivec, Konjščak. Za razliko od Krasa ima Vipavska dolina dve večji tekoči površinski reki. To ji omogočajo flišne plasti, ki so le delno prepustne za vodo.



Slika 19: Relief na področju Vipavske doline in Goriškega polja.

5.1.2. Nova Gorica

Sodobno urejeno mesto na robu Goriškega polja, tik ob slovensko-italijanski meji, je bilo ustanovljeno leta 1947, ko je z določitvijo meje med Italijo in Jugoslavijo Gorica ostala na italijanskem ozemlju, spodnja Soška dolina, Goriška brda in spodnja Vipavska dolina na jugoslovanski

strani meje pa so ostali brez naravnega središča [33,39,38]. Solkan, edino večje naselje spodnje Vipavske doline, je bilo tedaj kraj brez tradicionalnih središčnih funkcij. Nova Gorica je torej eno najmlajših slovenskih mest. Zaradi ugodne prometne lege na stičišču Vipavske in Soške doline, Krasa in italijanske Furlanske nižine se je hitro razvila v regionalno gospodarsko, kulturno, izobraževalno, upravno in prometno središče.

Mesto se razteza po prodnih nanosih Soče in glinastih usedlinah v srednjem delu porečja hudournika Koren in kaže, da se bo zraslo z naselji Kromberk, Pristava, Rožna dolina, Šempeter pri Gorici in Vrtojba. Nadmorska višina mesta znaša 92 m, število prebivalcev pa je večje od 35.000 [39]. Zemljevid mesta prikazuje slika 20.

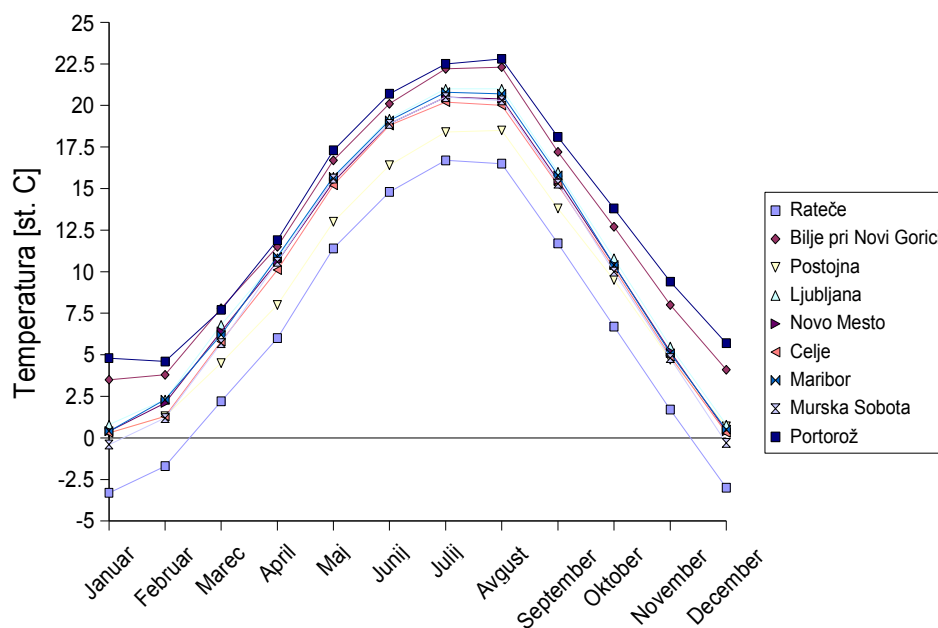


Slika 20: Zemljevid Nove Gorice

Skozi Novo Gorico teče promet s tujino. Železniška proga, tri ceste in sedem mejnih prehodov (pri Vrtojbi in v Rožni Dolini sta mednarodna) povezuje njeno širšo mestno območje s sosednjo Gorico v Italiji. Bližina meje in Gorice ji daje značaj obmejnega mesta.

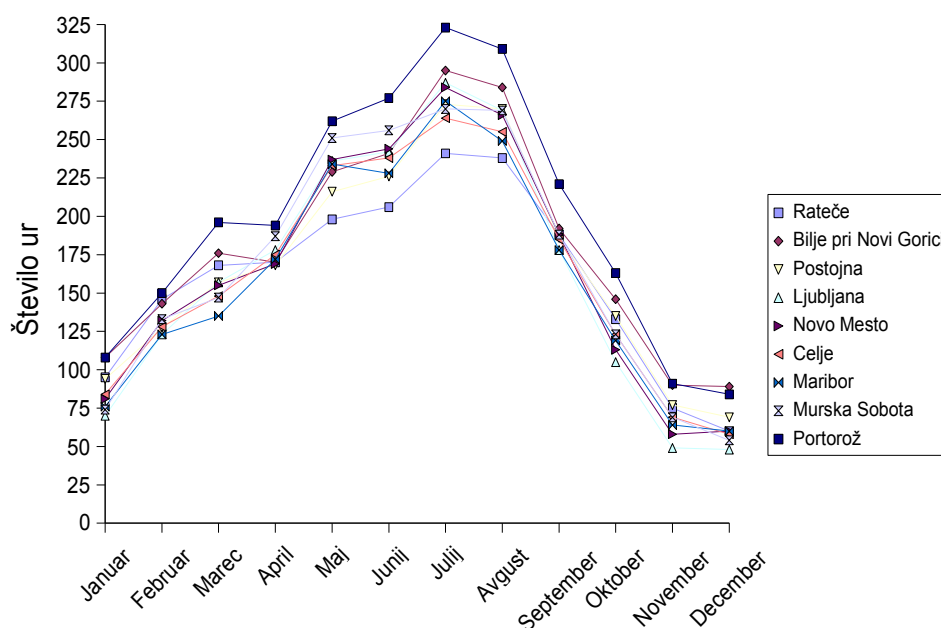
Podnebje je prehodno submediteransko z veliko sonca glede na večino ostalih krajev v Sloveniji, kot je to prikazano na grafu 22, ter povprečno milimi in vlažnimi zimami in vročimi poletji. Slednje nazorno prikazuje graf 21, kjer je prikazan letni potek mesečnih povprečnih temperatur

zraka v večjih krajih v Sloveniji za obdobje od leta 1991 do leta 2000. Podatki na omenjenih zadnjih dveh grafih ne vsebujejo podatkov izmerjenih v Novi Gorici, ker so bili na voljo samo podatki meteorološke merilne postaje Bilje v bližini Nove Gorice. Nova Gorica je pridobila avtomatsko merilno postajo za spremljanje onesnaženja ozračja in meteorologije šele v letu 2001 v okviru projekta Phare, kot je bilo že omenjeno.



Slika 21: Letni potek mesečnih povprečnih temperatur zraka v večjih krajih v Sloveniji za obdobje od leta 1991 do leta 2000

[3]



Slika 22: Letni potek mesečnega povprečnega trajanja sončnega obsevanja v urah v večjih krajih v Sloveniji za obdobje od leta 1991 do leta 2000 [3]

5.2.Meritve meteoroloških podatkov in koncentracij snovi v zunanjem zraku v Novi Gorici

Za identifikacijo modela so bili izbrani meteorološki in imisijski podatki, ki se merijo in shranjujejo na avtomatski okoljski postaji v Novi Gorici. Izmerjene polurne vrednosti se prenašajo iz postaje na centralno enoto Agencije RS za okolje (ARSO) ter zatem na mestno občino Nova Gorica (MO Nova Gorica). Okoljski sistem MO Nova Gorica ima poleg sistema shranjevanja polurnih podatkov vgrajen tudi mehanizem za izračun in shranjevanje vrednosti, ki se uporabljajo za vrednotenje stanja okolja glede na uredbo o ukrepih za ohranjanje in izboljšanje kakovosti zunanjega zraka [57,58,59,60]. V uredbi se za vrednotenje posameznih okoljskih veličin uporabljajo urne, dnevne, letne, osemurne drseče in štiriindvajseturne drseče vrednosti. Za ozon je predpisana opozorilna urna vrednost koncentracije $180 \mu\text{g}/\text{m}^3$ ter alarmna vrednost $240 \mu\text{g}/\text{m}^3$. Avtomatska merilna postaja meri in shranjuje meteorološke veličine:

- temperaturo zraka,
- relativno vlažnost,
- moč globalnega sončnega sevanja na enoto površine,
- vrednost zračnega pritiska,

- hitrost in smer vetra.

Poleg meteoroloških veličin se merijo in shranjujejo tudi podatki o koncentracija snovi v zunanjem zraku:

- koncentracija delcev $PM(10)^3$,
- koncentracija žveplovega dioksida (SO_2),
- koncentracija dušikovega monoksida (NO),
- koncentracija dušikovega dioksida (NO_2),
- koncentracija dušikovega oksida (NO_x),
- koncentracija ogljikovega monoksida (CO)
- in koncentracija ozona (O_3).

Dodatno pa je v postaji vgrajen tudi analizator hlapnih organskih spojin (VOC), ki meri sledeče parametre:

- koncentracijo benzena,
- koncentracijo toluena,
- koncentracija (orto)ksilena
- in vsoto koncentracij hlapnih organskih spojin.

Zaradi tehničnih težav z merilnikom podatki o izmerjenih vrednostih hlapnih organskih spojin (VOC) niso dostopni in jih na žalost tudi ni bilo mogoče uporabiti pri modeliranju. Tudi izmerjenih vrednosti koncentracije prahu v ozračju ni bilo mogoče uporabiti zaradi neenotne statistične obdelave izmerjenih vrednosti, ki se je v izbranem časovnem intervalu spreminjala.

Za modeliranje sem uporabil izmerjene urne povprečne vrednosti vseh omenjenih parametrov, ki so bili na voljo od dne 01.01.2002 do dne 31.12.2004. Vse podane urne vrednosti sem moral še dodatno urediti tako, da sem izpustil tiste urne intervale za katere niso bile na voljo urne vrednosti vseh izbranih parametrov. Po končanem urejanju mi je ostalo na voljo od 26304 možnih urnih vrednosti 14520 uporabnih urnih vrednosti (55.2%). V nekaterih urnih intervalih so manjkale izmerjene vrednosti določenih parametrov zaradi različnih vzrokov kot so kalibracije merilnika vsakih 23 ur (ker se običajno merilniki ne

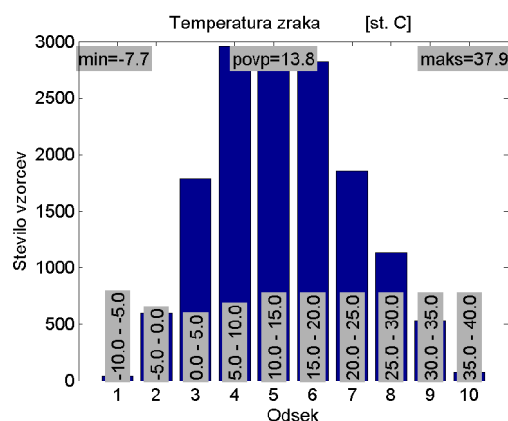
3 Delci $PM(10)$ so delci v zraku, ki jih prepušča filter s 50% nepropustnostjo za delce z aerodinamskim premerom $10\ \mu m$.

kalibrirajo hkrati, je izpad večkratno) ter popravila senzorjev in merilnikov, ki so včasih trajala tudi po več dni. Vse uporabne urne povprečne vrednosti meteoroloških in imisijskih parametrov sem tudi dodatno analiziral. Pri analizi sem za vsak parameter izrisal histogram izmerjenih vrednosti in določil minimalno vrednost, maksimalno vrednost in povprečno izmerjeno vrednost.

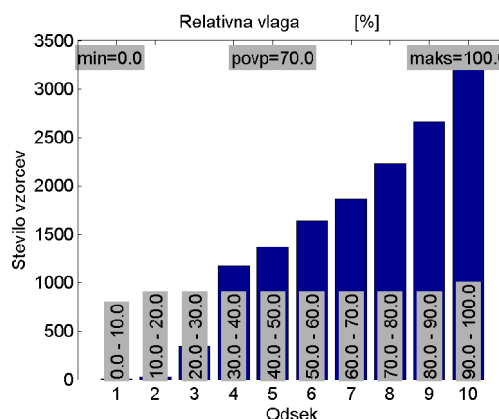
Na slikah od 23 do 26 so prikazani histogrami meteoroloških parametrov, ki se ujemajo s klimatološkimi značilnostmi Nove Gorice. Za statistično predstavitev vetra je uporabljena roža vetrov, ki v obliki frekvenčnega diagrama prikazuje odstotek vetra iz posamezne smeri. Roža vetrov je prikazana na sliki 27, statistika, iz katere je bila roža vetrov sestavljena, pa je prikazana na sliki 28.

Na slikah od 29 do 34 pa so prikazani histogrami koncentracij snovi v zunanjem zraku. Iz histogramov je razvidno, da so bile glede na veljavne uredbe [57,60,58,2] izmerjene relativno nizke koncentracije žveplovega dioksida (mejna urna vrednost po uredbi znaša $350 \mu\text{g}/\text{m}^3$) in ogljikovega monoksida (mejna 8-urna vrednost po uredbi znaša $10 \text{ mg}/\text{m}^3$). Medtem pa so izmerjene vrednosti dušikovega monoksida, duškovih oksidov in dušikovega dioksida relativno visoke, čeprav so še pod dovoljenimi vrednostmi (mejna urna vrednost po uredbi znaša za dušikov dioksid $200 \mu\text{g}/\text{m}^3$).

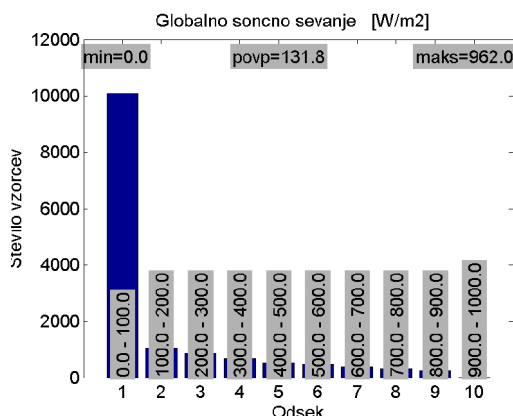
Pri analizi podatkov sem ugotovil, da so bile opozorilne urne koncentracije ozona kar 82 krat presežene. Pri tem je potrebno poudariti, da so visoke koncentracije nad opozorilno mejo nastopile v enem dnevu kar nekajkrat, kar pomeni, da je bilo v tem času glede na podane podatke kritičnih 19 dni.



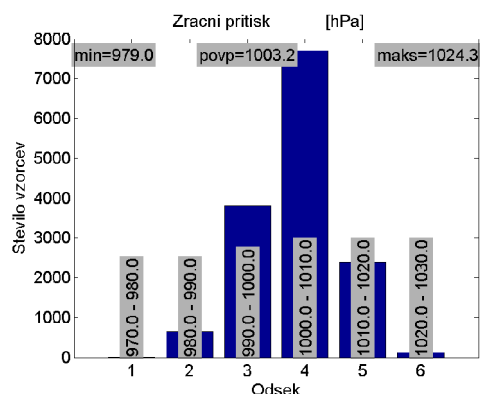
Slika 23: Histogram temperature zraka



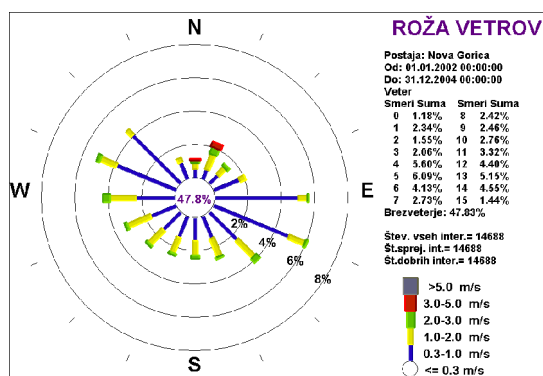
Slika 24: Histogram relativne vlage



Slika 25: Histogram moči globalnega sončnega sevanja na enoto površine



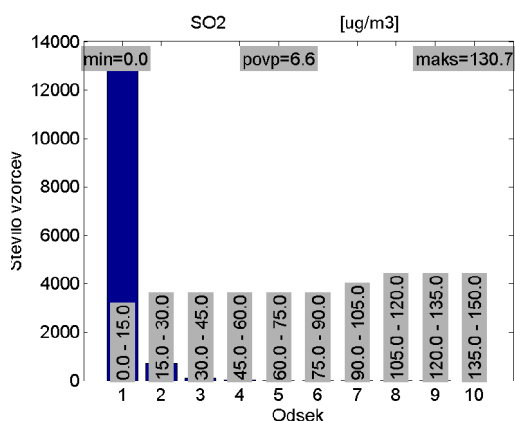
Slika 26: Histogram zračnega pritiska



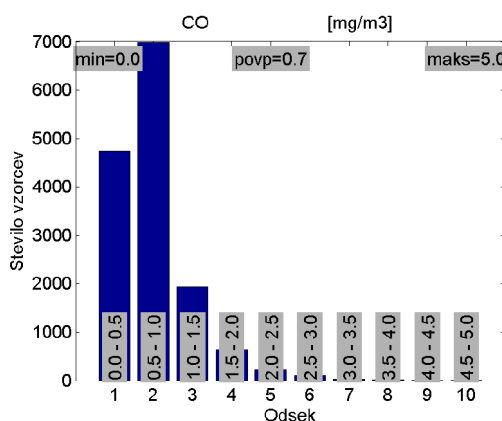
Slika 27: Roža vetrov

ROŽA VETROV							
Podatki od 01.01.2002 00:00:00 do vključno 31.12.2004 00:00:00 Postaja: Nova Gorica							
SMERI	< 0.3	0.3-1.0	1.0-2.0	2.0-3.0	3.0-5.0	>5.0	SKUPAJ
-11.3 - 11.3		0.52	0.33	0.18	0.10	0.04	1.17
11.3 - 33.8		0.61	0.96	0.44	0.31	0.01	2.33
33.8 - 56.3		0.71	0.57	0.21	0.06	0.00	1.55
56.3 - 78.8		1.63	0.38	0.04	0.01	0.00	2.06
78.8 - 101.3		4.92	0.54	0.11	0.03	0.00	5.60
101.3 - 123.8		4.83	1.02	0.14	0.10	0.00	6.09
123.8 - 146.3		2.30	1.42	0.34	0.07	0.00	4.13
146.3 - 168.8		1.36	1.19	0.13	0.05	0.00	2.73
168.8 - 191.3		1.36	0.84	0.22	0.00	0.00	2.42
191.3 - 213.8		1.32	0.99	0.14	0.01	0.00	2.46
213.8 - 236.3		1.46	1.14	0.16	0.01	0.00	2.76
236.3 - 258.8		1.64	1.48	0.17	0.02	0.00	3.33
258.8 - 281.3		2.31	1.57	0.45	0.07	0.01	4.41
281.3 - 303.8		3.86	1.04	0.22	0.03	0.00	5.15
303.8 - 326.3		4.06	0.44	0.05	0.00	0.00	4.55
326.3 - 348.8		1.07	0.28	0.07	0.02	0.00	1.44
SKUPAJ	47.83	33.96	14.20	3.08	0.88	0.06	
Štev. vseh int. = 14688 Št. dobrih int. = 14688 Št. sprej. int. = 14688							

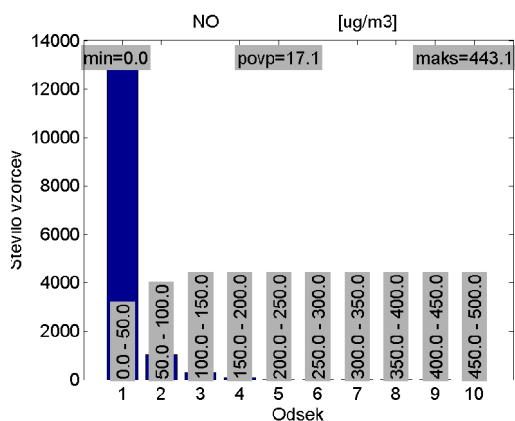
Slika 28: Statistika rože vetrov



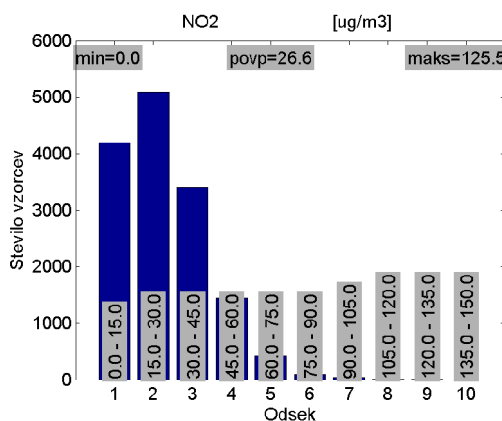
Slika 29: Histogram koncentracij žveplovega dioksida



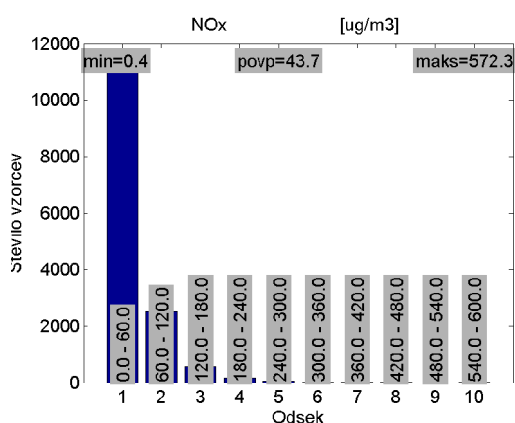
Slika 30: Histogram koncentracij ogljikovega monoksida



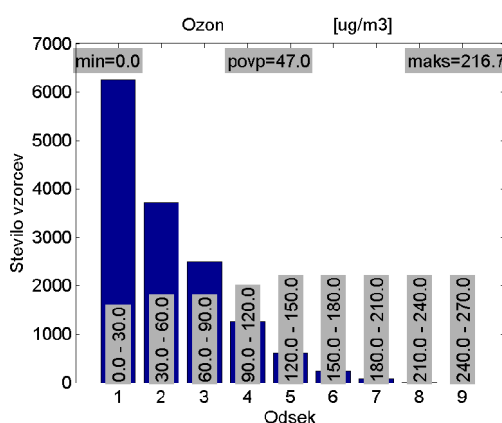
Slika 31: Histogram koncentracij dušikovega monoksida



Slika 32: Histogram koncentracij dušikovega dioksida



Slika 33: Histogram koncentracij dušikovih oksidov



Slika 34: Histogram koncentracij ozona

5.3.Izbira vhodnih spremenljivk modela

Za izgradnjo modela za napovedovanje maksimalne urne koncentracije ozona naslednjega dne glede na zbrane meterološke podatke in podatke o koncentracijah snovi v zunanjem zraku do 19 ure v tekočem dnevu sem moral najprej določiti vhodne spremenljivke v model in izhod.

Za izhod modela sem določil maksimalno urno koncentracijo ozona, ki je bila izmerjena v prihajajočem dnevu od 19:00 ure v tekočem dnevu do 19:00 ure v prihajajočem dnevu.

Glede na fizikalno kemijsko naravo procesa, priporočila in rezultate iz literature, izračunane vrednosti korelacijskih koeficientov med posameznimi vhodnimi veličinami in izhodno veličino (slika 44), ter na veliko število poizkusov so bili izbrane vhodne spremenljivke modela:

1. temperatura zraka (24 urna povprečna vrednost),
2. moč globalnega sončnega sevanja na enoto površine (24 urna povprečna vrednost),
3. koncentracija dušikovega monoksida (24 urna povprečna vrednost),
4. koncentracija dušikovega dioksida (24 urna povprečna vrednost),
5. koncentracija ozona (24 urna povprečna vrednost).

24 urna povprečna vrednost pomeni izračunano povprečje iz urnih povprečnih izmerjenih vrednosti od 19:00 ure prejšnjega dne do 19:00 ure današnjega dne.

Med procesom modeliranja se je izkazalo, da je možno model izboljšati z uvedbo dodatnih vhodnih spremenljivk, ki vsebujejo napovedane vrednosti parametra za dan vnaprej. Za potrebe modeliranja so bile pridobljene iz izmerjenih vrednosti prihajajočega dneva. V praktični uporabi modela pa je predvideno pridobivanje teh vrednosti iz modela ALADIN Agencije RS za okolje. Vrednosti modela so objavljene na internetnih straneh ARSO [5]. Dodatne vhodne spremenljivke so:

6. napovedana maksimalna urna vrednost temperature zraka v prihajajočem dnevu,
7. napovedana komponenta hitrosti vetra v smeri sever-jug, pozitivna vrednost pomeni komponento vetra iz severne smeri, negativna pa komponento vetra iz južne smeri (povprečna vrednost izračunana iz vektorske hitrosti vetra za čas od 19:00 ure današnjega dne do 19:00 ure naslednjega dne),

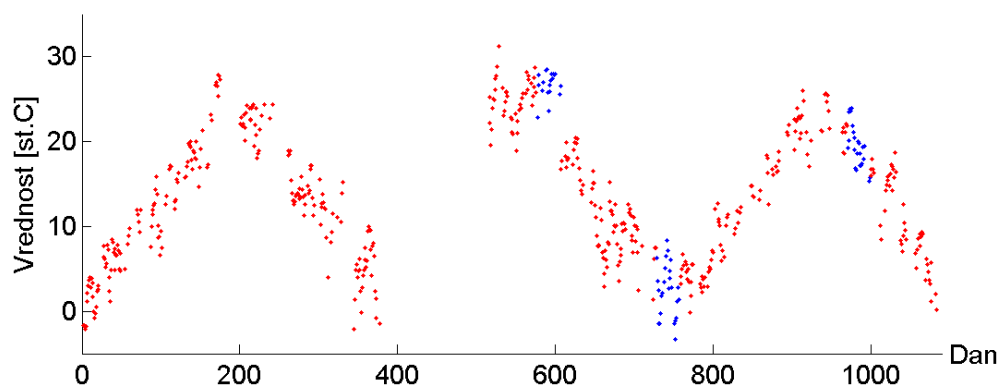
8. napovedana komponenta hitrosti vetra v smeri vzhod-zahod, pozitivna vrednost pomeni komponento vetra iz vzhodne smeri, negativna pa komponento vetra iz zahodne smeri (povprečna vrednost izračunana iz vektorske hitrosti vetra za čas od 19:00 ure današnjega dne do 19:00 ure naslednjega dne).

Vrednosti korelacijskih koeficientov med posameznimi vhodi in izhodom so prikazane na sliki 44 in so predstavljene v enakem vrstnem redu kot na predhodnem spisku izbranih vhodov. Pomen korelacijskega koeficienta in postopek izračuna je predstavljen v podpoglavju 6.1.1.

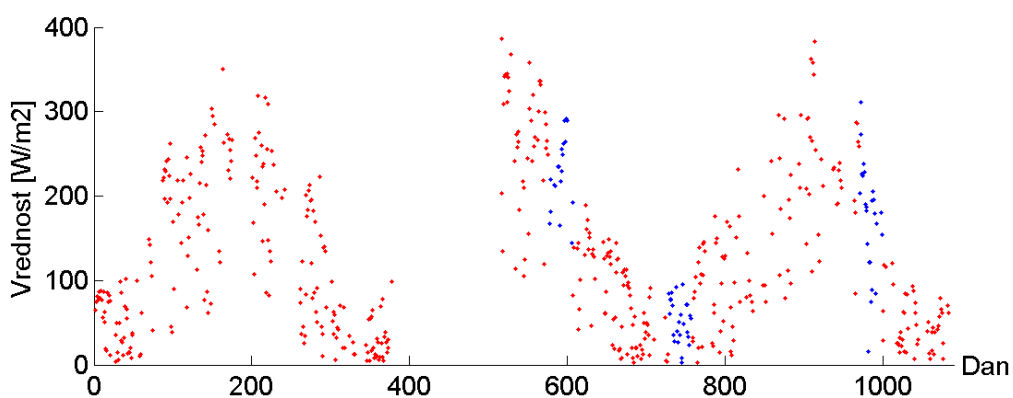
Množico vseh vzorcev sem za potrebe modeliranja razdelil v dve množici.

- Množico za učenje modelov, ki vsebuje vzorce iz štirih časovnih intervalov: 01.01.2002-31.07.2003, 01.09.2003-31.12.2003, 01.02.2004-31.08.2004, 01.10.2004-31.12.2004. Po izločevanju neuporabnih vzorcev zaradi različnih napak, ki so bile posledice različnih vzrokov (kalibracije merilnikov, okvare merilnikov in senzorjev, neuspešne komunikacije), je bilo možno za identifikacijo uporabiti 488 dobrih vzorcev izmed 1004 največ možnih vzorcev.
- Množico za vrednotenje modelov, ki vsebuje vzorce iz treh časovnih intervalov: 01.08.2003-31.08.2003, 01.01.2004-31.01.2004, 01.9.2004-30.09.2004. Po izločevanju neuporabnih vzorcev zaradi različnih napak, je bilo možno za preizkušanje uporabiti 68 dobrih vzorcev izmed 92 največ možnih vzorcev.

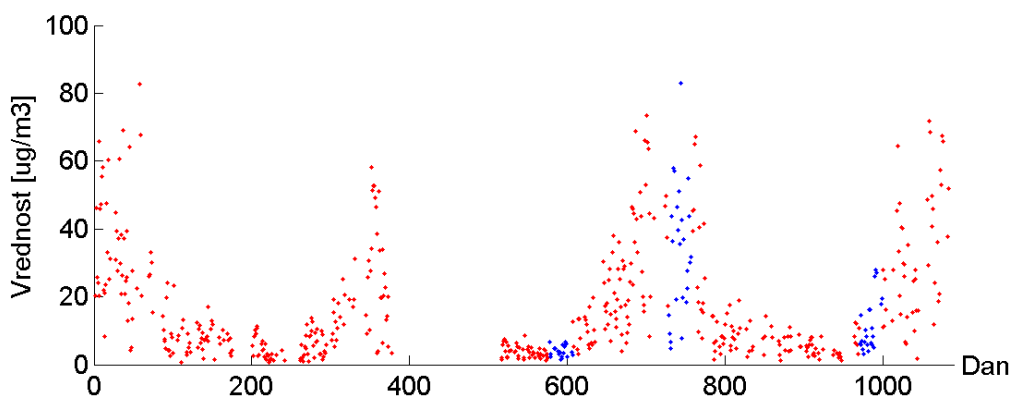
Potek posameznih vhodnih spremenljivk je prikazan na slikah od 35 do 42. Na sliki 43 pa je prikazan potek izhodne spremenljivke. Rdeča (svetlejša) krivulja označuje množico za učenje, modra (temnejša) krivulja pa množico za preizkušanje modelov. Horizontalna os grafov *Dan* označuje številko dneva šteto od dne 01.01.2002. To pomeni, da je na primer z indeksom 0 označen dan 01.01.2002, z indeksom 1096 pa 31.12.2004.



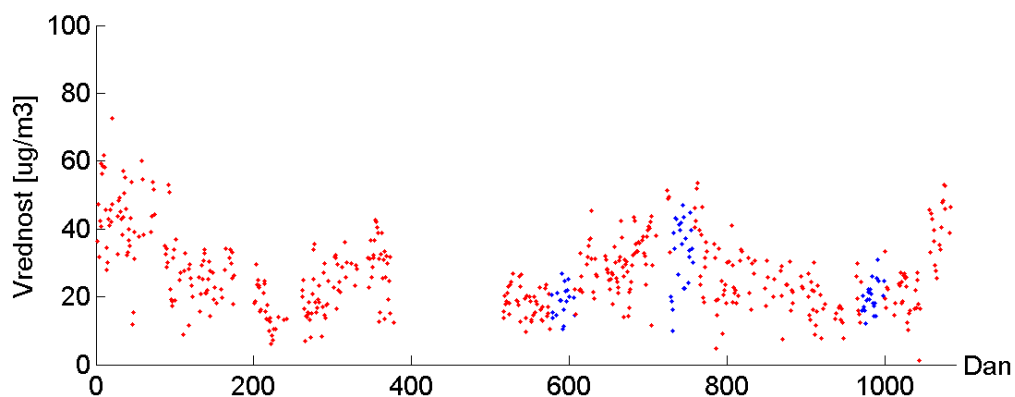
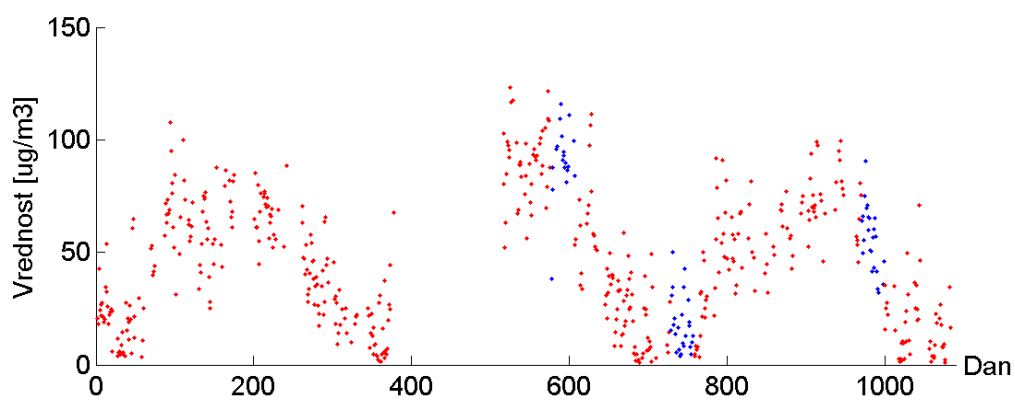
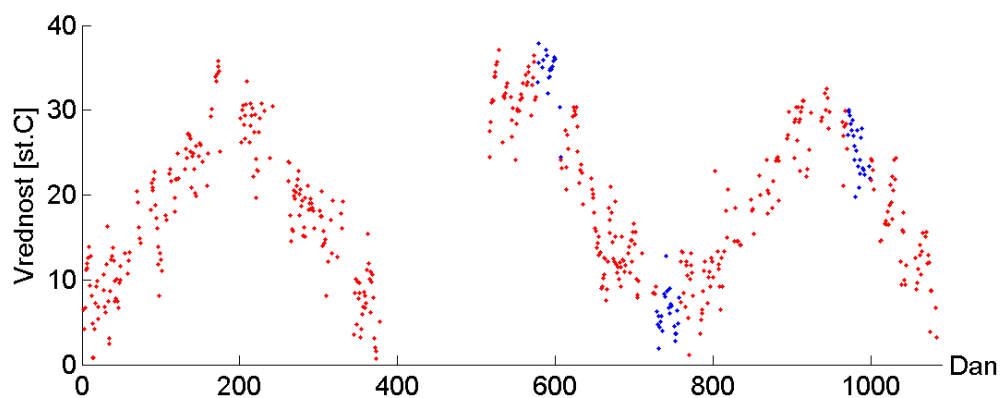
Slika 35: Vhod 1, Temperatura zraka



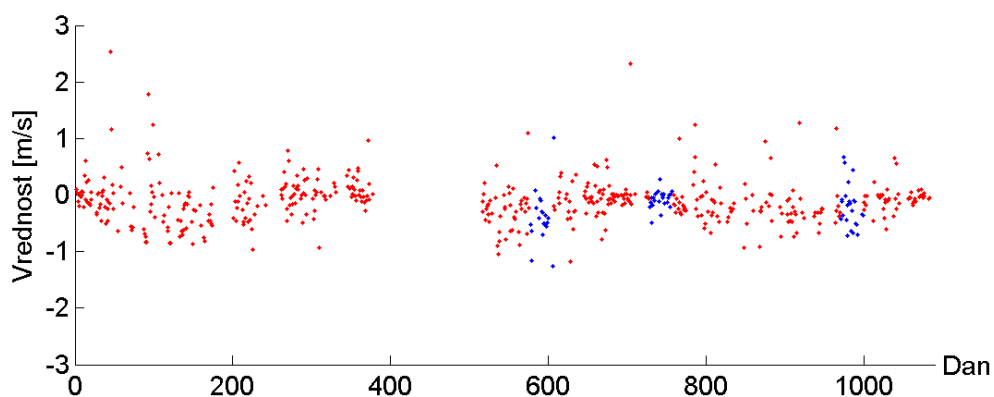
Slika 36: Vhod 2, Moč globalnega sončnega sevanja na enoto površine



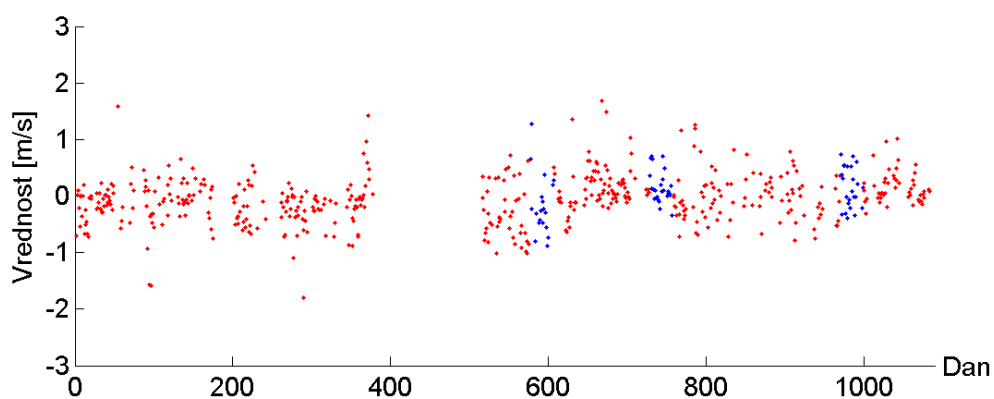
Slika 37: Vhod 3, Koncentracija dušikovega monoksida (NO)

Slika 38: Vhod 4, Koncentracija dušikovega dioksida (NO_2)Slika 39: Vhod 5, Koncentracija ozona (O_3)

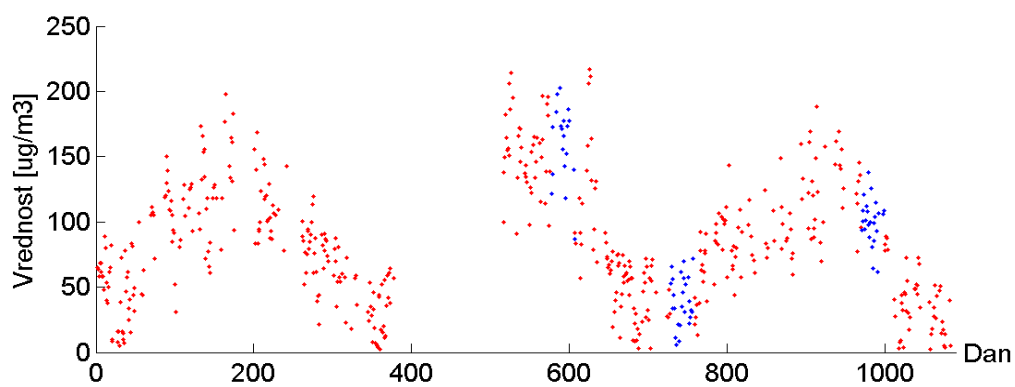
Slika 40: Vhod 6, Napovedana maksimalna temperatura zraka



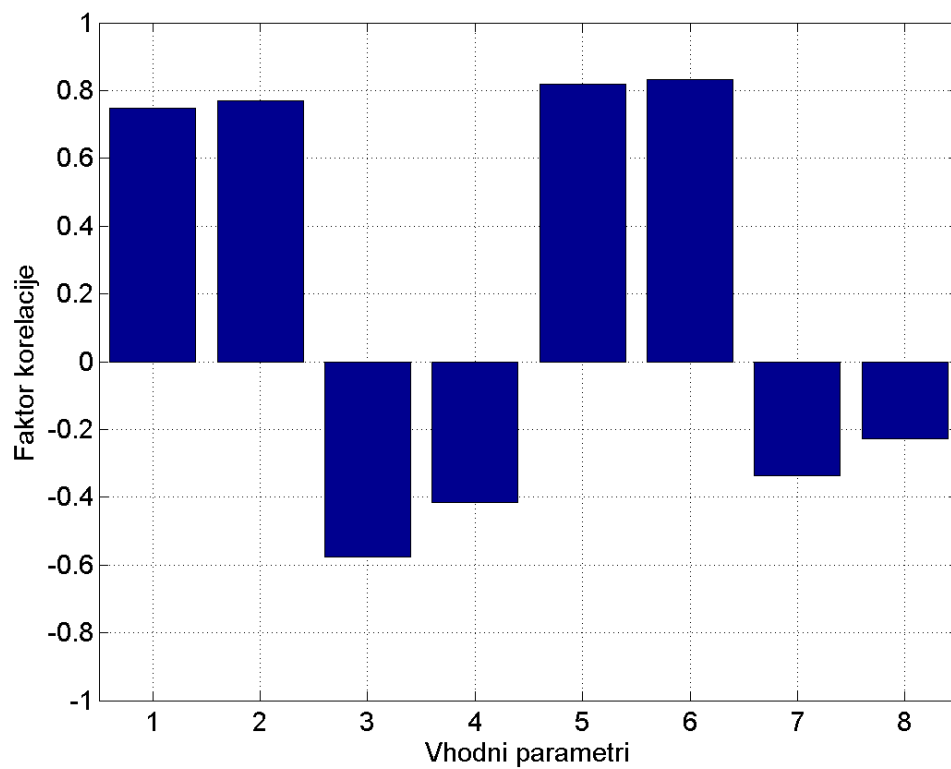
Slika 41: Vhod 7, Napovedan veter v smeri sever-jug



Slika 42: Vhod 8, Napoved vetra v smeri vzhod-zahod



Slika 43: Izhod, Maksimalne dnevne koncentracije ozona



Slika 44: Vrednosti korelacijskih koeficientov med posameznimi vhodnimi spremenljivkami v model in izhodno spremenljivko modela

6. Sinteza modelov za napovedovanje povišanih koncentracij ozona

Za napovedovanje povišanih koncentracij ozona sem glede na izbrane vhodne in izhodne spremenljivke zgradil tri različne modele z uporabo treh različnih metod modeliranja opisanih predhodno: z umetno nevronske mrežo, Gaussovimi procesi in mehko logiko. Za izgradnjo vseh treh modelov je bila uporabljena ista učna množica, ki je bila predstavljena v prejšnjem poglavju.

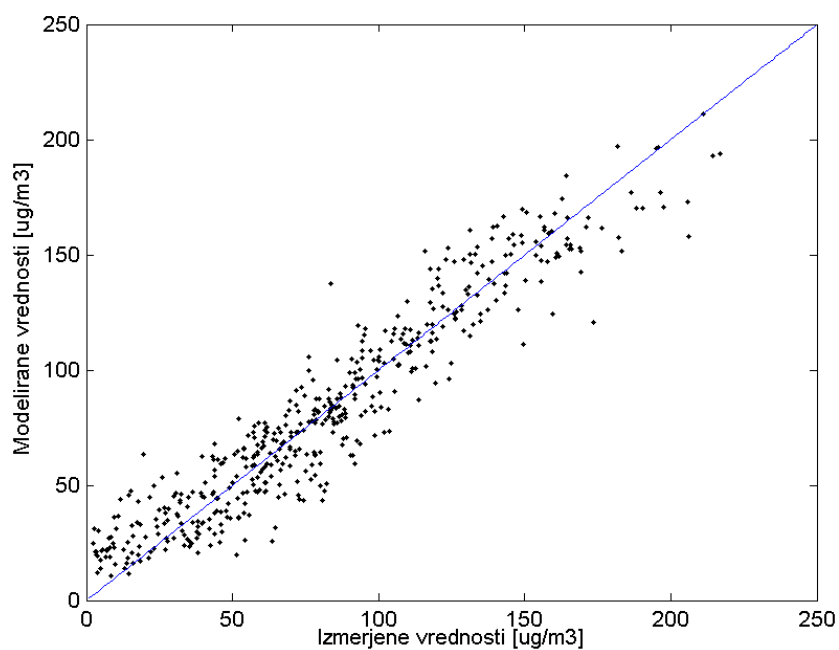
6.1. Sinteza modela z umetno nevronske mrežo

Za model je bila izbrana večnivojska perceptronska (multi-layer feed-forward) umetna nevronska mreža, ki je bila zgrajena s pomočjo orodja za uporabo umetnih nevronskih mrež v programskem paketu Matlab poimenovanega *Neural Network Toolbox* [54] (funkcija *newff()*). Model vsebuje dva nivoja, nevroni pa uporabljajo v obeh primerih tangens-sigmoidno prenosno funkcijo. Pri procesu modeliranja sem po velikem številu poizkusov določil, da je najbolj primerno število skritih nevronov 16. Začetne vrednosti uteži in odmikov so bile nastavljene z uporabo funkcije *initnw()*. Učenje nevronske mreže je potekalo s pomočjo dveh množic podatkov: množice za učenje, ki je vsebovala 437 vzorcev, ter množice za sprotno preverjanje podatkov, ki je vsebovala 51 vzorcev. Za učenje je bila izbrana metoda Levenberg-Marquardt. Vsi izmerjeni podatki v model so bili predhodno normalizirani.

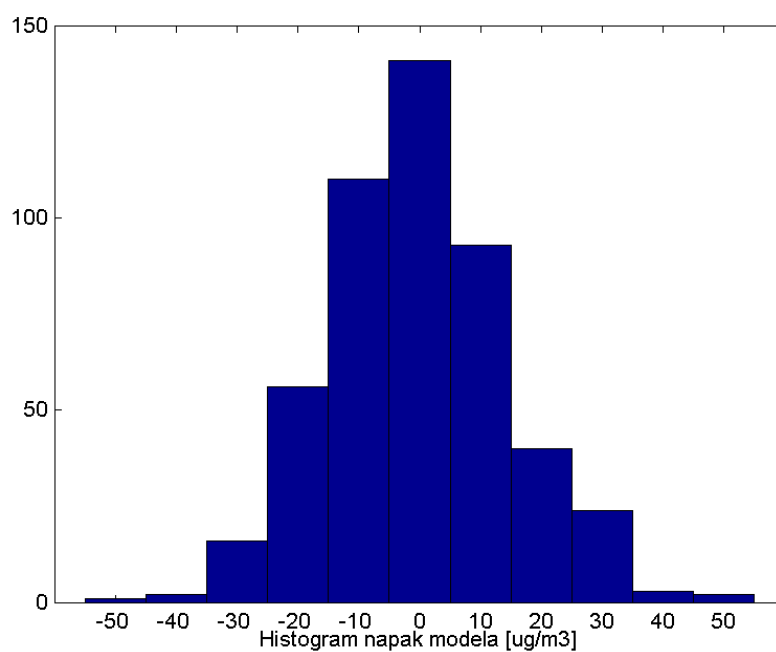
Po končanem učenju je bil najprej preizkušen model na podatkih iz učne množice. Odvisnost med izmerjenimi in izračunanimi vrednostmi je prikazana na sliki 45. Na sliki 46 pa je prikazan še histogram napak, ki kaže značilnosti normalne porazdelitve. Napake so izračunane kot razlika med izmerjeno in modelirano vrednostjo. Iz podatkov, ki so bili uporabljeni za indentifikacijo modela, sem izračunal tudi vrednosti cenilk, ki sem jih uporabljal za vrednotenje modela (poglavje 4.4):

$r=0.949$, $RMSE=14.8$, $p^6=0.93$, $SP_{180}=40.0$, $SR_{180}=85.7$, $SI_{180}=39.8$, $SP_{140}=86.1$, $SR_{140}=76.7$, $SI_{140}=82.1$.

Poudarim pa naj, da prikazane vrednosti ne odražajo prave kvalitete modela (v smislu zmožnosti posploševanja), ampak so samo informativne narave in prikazujejo kako dobro se je model naučil učne podatke.



Slika 45: Rezultati preizkusa modela z umetno nevronske mrežo s podatki iz učne množice



Slika 46: Histogram napak preizkusa modela z umetno nevronske mrežo s podatki iz učne množice

6.2.Sinteza modela z Gaussovimi procesi

Model z Gaussovimi procesi je bil zgrajen s pomočjo programskega orodja v okolju Matlab, ki je na voljo na domači spletni strani dr. C. E. Rasmussen-a [48]. Programsko orodje predstavljata dve Matlab funkciji.

- *gp01lik()* funkcija, ki omogoča optimizacijo hiperparametrov modela z Gaussovimi procesi, ki uporablja stacionarno kovariančno funkcijo. Optimizacija poteka po metodi največje podobnosti. Vhod v funkcijo predstavljajo začetne vrednosti hiperparametrov in množica učnih točk, ki je sestavljena iz parov $\{\text{izmerjeni vhodi}, \text{izmerjen izhod}\}$. Izhod funkcije pa so optimirane vrednosti hiperparametrov.
- *gp01pred()* funkcija, ki omogoča na podlagi podanih hiperparametrov in učnih točk napoved izhoda modela pri podanem vhodu. Vhod v funkcijo predstavljajo optimalne vrednosti hiperparametrov in množica učnih točk, ki je sestavljena iz parov $\{\text{izmerjeni vhodi}, \text{izmerjen izhod}\}$, ter množica vhodov, za katere želimo napovedati izhodne vrednosti. Izhod pa predstavlja množica napovedanih izhodnih vrednosti, ki jo sestavljajo pari $\{\text{napovedana srednja vrednost izhoda}, \text{napovedana dvojna vrednost standardne deviacije izhoda}\}$.

Glede na število vhodov, ki je znašalo v konkretnem primeru 8, sem določil 8 hiperparametrov GP modela (hiperparametri $\Theta_1 \dots \Theta_8$ predstavljajo dolžinska merila vsake dimenzije (vhoda), Θ_9 določa skupen vpliv vseh učnih točk na napovedane vrednosti, Θ_{10} določa minimalno negotovost napovedanih točk). Začetne vrednosti vseh hiperparametrov sem pred optimizacijo nastavljal na konstantno vrednost 0.5. Po optimizaciji so bile vrednosti hiperparametrov sledeče:

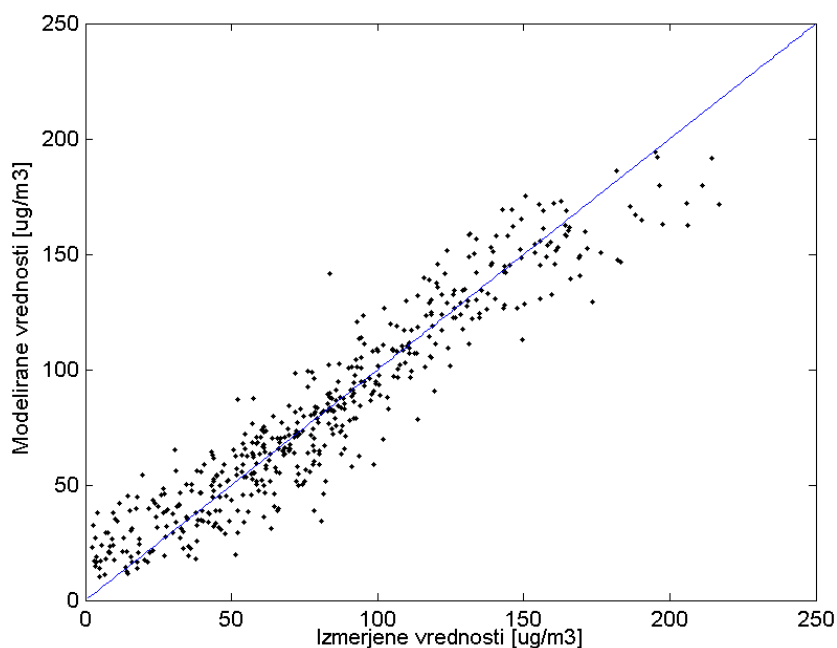
Θ_1	Θ_2	Θ_3	Θ_4	Θ_5	Θ_6	Θ_7	Θ_8	Θ_9	Θ_{10}
0,270	0,061	0,019	0,150	0,342	0,885	2,941	1,772	0,256	0,022

Tabela 6.1: Vrednosti hiperparametrov

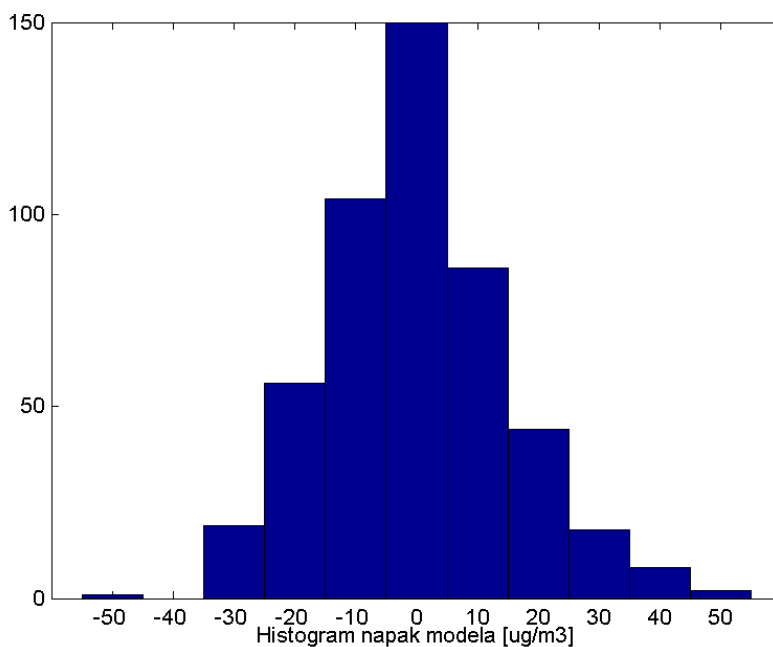
Po končanem učenju je bil najprej preizkušen model na podatkih iz učne množice. Odvisnost med izmerjenimi in izračunanimi vrednosti je prikazana na sliki 47. Na sliki 48 pa je prikazan še histogram napak, ki kaže značilnosti normalne porazdelitve. Napake so bile izračunane kot razlika med izmerjeno in modelirano vrednostjo. Iz podatkov, ki so bili uporabljeni za indentifikacijo modela sem izračunal tudi vrednosti cenilk, ki sem jih uporabljal za vrednotenje modela (poglavje 4.4):

$r=0.947$, $RMSE=15.1$, $p^6=0.93$, $SP_{180}=26.7$, $SR_{180}=100.0$, $SI_{180}=26.7$, $SP_{140}=84.6$, $SR_{140}=79.7$, $SI_{140}=81.3$.

Zopet poudarjam, da prikazane vrednosti ne odražajo prave kvalitete modela, ampak so samo informativne narave.



Slika 47: Rezultati preizkusa modela z Gaussovimi procesi s podatki iz učne množice



Slika 48: Histogram napak preizkusa modela z Gaussovimi procesi s podatki iz učne množice

6.3.Sinteza modela z mehko logiko

Mehki model je bil zgrajen s pomočjo orodja FMID (Fuzzy Modeling and Identification Toolbox) v programskem okolju Matlab, katerega avtor je prof. dr. Robert Babuška [10]. Orodje sestavlja skupina Matlab funkcij za izgradjo Takagi-Sugeno (TS) mehkih modelov iz množice izmerjenih podatkov. Model temelji na grupiranju izmerjenih podatkov z mehko kovariančno matriko s pomočjo Gustafson-Kessel algoritma [51, 54]. Algoritem predstavlja razširitev običajnega algoritma mehke metode c-mehkih srednjih vrednosti (fuzzy c-means) z uporabo prilagodljive dolžinske norme. S tem je možno v množici vseh izmerjenih podatkov detektirati podmnožnice podatkov različnih geometrijskih oblik.

Za izgradjo modela je bila uporabljena funkcija *fmclust()*. Končne vrednosti vhodnih parametrov funkcije (podrobnejši opis je podan v [10]) so bile določene glede na izkušnje pridobljene med procesom modeliranja:

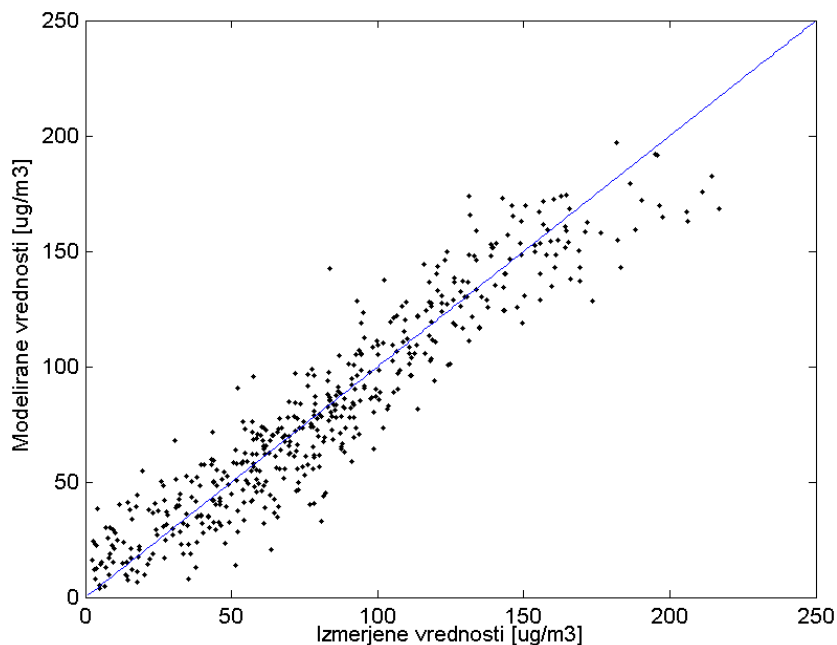
- število grup (določa tudi število pravil): 3,
- parameter mehkosti (večja vrednost določa mehkejšo (bolj prekrivajočo se grupe): 2,
- tip pripadnostnih funkcij pogojnega dela mehkega modela: 2 (uporaba projeciranih pripadnostnih funkcij, ki naj bi glede na priložena navodila orodja omogočala natančnejši model),
- tip posledičnega dela mehkega modela: 1 (uporaba globalne metode najmanjših kvadratov).

Parametri modela so bili dodatno izboljšani z uporabo funkcije *anfis()* [54], ki se poslužuje metode povratnega učenja. Za proces učenja je bila učna množica podatkov razdeljena na množico za učenje (437 vzorcev) in množico za sprotno preverjanje učenja (51 vzorcev).

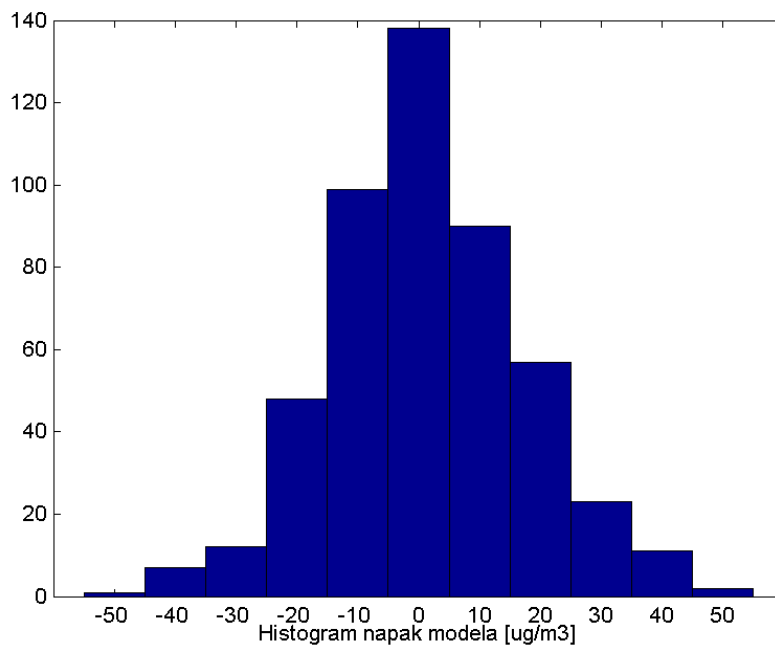
Po končanem učenju je bil najprej preizkušen model na podatkih iz učne množice. Odvisnost med izmerjenimi in izračunanimi vrednostmi je prikazana na sliki 49. Na sliki 50 pa je prikazan še histogram napak, ki kaže značilnosti normalne porazdelitve. Napake so bile izračunane kot razlika med izmerjeno in modelirano vrednostjo. Iz podatkov, ki so bili uporabljeni za indentifikacijo modela sem izračunal tudi vrednosti cenilk, ki sem jih uporabljal za vrednotenje modela (poglavje 4.4):

$$r=0.941, RMSE=16.0, p^6=0.92, SP_{180}=26.7, SR_{180}=100.0, SI_{180}=26.7, SP_{140}=81.5, SR_{140}=76.8, SI_{140}=77.8.$$

Zopet poudarjam, da prikazane vrednosti ne odražajo prave kvalitete modela, ampak so samo informativne narave.



Slika 49: Rezultati preizkusa modela z mehko logiko s podatki iz učne množice



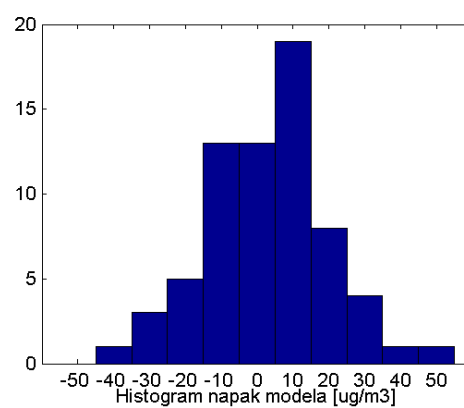
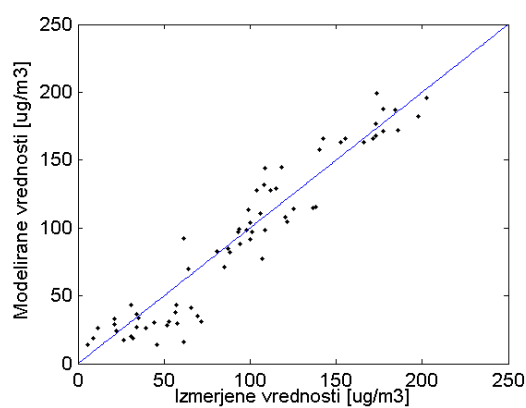
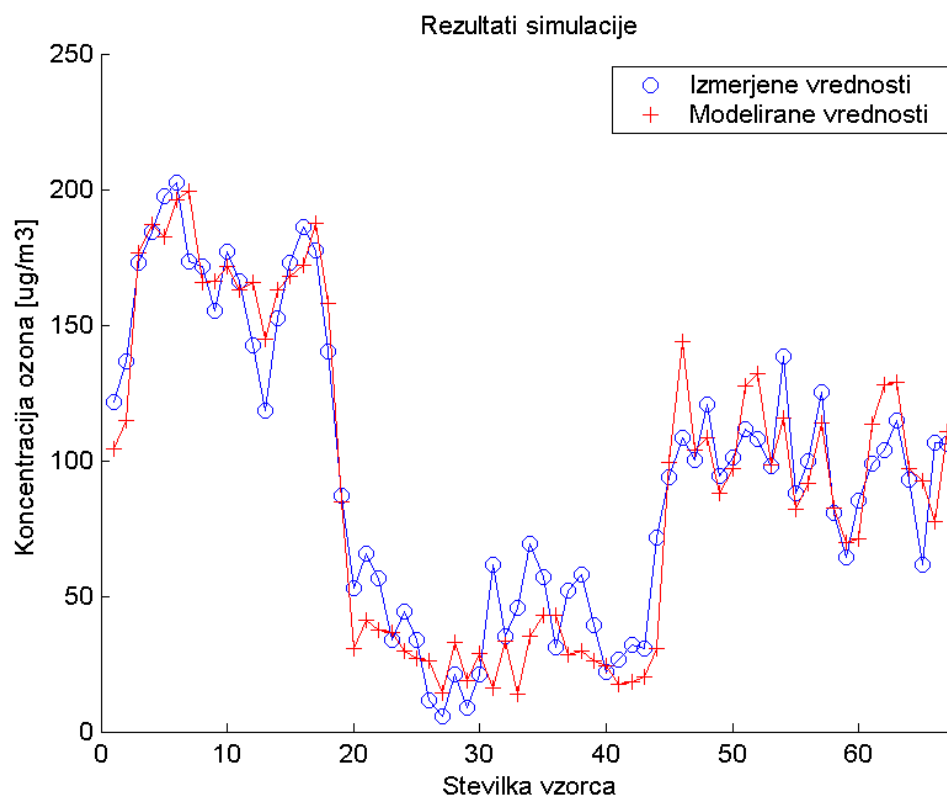
Slika 50: Histogram napak preizkusa modela z mehko logiko s podatki iz učne množice

7.Vrednotenje modelov

Po končanem procesu učenja sem vse tri zgrajene modele tudi preizkusil s pomočjo preizkusne množice, ki je vsebovala podatke za mesece: avgust 2003, januar 2004 in september 2004. Kot sem že predhodno omenil, ti podatki niso bili uporabljeni v procesu učenja modelov. Za vse tri modele sem pri preizkušanju izračunal tudi vse štiri vrste cenilk, ki so predstavljene v prejšnjem poglavju.

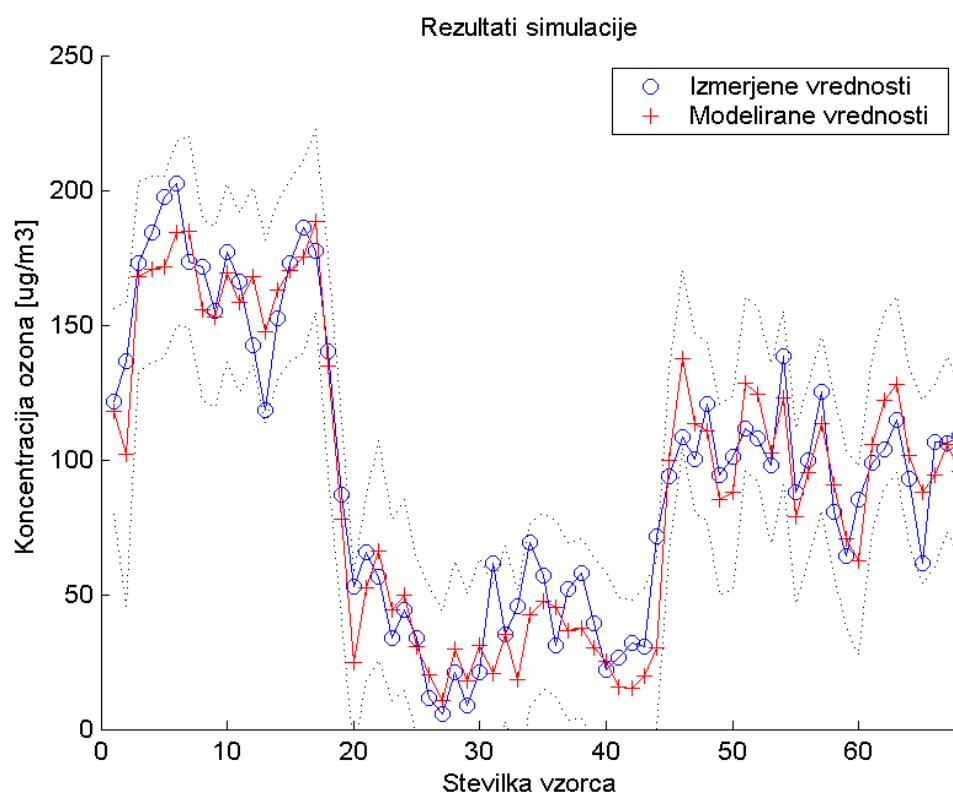
Na slikah v nadaljevanju so prikazani rezultati preizkušanja modelov zgrajenih z umetno nevronske mrežo (slika 51), Gaussovimi procesi (slika 54) in mehko logiko (slika 57). Z modro (temnejšo) krivuljo in krogi so prikazane dejansko izmerjene vrednosti, z rdečo (svetlejšo) krivuljo in križci pa so prikazane izračunane (napovedane) vrednosti izhoda modela. Model z Gaussovimi procesi podaja tudi dvojno vrednost standardne deviacije za vsako točko napovedi, kar je na grafu prikazano s črtkano črto. Na slikah 55, 58 in 61 je za vsak model predstavljeno tudi odstopanje med izmerjenimi in izračunanimi vrednostmi. Na slikah 53, 56 in 59 so za vsak model prikazani še histogrami napak.

7.1.Preizkušanje modela na osnovi umetne nevronske mreže

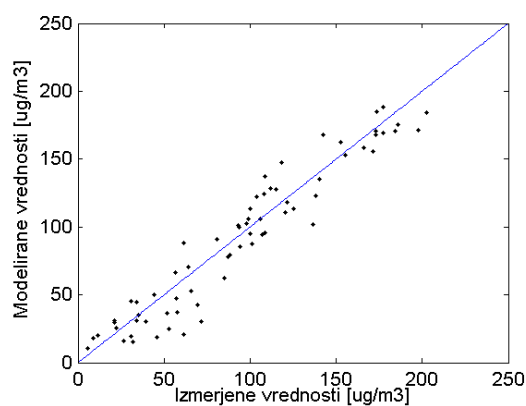


$r=0.955$, $RMSE=17.5$, $p^6=1.00$, $SP_{180}=75.0$, $SR_{180}=60.0$, $SI_{180}=71.8$,
 $SP_{140}=100.0$, $SR_{140}=88.2$, $SI_{140}=96.2$

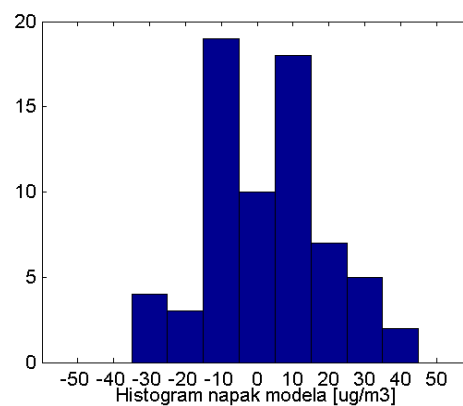
7.2.Preizkušanje modela na osnovi Gaussovih procesov



Slika 54: Rezultat preizkušanja modela na osnovi Gaussovih procesov



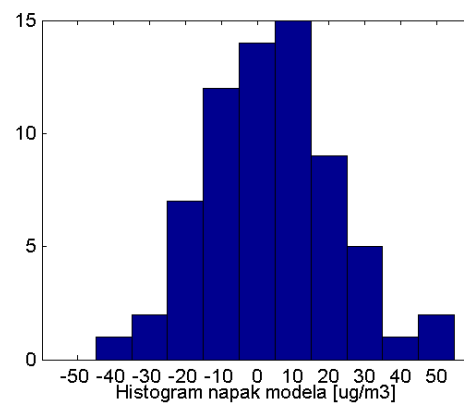
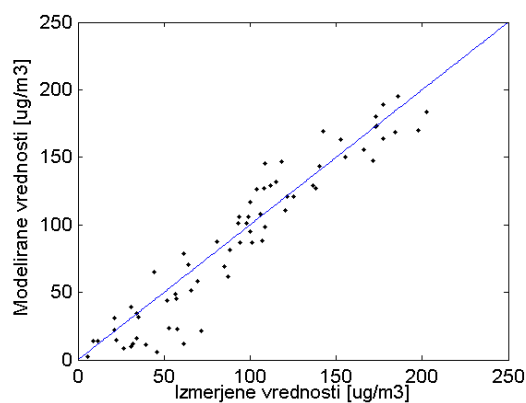
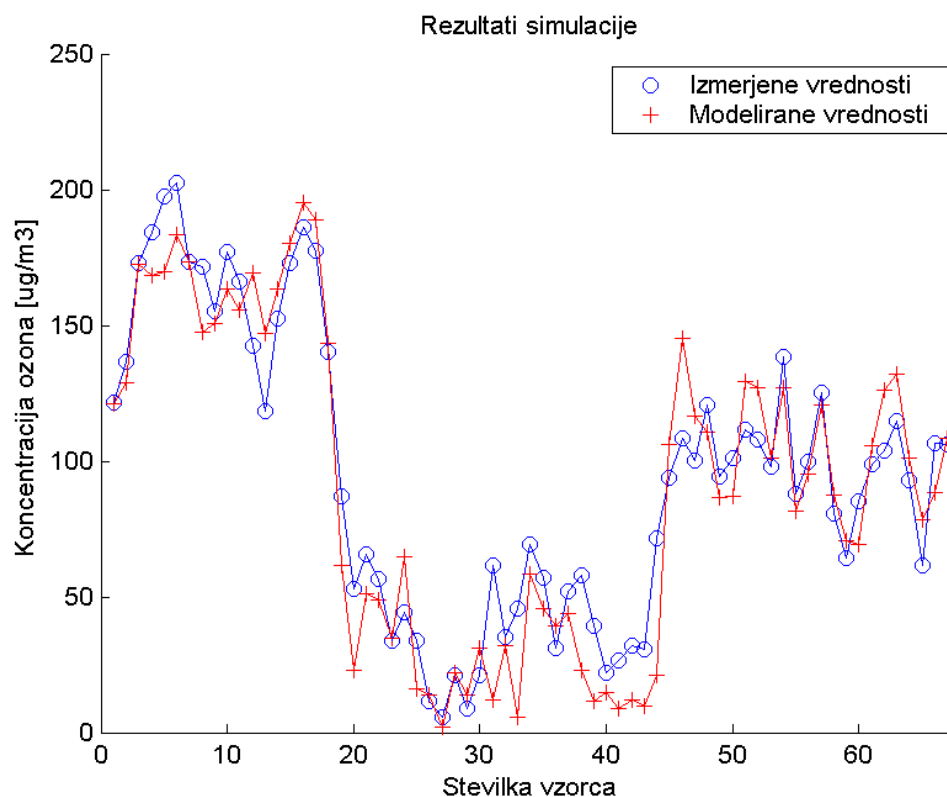
Slika 55: Odstopanje med izmerjenimi in napovedanimi vrednostmi



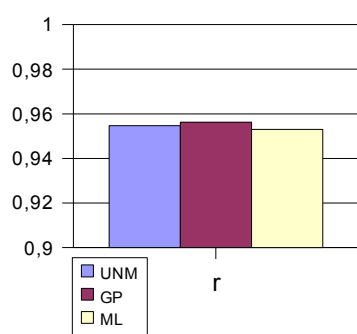
Slika 56: Histogram napak preizkušanja modela na osnovi Gaussovih procesov

$r=0.956$, $RMSE=16.2$, $p^6=0.95$, $SP_{180}=25.0$, $SR_{180}=33.3$, $SI_{180}=21.9$,
 $SP_{140}=93.3$, $SR_{140}=93.3$, $SI_{140}=91.5$

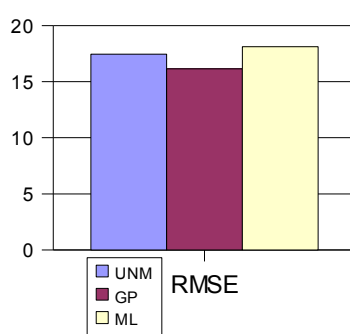
7.3.Preizkušanje modela na osnovi mehke logike



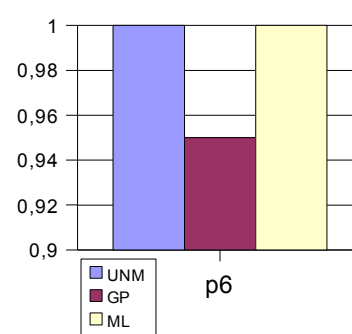
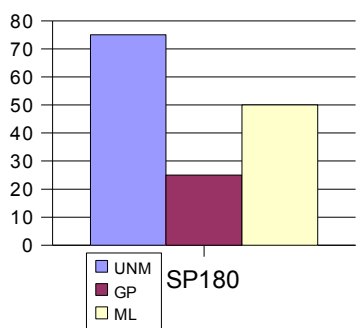
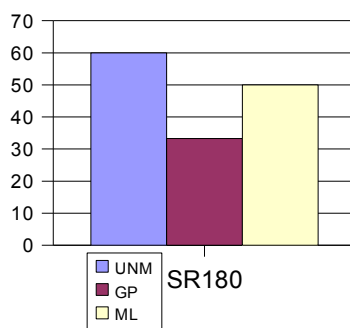
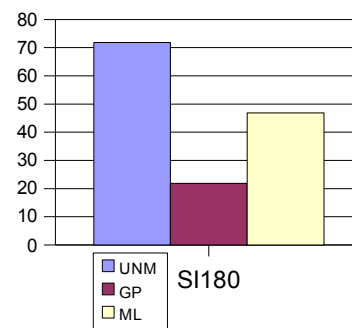
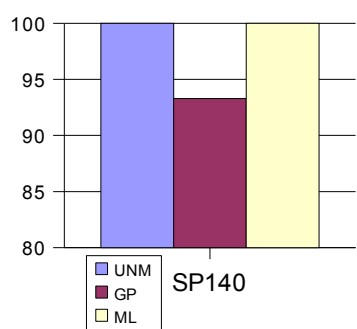
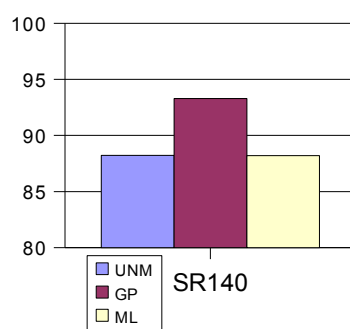
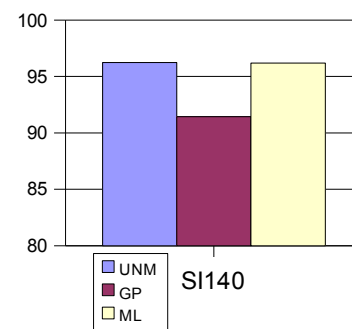
$r=0.953$, $RMSE=18.1$, $p^6=1.00$, $SP_{180}=50.0$, $SR_{180}=50.0$, $SI_{180}=46.8$,
 $SP_{140}=100.0$, $SR_{140}=88.2$, $SI_{140}=96.2$

7.4.Primerjava rezultatov preizkušanja modelov

Slika 60: Primerjava korelacijskega faktorja



Slika 61: Primerjava RMSE

Slika 62: Primerjava cenilke p^6 Slika 63: Primerjava merila pravih napovedi SP pri meji 180 mg/m^3 Slika 64: Primerjava merila uresničenih napovedi SR pri meji 180 mg/m^3 Slika 65: Primerjava indeksa uspeha SI pri meji 180 mg/m^3 Slika 66: Primerjava merila pravih napovedi SP pri meji 140 mg/m^3 Slika 67: Primerjava merila uresničenih napovedi SR pri meji 140 mg/m^3 Slika 68: Primerjava indeksa uspeha SI pri meji 140 mg/m^3

Primerjave posameznih cenilk so v splošnem pokazale, da mi je uspelo zgraditi tri skoraj enakovredne modele. Primerjava korelacijskega faktorja (slika 60) in kvadratnega korena povprečne kvadratne napake (slika 61) kaže, da je morda model z Gaussovimi procesi za odtenek boljši od ostalih dveh. Vendar pa je glede na glavni namen te naloge, ki je napovedovanje povišanih koncentracij ozona, vrednost cenilke p^6 , ki določa verjetnost uspešne napovedi visokih koncentracij, izkazalo da se je model na preizkusni množici odrezal malce slabše.

Primerjava parametrov sposobnosti pri kritični meji koncentracije 180 mg/m^3 je sicer pokazala, da se je v primeru ocenjevanja vseh treh parametrov sposobnosti SP , SR in SI , najbolje izkazal model na osnovi umetne nevronske mreže. Grafikoni rezultatov so prikazani na slikah 63, 64 in 65.

Ker pa so omenejni rezultati zaradi majhnega števila vzorcev, ki so prekoračili kritično mejo (4 vzorci) vprašljivi, sem napravil tudi oceno parametrov sposobnosti pri nižji kritični meji koncentracije 140 mg/m^3 . Grafikoni rezultatov so prikazani na slikah 66, 67 in 68. V tem primeru se je zopet izkazalo, da je morda model na osnovi Gaussovih procesov za odtenek slabši od ostalih dveh, v glavnem pa se je izkazalo, da imamo opravka z enakovrednimi modeli.

8. Zaključek

Prikazani rezultati primerjave vrednotenja modelov v prejšnjem poglavju kažejo, da je mogoče z različnimi pristopi modeliranja zgraditi povsem enakovredne modele. Pri tem je seveda potrebno upoštevati, da so bili vsi modeli zgrajeni na podlagi istih učnih podatkov in ovrednoteni na podlagi istih podatkov za preizkušanje.

Postopek izgradnje modela je za vse tri postopke modeliranja po mojih izkušnjah precej podoben in zahteva približno enako mero predznanja. Kar je tudi razumljivo, saj so si metode sorodne in vse tri spadajo v isto skupino metod modeliranja. Sorodnost med metodami pa mi je pomagala, da sem po pridobitvi znanja in izkušenj ob proučevanju in učenju umetnih nevronske mreže hitreje osvojil modeliranje z Gaussovimi procesi in mehko logiko. Pri učenju modelov z istimi učnimi podatki se je tudi izkazalo, da je bil za metodi modeliranja z umetno nevronske mrežo in mehko logiko potreben približno enak čas učenja, ki je odvisen predvsem od računske zmogljivosti računalnika. Pri metodi učenja modela z Gaussovimi procesi pa je prišla računska zmogljivost računalnika še bolj do izraza, saj le ta pri optimizaciji hiperparametrov raste s tretjo potenco velikosti učne množice podatkov [53]. Za vse tri modele je bil potreben čas za izračun nove napovedi enak.

Prednost modela z Gaussovimi procesi v primerjavi z modeloma z umetno nevronske mrežo in mehko logiko vidim v tem, da je predikcija izhoda modela na osnovi Gaussovih procesov normalna verjetnostna porazdelitev s srednjo vrednostjo in varianco. Velikost variance, ki je odvisna od območja vhodnega prostora, podaja mero zaupanja v napovedano vrednost izhoda modela. Varianca je majhna na območjih, kjer je bila gostota točk učne množice velika, raste pa s približevanjem področjem, o katerih ima GP model relativno malo informacij.

Druga prednost modela na osnovi Gaussovih procesov pa je optimizacija precej manjšega števila hiperparametrov, kot v primeru umetne nevronske mreže in mehke logike. Število hiperparametrov je vedno vnaprej določeno z obliko kovariančne funkcije in je običajno enako številu vhodov.

Slabost, ki se je izkazala pri optimizaciji modela na osnovi Gaussovih procesov, pa je nastopala v primeru neustreznih začetnih vrednosti hiperparametrov. Uporabljena metoda največje podobnosti je v primeru slabo nastavljenih začetnih vrednosti hiperparametrov zelo hitro postala numerično nestabilna ali pa končala v lokalnem minimumu.

Izkušnje pridobljene pri izdelavi naloge kažejo, da se je potrdila trditev iz [40,15], da je dobra identifikacija modela odvisna predvsem od pravilne izbire vhodnih spremenljivk. Glede na pregled priporočil iz literature se je izkazalo, da imajo nekatere vhodne spremenljivke bistveno manjši vpliv na izhod, kot je bilo omenjeno, ali pa celo onemogočajo dobro identifikacijo. Vhodni spremenljivki kot sta prah in žveplov dioksid v konkretnem primeru modeliranja za Novo Gorico pa celo poslabšujeta identifikacijo modela. Precej pa se je identifikacija izboljšala z uvedbo napovedi temperature in vetra. Med izvajanjem naloge se je izkazalo, da je zelo močna tudi povezava med koncentracijo ozona in temperaturo zraka, kar je tudi smiselno. Fotokemična reakcija nastajanja in razgradnje potrebuje določeno energijo, ki je v glavni meri sorazmerna temperaturi.

Meje pri vrednotenju s cenilko p^6 so bile postavljene razmeroma strogo zaradi primerjanja med različnimi modeli med procesom identifikacije. Dobljeni rezultati sicer povsem ustrezajo uredbi [4], ki predpisuje dovoljeno relativno napako napovedovanja ozona 50%, kljub temu pa rezultati kažejo, da bi se dalo model še malce izboljšati z večjo učno množico predvsem prekoračenih vrednosti, kar bo možno v prihodnosti. Obenem pa bo ob večji učni množici podatkov za preizkušanje tudi priložnost za sprotno ocenjevanje modela.

Nova Gorica se bo kot urbano okolje v prihodnosti verjetno še širila, kar morda pomeni povečevanje števila prekoračenih vrednosti koncentracij ozona. Ker za vpliv na predvidene povišane koncentracije ozona v urbanem okolju zaenkrat še nimamo direktnih vzvodov (zaenkrat študije kažejo, da zmanjšanje prometa v mestih samo še poveča koncentracije ozona, ker se zmanjša koncentracija dušikov oksidov, ki porabljajo ozon), se mi zdi uporaba modela za napovedovanje povišanih koncentracij ozona edino smiselna in upravičena za varovanje zdravja ljudi. Zgrajene modele bi se dalo po mojem mnenju in glede na priporočila iz literature dodatno izboljšati z uporabo dodatnih vhodnih spremenljivk. Takšno skupino vhodnih spremenljivk, ki bi lahko glede na priporočila iz literature vplivala na izboljšavo predstavljajo koncentracije hlapnih organskih spojin (VOC). Merilna postaja v Novi Gorici je sicer imela ob postavitvi vgrajen merilnik hlapnih organskih spojin, vendar pa

zaradi tehničnih težav z merilnikom zaenkrat ni mogoče pridobiti uporabnih meritev. Pri tem pa bi bilo tudi potrebno raziskati vpliv vsake posamezne hlapne organske spojine na nastajanje ozona in njen pomen za potrebo izboljšanja napovedovanja povišanih koncentracij ozona.

9.Viri in literatura

- [1] Abdul-Wahab S. A., Al-Alawi S.M., Assessment and Prediction of Tropospheric Ozone Concentration Levels using Artificial Neural Networks, *Environmental Modelling & Software*, 17(3), 219-228, 2002.
- [2] Agencija Republike Slovenije za okolje, Mesečni bilten, http://www.arso.gov.si/o_agenciji/knji~znica/publikacije/bilten.htm, 12.11.2004
- [3] Agencija Republike Slovenije za okolje, Podatki za nekatere postaje v obdobju 1991-2000, http://www.arso.gov.si/podro~cja/vreme_in_podnebje/napovedi_in_podatki/podneb_10_tabele.html, 21.12.2004
- [4] Agencija Republike Slovenije za okolje, Poročilo o stanju okolja v Sloveniji, http://www.arso.gov.si/poro~cila/Poro~cila_o_stanju_okolja_v_Sloveniji/zrak.pdf, 12.11.2004
- [5] Agencija Republike Slovenije za okolje, Vreme in podnebje – napovedi in podatki, http://www.arso.gov.si/podro~cja/vreme_in_podnebje/napovedi_in_podatki/index.html, 10.01.2004
- [6] Ainslie B. , Steyn D. G., Scaling analysis of ozone precursor relationships, *Air pollution modeling and its application XVI*, Edited by Borrego and Incecik, Kluwer Academic/Plenum Publishers, New York, 321-328, 2004
- [7] Aleksić N. M., Milutinović M., Milutinović P., Informacioni sistemi za kontrolu i upravljanje kvalitetom vazduha u urbanim sredninama, Institut za fiziku, Beograd
- [8] Atmospheric Research and Technology, LLC, About SODAR, http://www.sodar.com/about_sodar.htm, 12.12.2004
- [9] Ažman K., Identifikacija dinamičnih sistemov z Gaussovimi procesi z vključenimi linearnimi modeli, Magistrsko delo, Fakulteta za elektrotehniko, 2004.
- [10] Babuška R., Fuzzy Identification Toolbox for MATLAB 6.5 (ver. 3.03). <http://www.dcsc.tudelft.nl/~babuska/>, 28.02.2004

- [11] Ballester E. B., Valls G. C., Carrasco-Rodriguez J. L., Olivas E. S., S. del Valle-Tascon, Effective 1-day ahead prediction of hourly surface ozone concentrations in eastern Spain using linear models and neural networks, *Ecological Modelling* 156, 27-41, Elsevier, 2002
- [12] Bochner S., *Harmonic analysis and the theory of probability*, Berkeley, 1979.
- [13] Box George E. P., Jenkins Gwilym M., *Time series analysis forecasting and control*, San Francisco, Holden-Day, 1976
- [14] Božnar M. Z., Mlakar P., *Use of neural networks in the field of air pollution modelling*, Ath, Edited by Borrego and Schayes, Kluwer Academic/Plenum Publishers, New York, 2002
- [15] Božnar M., *Izbira učnih vzorcev za model napovedovanja ozračja na osnovi nevronske mreže*, Doktorska disertacija, Fakulteta za elektrotehniko, Univerza v Ljubljani, 1997
- [16] Božnar 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* Vol. 27B, Pergamon Press Ltd, Great Britain, 221-230, 1993
- [17] Comrie A. C., Comparing neural networks and regression models for ozone forecasting, *Journal of Air & Waste Management Association* 47, 653-663, 1997
- [18] Dodge M., Chemical oxidant mechanisms for air quality modeling: critical review. *Atmospheric Environment* 34, 2103-2130, 2000
- [19] EPA – Environmental Protection Agency, <http://www.epa.gov/air/oaqps/eog/ozonehealth/aqi.html>, 06.11.2004
- [20] EPA – Environmental Protection Agency, <http://www.epa.gov/airnow/>, 06.11.2004
- [21] Finzi G., Nunnari G., Air quality forecast and alarm systems, *Air Quality Modelling- Theories, Methodologies, Computational Techniques and Available Databases and Software*, Vol. II, Ch. 16, EnviroComp Institute and Air & Waste Management Association co-publishers, 2004
- [22] Finzi G., Tebaldi G., A mathematical model for air pollution forecast and alarm in an urban area, *Atmospheric Environment*, vol. 16, 9,

2055-2059, 1982

- [23] Gardner M. W., Dorling S. R., Artificial neural networks (the multilayer perceptron)-A review of applications in the atmospheric sciences, Atmospheric Environment Vol. 32, Elsevier Science Ltd., Great Britain, 1998
- [24] Gardner M. W., Dorling S. R., Neural network modelling of the influence of local meteorology on surface ozone concentrations, Proceedings 1st International Conference on GeoComputation, University of Leeds, 359-370, 1996.
- [25] Geai P., Methode D'interpolation et de Reconstruction Tridimensionnelle d'un Champ de Vent, le Code d'Analyse Objective MINERVE, Technical Report EDF/DER/HE/34-87.03
- [26] Gibbs M. N., Bayesian Gaussian Processes for Regression and Classification, Ph.D. Disertation, Cambridge University, 1997.
- [27] Gutfreund P., Liu C., Nicholson B., Roberts E., COMPLEX I and II Model Performance Evaluation in Nevada and New Mexico, Journal of Air Pollution Control Association, Vol. 33, 846-871, 1983
- [28] Haagen-Smit A. J., Chemistry and physiology of Los Angeles Smog. Industrial and Engineering Chemistry 44, 1342-1346, 1952
- [29] Hornik K., Stinchcombe M., White H., Multi-layer feedforward networks are universal approximators, Neural Networks 2, 1989.
- [30] How Ozone Affects Bacteria, Fungus, Molds And Viruses, <http://mold-help.org/content/view/436/>, 11.12.2004
- [31] IAEA Training course in Planning, Preparedness & Response to Radiological Emergencies, lecture 20.5.3, Basic meteorology – plume transport, Argonne national laboratory, Argonne, Illinois, United States of America, 1982
- [32] Kloss S. D., A Discussion on Ozone Chemistry, <http://www.environmental-expert.com/article316.htm>, 06.11.2004
- [33] Krušič M., Slovenija, Turistični vodnik, Mladinska knjiga, Ljubljana, 2002
- [34] Lamb B., Gaseous pollutant characteristics. Handbook of air pollution technology. John Wiley and Sons, Inc. USA, 1984
- [35] Lawrence J., Introduction to Neural Networks, California Scientific Software, Grass Valley, USA, 1991

- [36] Likar B., Prediktivno vodenje nelinearnih sistemov na osnovi Gaussovih procesov, Magistrsko delo, Fakulteta za elektrotehniko, 2004.
- [37] MacKay D. J. C., Hyperparameters: Optimize, or integrate out? In Maximum Entropy and Bayesian Methods, Santa Barbara 1993, ed. by G. Heidbreder, pp. 43--60, Dordrecht, Kluwer, 1996.
- [38] Mestna občina Nova Gorica, Spletna stran mestne občine Nova Gorica, <http://www.nova-gorica.si>, 22.12.2004
- [39] Miklavčič Brezigar I., Nova Gorica, Vodnik po mestu, IKI – Institut za komunikacije in informatiko, Ljubljana, 1998
- [40] Mlakar P., Izbira značilnk vzorcev za model kratkoročnega napovedovanja onesnaževanja ozračja, Doktorska disertacija, Fakulteta za elektrotehniko, Univerza v Ljubljani, 1996
- [41] Moreno J., Moreno-Grau S., Bayo J., M. Angosto J., Jimenez J. E., Moreno-Clavel J., Photochemical pollution precursors and ozone in the athmosphere of Cartagena from 1995 to 1997. Wessex institute of techology, Air pollution 98, 1998.
- [42] Neal R. M., Bayesian Learning for Neural Networks, Number 118 in Lecture Notes in Statistics, New York, Springer, 1996.
- [43] NeuroShell 2, Users manual, Ward Systems Group, Inc., Frederick, Maryland, 1993
- [44] Paine R. J., Eagan B. A., User's guide to Rough Terrain – Diffusion Model (RTDM), EPA/SW/MT-88/041a, U.S. EPA Research Triangle Park, NC, 1987
- [45] Periodic elements, Oxygen, <http://pearl1.lanl.gov/periodic/elements/8.html>, 06.11.2004
- [46] Perry S. G., CTDMPLUS: A Dispersion Model for Source near Complex Topography, Part I: Technical Formulation, Journal of Applied Meteorology, Vol. 31, 633-645, 1992
- [47] Rasmussen C. E., Evaluation of Gaussian Processes and other Methods for Non-Linear Regression, Ph.D. Disertation, Graduate department of Computer Science, University of Toronto, Toronto, 1996.
- [48] Rasmussen C. E., Home page, Gaussian Processes for Regression software code, <http://www.gatsby.ucl.ac.uk/~edward/code/gp/>,

10.10.2004

- [49] Rojko A., Položajno vodenje nelinearnih mehanizmov z uporabo mehke logike, Doktorat, Fakulteta za elektrotehniko, računalništvo in informatiko, Univerza v Mariboru, 2002
- [50] Rumelhart D. E., G. E. Hinton, R. J. Williams, Learning internal representations by error propagation, Paraller distributed processing: Explorations in the Microstructure of Cognition, Vol. 1, MIT Press, Cambridge, MA, 1986.
- [51] Škrjanc I., Fuzzy modelling and identification, Fakulteta za elektrotehniko, Univerza v Ljubljani, 2003
- [52] Takagi T., Sugeno M., Fuzzy Identification of Systems and Its Applications to Modeling and Control, IEEE Trans. on Systems, man and Cybernetics, vol.15, no.1, str. 116-132, januar 1985.
- [53] The Math Forum @ Drexel, Ask Dr. Math, Questions and Answers from our archives, Complexity of Matrix Inversion, <http://mathforum.org/library/drmath/view/51908.html>, 06.12.2005
- [54] The MathWorks, Inc. Matlab® version 6.5. Matlab documentation.
- [55] Tinarelli G., et all, Lagrangian Particle Simulation of Tracer Dispersion in the Lee of a Schematic Two-dimensional Hill, Journal of Applied Meteorology, Vol. 33, No. 6, 744-756, 1994
- [56] Tomšič G., Matematika IV, Fakulteta za elektrotehniko, 1998.
- [57] Uredba o benzenu in ogljikovemu monoksidu v zunanjem zraku. Ur. list RS, št. 52/2002.
- [58] Uredba o ozonu v zunanjem zraku. Ur. list RS, št. 8/2003.
- [59] Uredba o ukrepih za ohranjanje in izboljšanje kakovosti zunanjega zraka. Ur. list RS, št. 52/2002.
- [60] Uredba o žveplovem dioksidu, dušikovih oksidih, delcih in svincu v zunanjem zraku. Ur. list RS, št. 52/2002.
- [61] Van Aalst R.M., F.A.A.M. de Leeuw (urednika), National Ozone Forecasting System And International Data Exchange In Northwest Europe, European Topic Centre on Air Quality, <http://reports.eea.eu.int/TEC09/en/tech09.pdf>, september 1997
- [62] Wikipedia, <http://www.wikipedia.org>, 06.11.2004
- [63] WMO/UNEP Scientific Assessment of Ozone Depletion: 2002,

<http://www.al.noaa.gov/WWHD/Pubdocs/assessment02.html>,
12.12.2004

- [64] Yi J., Prybutok R., A neural network model forecasting for prediction of daily maximum ozone concentration in an industrialized urban area. *Environmental Pollution* 92(3), 349-357, 1996.
- [65] Zadeh L. A., Fuzzy logic = Computing with worlds, *IEEE Trans. Fuzzy Syst.*, vol. 4, no.2, str. 103-110, Maj 1996.
- [66] Zadeh L. A., Fuzzy sets, *Information and Control*, 8, str. 338-353, 1965

Dodatek A: Osnove verjetnostnega modeliranja

A.1. Naključna spremenljivka

Naključna spremenljivka je spremenljivka, katere vrednost je odvisna od naključja. O naključnem procesu (ang. stochastic process) govorimo, ko imamo več neodvisnih (časovno, prostorsko itd.) realizacij naključne spremenljivke. Porazdelitve vrednosti posameznih realizacij naključne spremenljivke so lahko poljubne, opišemo jih s porazdelitvenim zakonom in zalogo vrednosti, s čimer je naključna spremenljivka tudi popolnoma določena. Glede na zalogo vrednosti ločimo diskretne in zvezne naključne spremenljivke, glede na obliko porazdelitvenega zakona pa različne standardne (enakomerna, normalna ali Gaussova, Studentova itd) in nestandardne porazdelitve [56,36]. Splošna oblika porazdelitvenega zakona za naključno spremenljivko X je porazdelitvena funkcija ali distribucija $F(x)$:

$$F(x) = P(X < x) \quad (-\infty < x < \infty) \quad . \quad (\text{A.1})$$

kjer je $P(X < x)$ verjetnost dogodka, da naključna spremenljivka X zavzame vrednost manjšo od x . V primeru zvezne naključne spremenljivke (zaloga vrednosti spremenljivke je zvezna množica) lahko porazdelitveno funkcijo spremenljivke X zapišemo tudi v obliki

$$F(x) = \int_{-\infty}^x p(t) dt \quad , \quad (\text{A.2})$$

kjer je $p(x)$ gostota verjetnosti ali verjetnostna gostota (ang. PDF-probability density function). Verjetnost, da zvezna naključna spremenljivka X zavzame vrednost na intervalu $[a, b]$ zapišemo

$$P(a \leq X \leq b) = F(b) - F(a) = \int_a^b p(t) dt \quad . \quad (\text{A.3})$$

Za vsako porazdelitveno funkcijo pa velja, da je v primeru nemogočega dogodka ($X \rightarrow -\infty$)

$$F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = 0 \quad (\text{A.4})$$

in v primeru gotovega dogodka ($X \rightarrow \infty$)

$$F(\infty) = \lim_{x \rightarrow \infty} F(x) = 1 \quad . \quad (\text{A.5})$$

A.2. Analiza porazdelitev naključnih spremenljivk

Pri analizi porazdelitev naključnih spremenljivk si pomagamo z različnimi številiškimi karakteristikami. Najpomembnejše so predstavljene v nadaljevanju.

Matematično upanje ali povprečna vrednost naključne spremenljivke X , $E(X)$:

$$E(x) = \int_{-\infty}^{\infty} xp(x) dx \quad . \quad (A.6)$$

Varianca ali disperzija naključne spremenljivke $D(X)$ je definirana kot povprečje kvadratov odstopanj spremenljivke X od njene povprečne vrednosti $E(X)$:

$$\text{var}(X) = E((X - E(X))^2) = \int_{-\infty}^{\infty} (x - E(X))^2 p(x) dx = E(X^2) - (E(X))^2 \quad (A.7)$$

Standardna deviacija je definirana kot pozitivni kvadratni koren variance naključne spremenljivke X :

$$\sigma_X = \sigma(X) = +\sqrt{\text{var}(X)} \quad . \quad (A.8)$$

Kovarianca naključnih spremenljivk X in Y je matematično upanje produkta odklonov spremenljivk od njunih povprečnih vrednosti in je mera za linearno odvisnost spremenljivk:

$$\text{cov}(X, Y) = E((X - E(X))(Y - E(Y))) \quad . \quad (A.9)$$

Korelacijski koeficient naključnih spremenljivk X in Y je s standardnima deviacijama normirana kovarianca spremenljivk:

$$r(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} \quad . \quad (A.10)$$

Naključni spremenljivki sta nekorelirani (linearno neodvisni) natanko takrat, kadar je $r(X, Y) = 0$.

A.3. Normalna (Gaussova) porazdelitev

Normalno ali Gaussovo porazdeljeno spremenljivko, definirano z gostoto verjetnosti (A.1) na kratko zaznamujemo z $N(\mu, \sigma)$, kjer je μ srednja vrednost, σ pa standardna deviacija naključne spremenljivke:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad . \quad (A.11)$$

A.4. Naključni vektorji

Naključni vektor je vektor, katerega komponente so naključne spremenljivke. Če imamo dve ali več med seboj povezanih naključnih spremenljivk, njihovo realizacijo imenujemo naključni proces (ang. stochastic process)

$$X = (X_1, X_2, \dots, X_n) \quad (A.12)$$

Podobno kot za naključno spremenljivko, lahko za naključni vektor zapišemo porazdelitveni zakon v obliki:

$$F(x_1, x_2, \dots, x_n) = P(X_1 < x_1, X_2 < x_2, \dots, X_n < x_n) \quad (A.13)$$

Če so komponente naključnega vektorja zvezne naključne spremenljivke, potem velja:

$$F(x_1, x_2, \dots, x_n) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} p(x_1, \dots, x_n) dx_1 \dots dx_n \quad (A.14)$$

kjer je $p(x_1, \dots, x_n)$ n -dimenzionalna gostota verjetnosti (nenegativna zvezna funkcija). Verjetnost, da naključni vektor zavzame vrednost na določenem intervalu imenujemo skupna ali totalna verjetnost in jo zapišemo:

$$P(a_1 \leq x_1 \leq b_1, \dots, a_n \leq x_n \leq b_n) = \int_{a_1}^{b_1} \dots \int_{a_n}^{b_n} p(x_1, \dots, x_n) dx_1 \dots dx_n \quad (A.15)$$

Porazdelitve posameznih komponent naključnega vektorja imenujemo tudi robne ali marginalne porazdelitve:

$$p(\infty, \dots, \infty, x_i, \infty, \dots, \infty) = p_i(x_i) = \int_{-\infty}^{i-1} \dots \int_{-\infty}^{i+1} p(x_1, \dots, x_n) dx_1 \dots dx_n \quad (A.16)$$

Enako velja za pogojno verjetnostno gostoto:

$$p(X_i | X_j) = \frac{p(X_i, X_j)}{p(X_j)} \quad (A.17)$$

kjer je $P(X_i, X_j)$ skupna verjetnostna gostota in $P(X_j)$ marginalna verjetnostna gostota spremenljivke (komponente) X_j naključnega vektorja.

Najpomembnejši med številiškimi karakteristikami naključnega vektorja sta:

Matematično upanje ali povprečna vrednost naključnega vektorja X je vektor $E(X)$, sestavljen iz povprečnih vrednosti posameznih komponent

naključnega vektorja:

$$E(\mathbf{X}) = (E(X_1), \dots, E(X_n)) \quad (\text{A.18})$$

Varianca ali disperzija naključnega vektorja $\mathbf{X}$ je definirana s kovariančno matriko $\Sigma(\mathbf{X})$, sestavljeno iz kovarianc med posameznimi komponentami vektorja:

$$\Sigma(\mathbf{X}) = \begin{bmatrix} K_{11} & K_{12} & \dots & K_{1n} \\ K_{21} & K_{22} & \dots & K_{2n} \\ \dots & \dots & \dots & \dots \\ K_{n1} & K_{n2} & \dots & K_{nn} \end{bmatrix} \quad (\text{A.19})$$

Velja:

1. $K_{ij} = \text{cov}(X_i, X_j)$,
2. $K_{ij} = K_{ji}$, $\Sigma(\mathbf{X})$ je simetrična,
3. $K_{ii} = \text{var}(X_i)$, diagonalni elementi $\Sigma(\mathbf{X})$ predstavljajo varianco posamezne komponente,
4. iz lastnosti variance, ki je vedno nenegativno število, sledi, da je izpolnjen eden izmed potrebnih pogojev (vendar ne zadostnih pogojev), da je $\Sigma(\mathbf{X})$ pozitivno definitna matrika.

Normalno ali Gaussovo porazdeljen naključni vektor je določen z verjetnostno gostoto:

$$p(\mathbf{x}) = \frac{1}{(\sqrt{2\pi})^n \sqrt{|\Sigma|}} e^{\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)} \quad (\text{A.20})$$

kjer je:

- n dimenzija naključnega vektorja $\mathbf{X}$,
- $\boldsymbol{\mu}$ vektor srednjih vrednosti naključnega vektorja $\mathbf{X}$,
- Σ kovariančna matrika naključnega vektorja $\mathbf{X}$.

Normalna porazdelitev je funkcija vektorja srednjih vrednosti in kovariančne matrike, zato pogosto uporabljamo skrajšani zapis $N(\boldsymbol{\mu}, \Sigma)$.

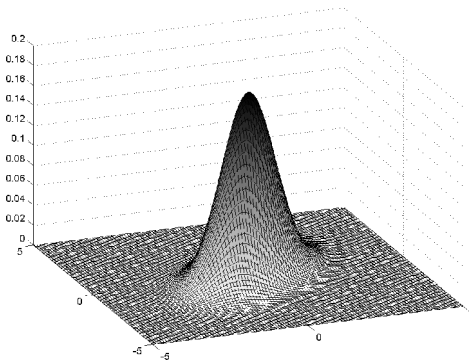
Dvodimenzionalno normalno porazdelitev lahko zapišemo v obliki:

$$p(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{(1-r^2)}} e^{-\frac{1}{2(1-r^2)}\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - \frac{2r(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)} \quad (\text{A.21})$$

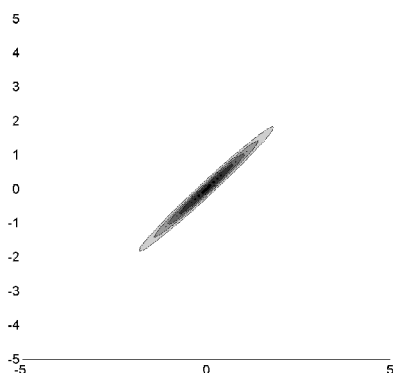
kjer r predstavlja korelacijski koeficient med naključnima spremenljivkama. V tem primeru je:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & r\sigma_1\sigma_2 \\ r\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \quad (\text{A.22})$$

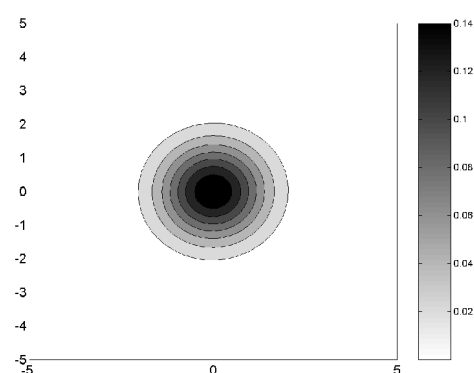
Na slikah od 69 do 71 so predstavljene nekatere značilne oblike dvodimenzionalnih normalnih porazdelitev. Na sliki 69 je prikazan splošen primer dvodimenzionalne Gaussove porazdelitvene funkcije v obliki tridimezionalnega grafa iz stranskega pogleda. Na ostalih slikah od 70 do 71 pa so prikazane dvodimenzionalne Gaussove porazdelitvene funkcije tudi v obliki tridimezionalnega grafa, zaradi nazornosti pa je bil izbran pogled iz florisa.



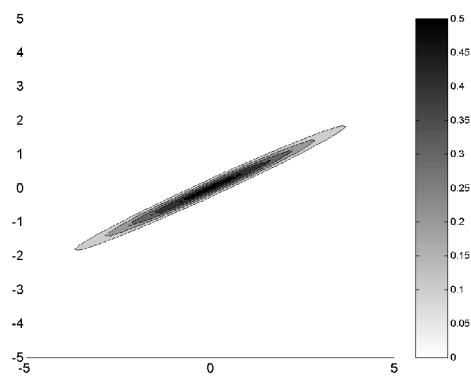
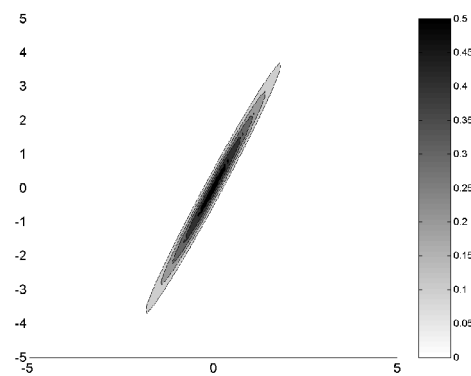
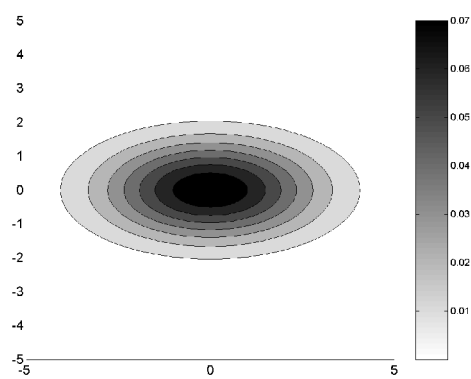
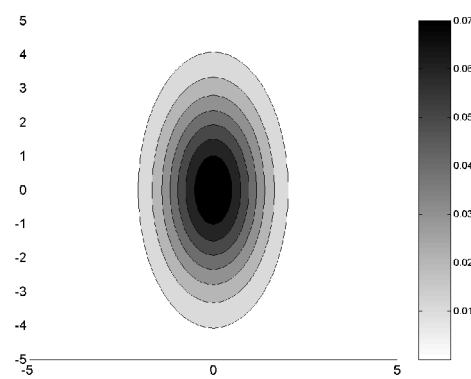
Slika 69: Primer dvodimenzionalne Gaussove porazdelitvene funkcije



Slika 70: $\sigma_1=1.0$, $\sigma_2=1.0$, $r=0.99$



Slika 71: $\sigma_1=1.0$, $\sigma_2=1.0$, $r=0.00$

Slika 72: $\sigma_1=1.0, \sigma_2=2.0, r=0.99$ Slika 73: $\sigma_1=2.0, \sigma_2=1.0, r=0.99$ Slika 74: $\sigma_1=1.0, \sigma_2=2.0, r=0.0$ Slika 75: $\sigma_1=2.0, \sigma_2=1.0, r=0.0$

Dodatek B: Preprost primer modeliranja nelinearne funkcije z Gaussovimi procesi

B.1. Identifikacija modela

Za praktičen primer modeliranja z Gaussovimi procesi sem si izbral nelinearno enačbo (B.1), ki opisuje temperaturo rosišča. Temperatura rosišča je tista temperatura, pri kateri zrak ali plin doseže 100 % relativno vlago. Če se pri nespremenjeni vsebnosti vodne pare v zraku njegova temperatura zniža, bo določen del vlage kondenziral.

$$\begin{aligned} a &= 15.0 \cdot b - 2.1 \cdot (Tz - 0.16) + 2711.5 \\ b &= \log\left(\frac{100.0}{Rv}\right) \\ c &= 273.0 + Tz \\ Tr &= \left(\frac{c \cdot a}{a + \frac{c \cdot b}{2}}\right) - 273.0 \end{aligned} \quad (B.1)$$

Pomen oznak v enačbi je sledeč:

Tz ...temperatura zraka,

Rv ...relativna vlaga zraka in

Tr ...temperatura rosišča.

V nadaljevanju so prikazani različni modeli nelinearne funkcije temperature rosišča, ki so bili modelirani z Gaussovimi procesi.

Skupne lastnosti vseh modelov so:

- zgrajeni so bili pri konstantni temperaturi zraka $Tz=20^{\circ}C$,
- imajo samo eno vhodno spremenljivko, to je relativna vlaga zraka Rv ,
- za indentifikacijo je bilo uporabljeno enako število učnih točk,
- izhod modela predstavlja temperatura rosišča v odvisnosti od relativne vlage zraka pri konstatni temperaturi zraka.

Izhod vsakega modela je prikazan na dveh slikah. Na prvih slikah so prikazani odzivi modela:

- z znakom o (krogec) so prikazane učne točke,
- z znakom . (pika) pa izhodi modela,
- pikčasta črta prikazuje vrednost dvojne standarne deviacije izhoda iz modela,

- polna črta prikazuje temperaturo rosišča glede na enačbo (B.1),

Druga slika pa prikazuje:

- napako, ki je izračunana kot razlika med izhodom modela in pravo vrednostjo glede na enačbo (B.1),
- pikčasta črta prikazuje vrednost dvojno standardno deviacijo izhoda iz modela,

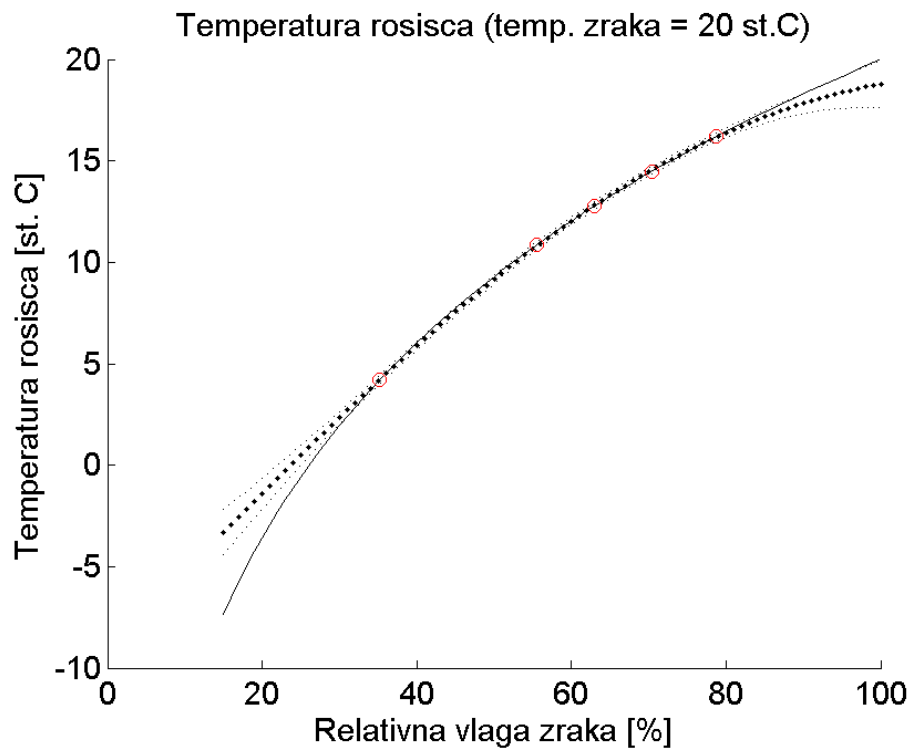
Prvi model je bil identificiran s pomočjo petih učnih točk, ki so bile naključno izbrane na območju relativne vlage od 20% do 100%. Prikazan je na slikah 76 in 77.

Drugi model je bil identificiran s pomočjo petih učnih točk, ki so bile naključno izbrane na območju relativne vlage od 20% do 100%. Od prvega modela se razlikuje v tem, da so je bil učnim točkam dodan še šum. Model je prikazan na slikah 78 in 79.

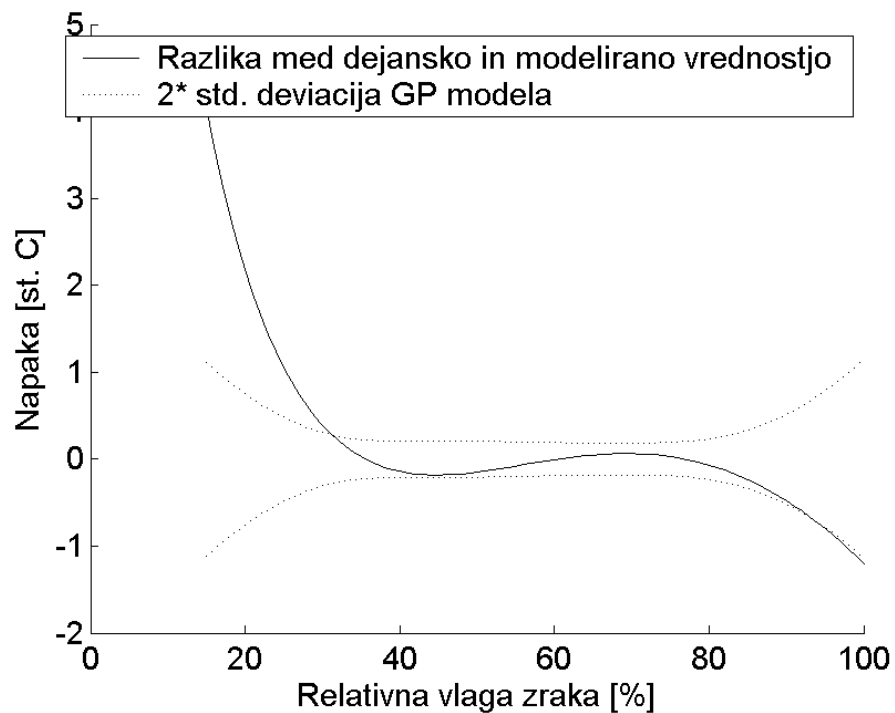
Tretji model je bil identificiran s pomočjo petih učnih točk, ki so bile izbrane na območju relativne vlage od 20% do 100% v enakomernih presledkih. Prikazan je na slikah 80 in 81.

Četrty model je bil identificiran s pomočjo petih učnih točk, ki so bile izbrane na območju relativne vlage od 20% do 100% v enakomernih presledkih. Od tretjega modela se razlikuje v tem, da so je bil učnim točkam dodan še šum. Prikazan je na slikah 82 in 83.

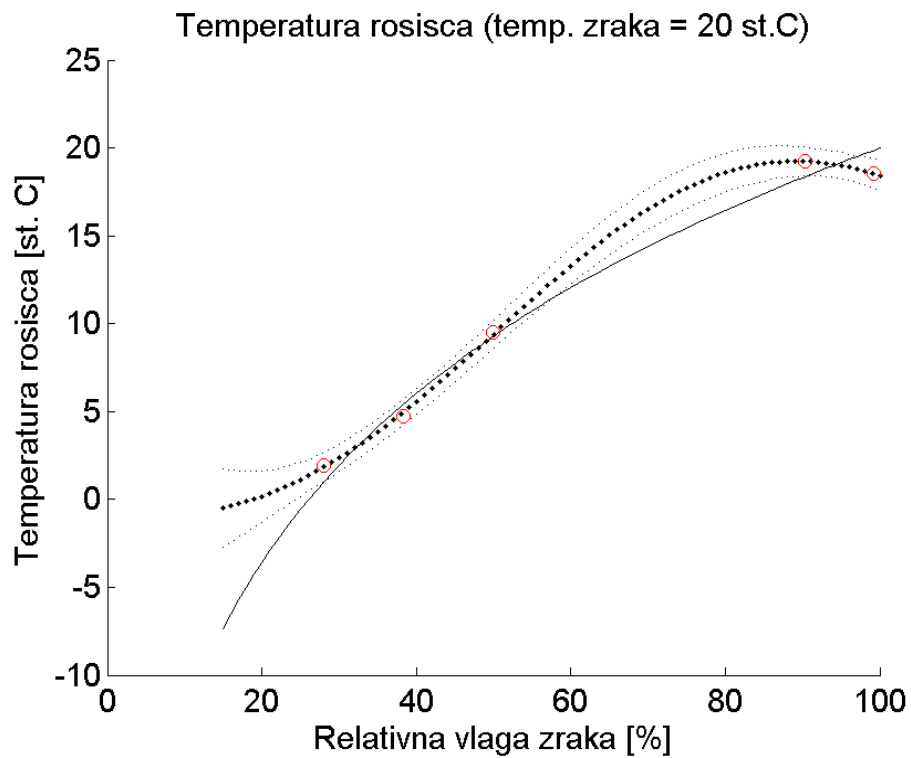
Peti model je bil identificiran s pomočjo petnajstih učnih točk, ki so bile izbrane na območju relativne vlage od 20% do 100% v enakomernih presledkih. V vsakem presledku so bile določene tri učne točke, ki so imele dodan šum. Prikazan je na slikah 84 in 85.



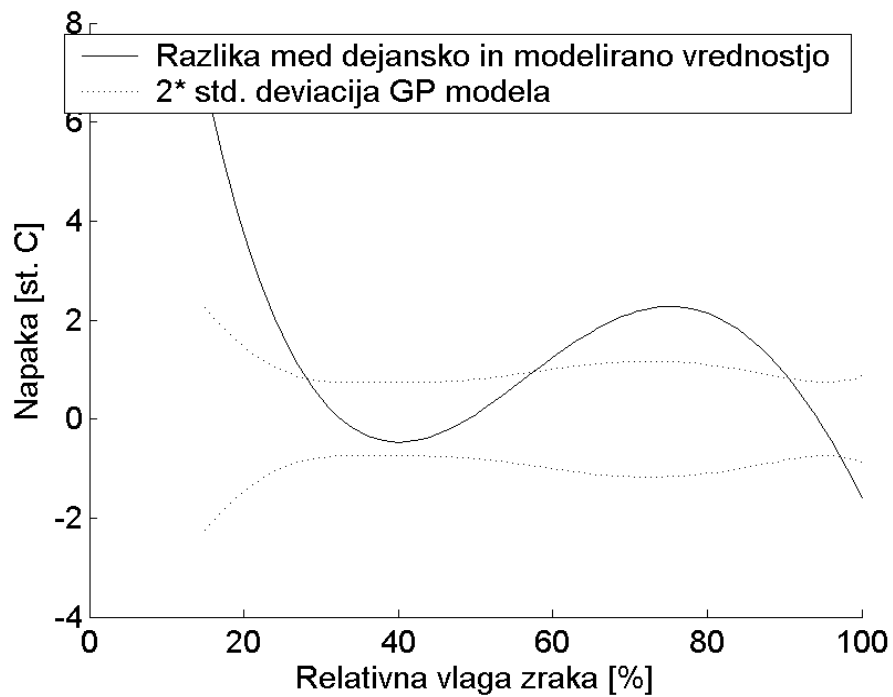
Slika 76: Prvi model temperature rosišča



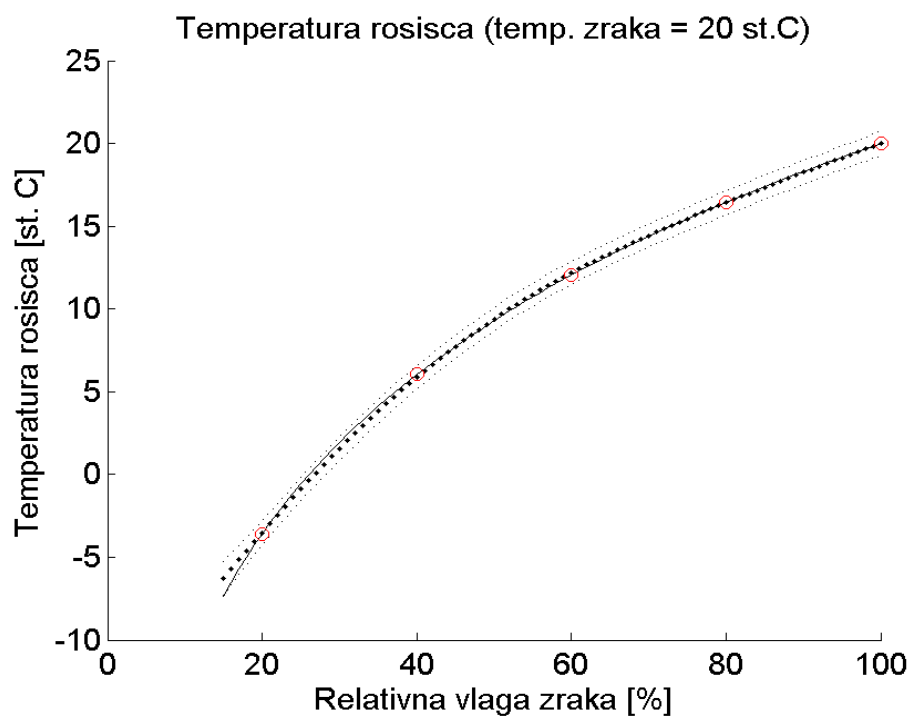
Slika 77: Napaka prvega modela temperature rosišča



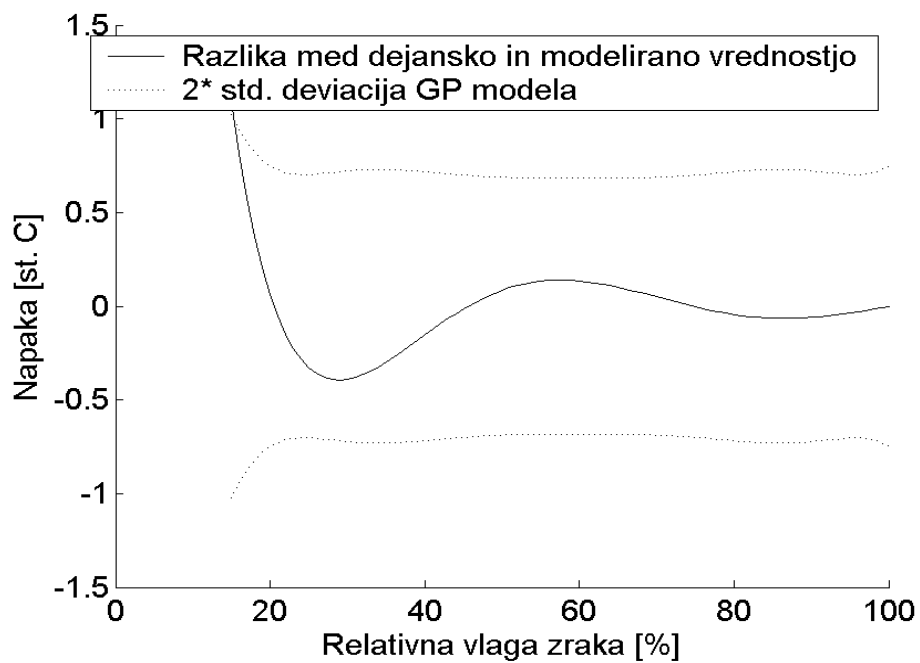
Slika 78: Drugi model temperature rosišča



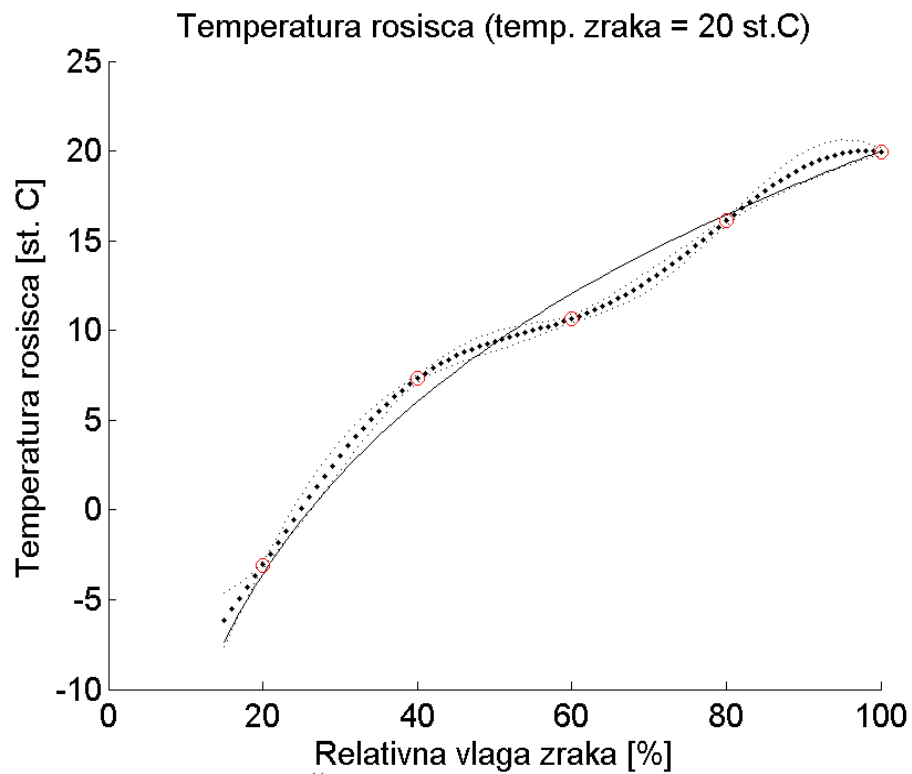
Slika 79: Napaka drugega modela temperature rosišča



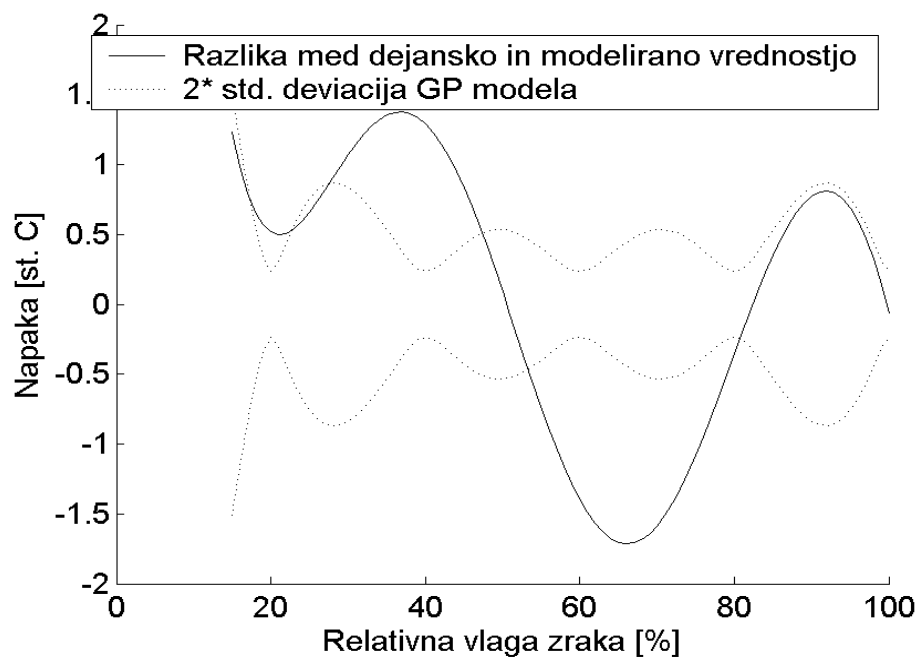
Slika 80: Tretji model temperature rosišča



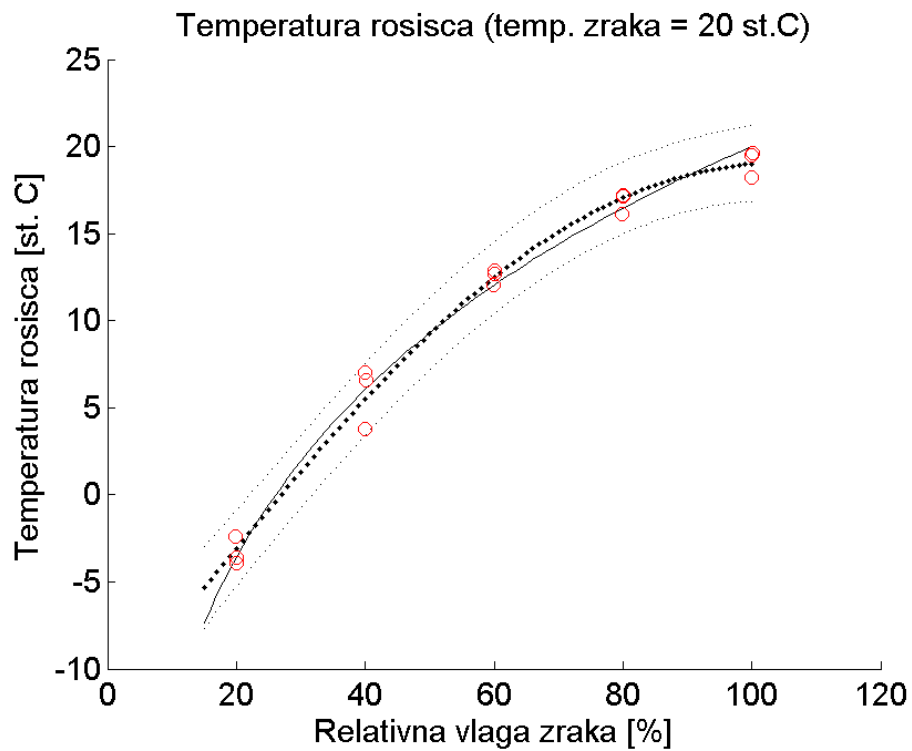
Slika 81: Napaka tretjega modela temperature rosišča



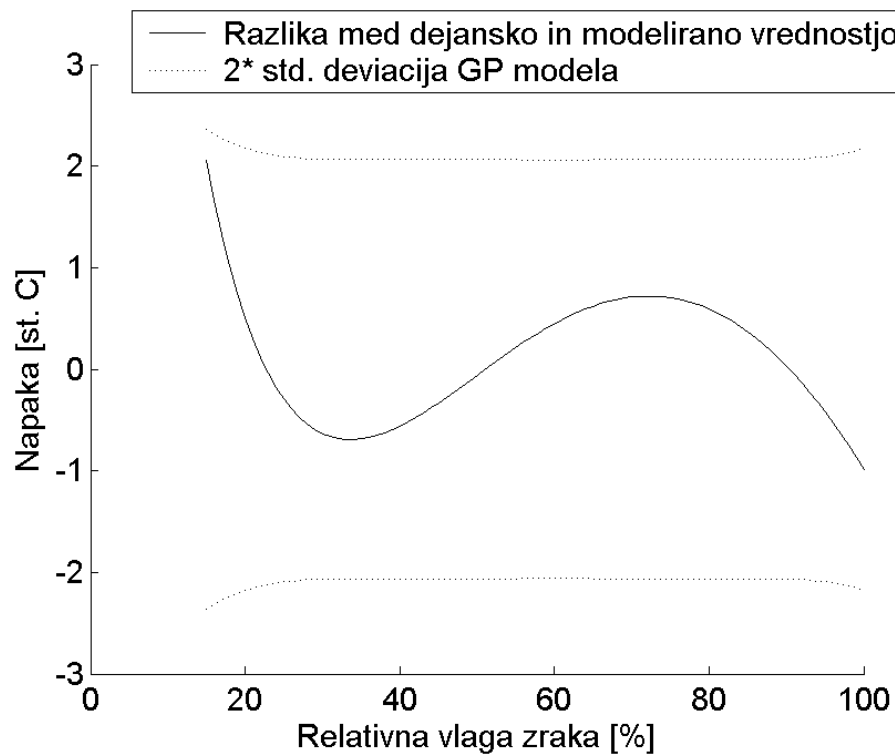
Slika 82: Četrty model temperature rosišča



Slika 83: Napaka četrtega modela temperature rosišča



Slika 84: Peti model temperature rosišča



Slika 85: Napaka petega modela temperature rosišča

B.2. Prikaz vpliva posameznih hiperparametrov modela

Model, ki je bil identificiran s pomočjo petih učnih točk, ki so bile izbrane na območju relativne vlage od 20% do 100% v enakomernih presledkih in je prikazan na slikah 80 in 81, sem uporabil za praktičen prikaz vpliva posameznih hiperparametrov v modelu. Za modeliranje je bila uporabljena oblika kovariančne funkcije

$$C_f(\mathbf{x}_m, \mathbf{x}_n; \Theta) = \Theta_{N+1} e^{-\frac{1}{2} \sum_{l=1}^N \frac{(x_m^{(l)} - x_n^{(l)})^2}{2\Theta_l^2}} + \delta_{mn} \Theta_{N+2}, \quad (\text{B.2})$$

kjer N označuje dimezijo vhodnega vektorja. V praktičnem primeru je bila dimezija $N=1$. To pomeni, da je bilo potrebno poiskati s pomočjo optimizacijskega algoritma, ki uporablja metodo največje podobnosti, optimalne vrednosti treh hiperparametrov. Po indentifikaciji modela so bile določene vrednosti hiperparametrov sledeče

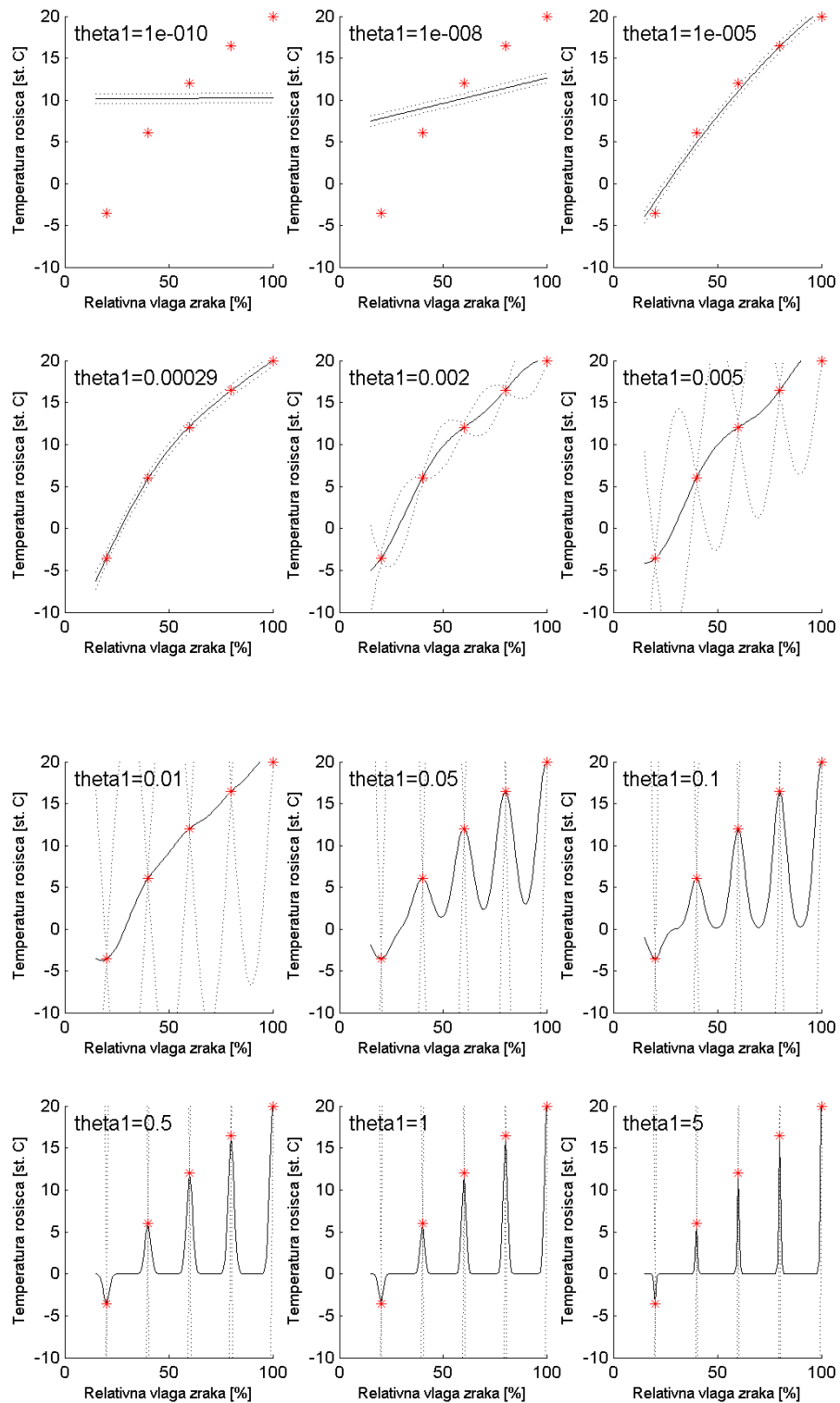
$$\Theta_1 = 0.00029, \quad \Theta_2 = 471, \quad \Theta_3 = 0.071. \quad (\text{B.3})$$

Za prikaz pomena posameznega hiperparametra sem napravil tri poizkuse. V vsakem poizkusu sem spreminjal vrednost enega hiperparametra, vrednosti preostalih dveh hiperparametrov pa se med poizkusom niso spreminjale in so bile nastavljene na optimalne vrednosti.

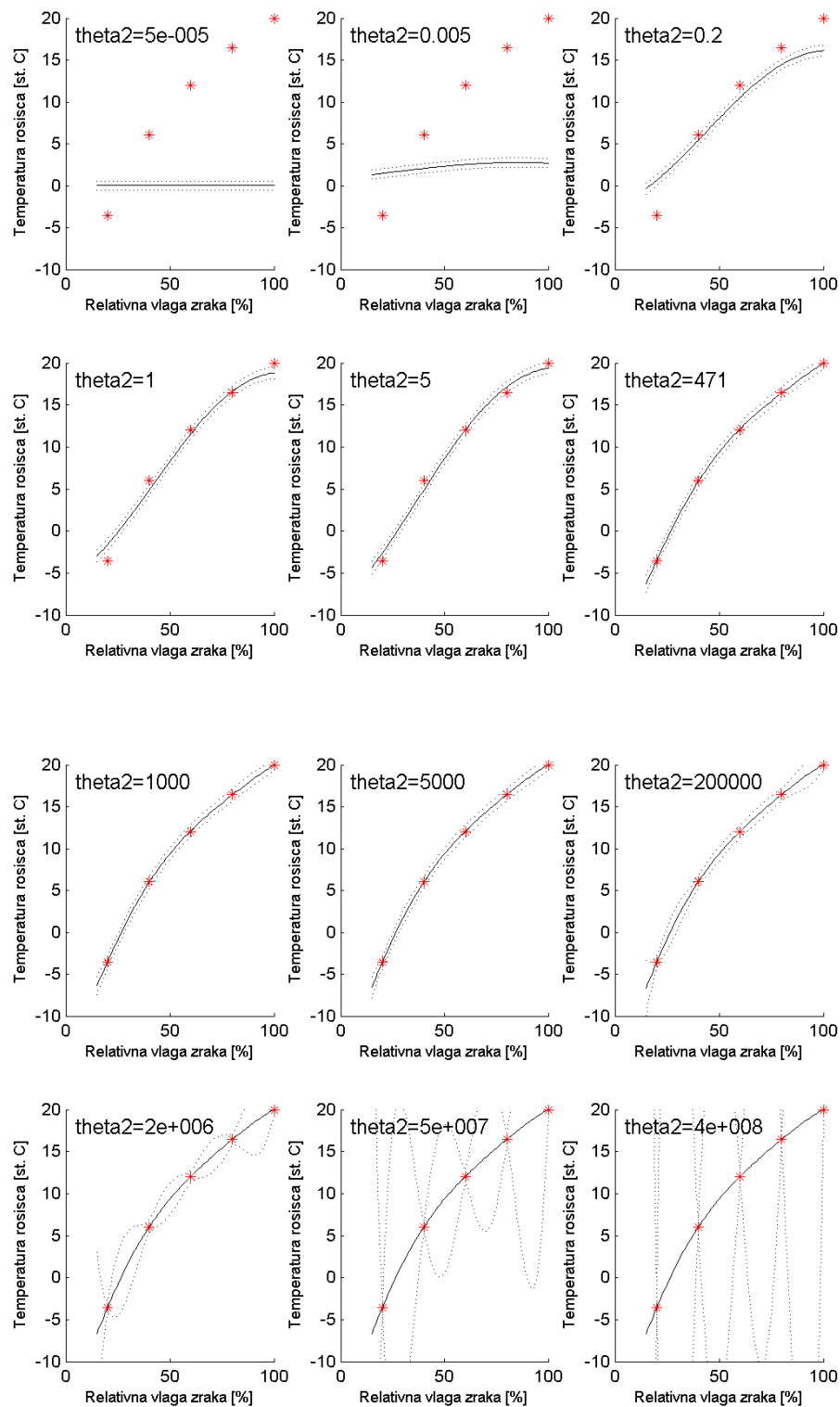
V prvem poizkusu, ki je prikazan na sliki 86 sem spreminjal vrednost prvega hiperparametra. Prvi hiperparameter Θ_1 predstavlja dolžinsko merilo dimezije. Iz slike je razvidno, da hiperparameter v bistvu določa področje (območje) vpliva posamezne učne točke. V primeru majhne vrednosti je območje zelo široko, kar v skrajnem primeru zelo majhne vrednosti pomeni, da bodo vse napovedane točke enake srednji vrednosti vseh učnih točk. Medtem, ko so v drugem skrajnem primeru zelo velike vrednosti hiperparametra napovedi z nizko varianco možne samo v neposredni bližini učnih točk.

V drugem poizkusu, ki je prikazan na sliki 87 sem spreminjal vrednost drugega hiperparametra. Drugi hiperparameter Θ_2 določa skupen vpliv vseh učnih točk na napovedane vrednosti. V skrajnem primeru zelo majhne vrednosti hiperparametra imajo učne točke skoraj ničel vpliv na napovedi (vse vrednosti napovedi so blizu konstantne vrednosti 0). V nasprotnem skrajnem primeru velike vrednosti hiperparametra pa imajo učne točke zelo velik vpliv na napovedane vrednosti v taki meri, da imajo napovedi v točkah, ki so malce oddaljene od učnih točk, že zelo veliko negotovost.

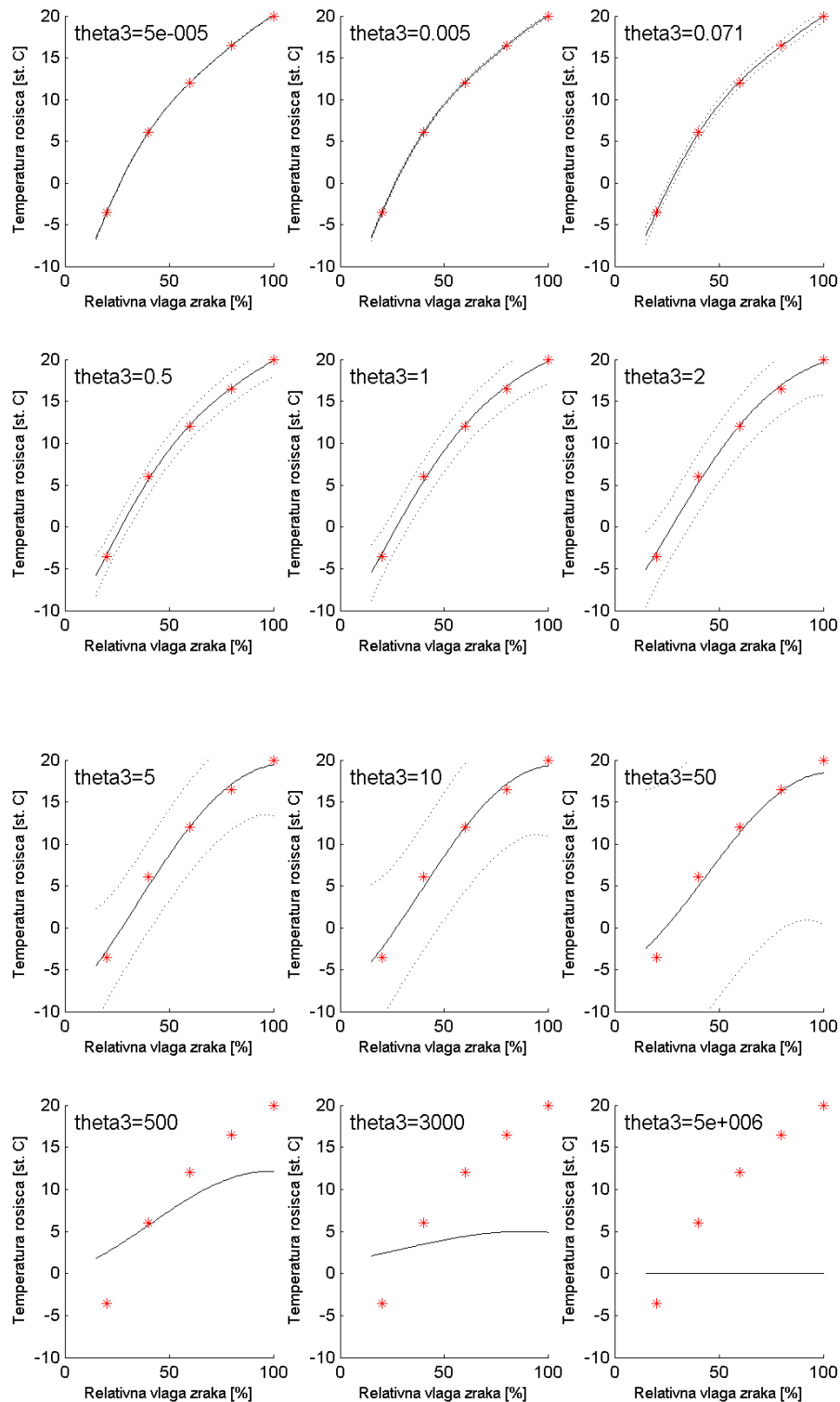
V tretjem poizkusu, ki je prikazan na sliki 88 sem spreminjal vrednost tretjega hiperparametra. Tretji hiperparameter θ_3 določa minimalno negotovost napovedanih točk, kar pomeni, da nobena napovedana vrednost ne more imeti manjše negotovosti. V skrajnem primeru zelo majhne vrednosti hiperparametra imajo tudi vse napovedi zelo majhno negotovost, v drugem skrajnem primeru velike vrednosti hiperparametra pa imajo vse napovedi izredno velike negotovosti, vrednosti napovedi pa blizu konstante vrednosti 0.



Slika 86: Vpliv prvega hiperparametra na model



Slika 87: Vpliv drugega hiperparametra na model



Slika 88: Vpliv tretjega hiperparametra na model

Dodatek C: Odvisnost negotovosti od kovariančne funkcije

Negotovost napovedi, ki je odvisna od oblike kovariančne funkcije, opisuje enačba

$$\sigma_{i_{N+1}}^2 = \kappa - \mathbf{k}_{N+1}^T \mathbf{C}_N^{-1} \mathbf{k}_{N+1} \quad , \quad (\text{C.1})$$

kjer je pomen oznake

$$\kappa = C(\mathbf{x}_{N+1}, \mathbf{x}_{N+1}; \Theta) \quad (\text{C.2})$$

ter oznake

$$\mathbf{k}_{N+1} = (C(\mathbf{x}_1, \mathbf{x}_{N+1}; \Theta), \dots, C(\mathbf{x}_N, \mathbf{x}_{N+1}; \Theta)) \quad . \quad (\text{C.3})$$

V nadaljevanju bom predstavil, kako se negotovost napovedi spreminja v odvisnosti od oddaljevanja iz področja učnih točk. V prvem primeru je bila za predstavitev uporabljena kovariančna funkcija brez modela šuma, v drugem primeru pa kovariančna funkcija z modelom šuma.

C.1. Uporaba stacionarne kovariančne funkcije brez upoštevanja šuma

Če uporabimo pri modeliranju z Gaussovimi procesi kovariančno funkcijo oblike

$$C_f(\mathbf{x}_m, \mathbf{x}_n; \Theta) = \Theta_1 e^{-\frac{1}{2} \sum_{i=1}^L \frac{(\mathbf{x}_m^{(i)} - \mathbf{x}_n^{(i)})^2}{2\tau_i^2}} \quad , \quad (\text{C.4})$$

ki ne upošteva šuma, se negotovost napovedi z oddaljevanjem iz področja učnih točk ekponentno povečuje od vrednosti 0 do neke končne vrednosti, ki je enaka Θ_1 .

V takšnem primeru stacionarne kovariančne funkcije je vrednost vseh diagonalnih elementov kovariančne matrike (elementi predstavljajo varianco posameznih učnih točk) enaka Θ_1 (saj je vrednost ekponentnega dela v enačbi (C.4) enaka 1, ker je oddaljenost med točkama enaka 0). Poleg tega pa je tudi vrednost κ enaka Θ_1 .

V enem skrajnem primeru, ko je vhod v model enak enemu izmed učnih vhodov (na vhod modela pripeljemo vrednost vhoda s katero smo ga pred tem učili), postane en element kovariančnega vektorja $\mathbf{k}_N$ enak varianci učne točke, ostali elementi pa so enaki kovariancam med učnimi točkami in napovedano točko

$$\mathbf{k}_N = [C(\mathbf{x}_1, \mathbf{x}_{N+1}; \Theta), \dots, \Theta_1, \dots, C(\mathbf{x}_N, \mathbf{x}_{N+1}; \Theta)] \quad . \quad (\text{C.5})$$

Vrednost omenjenega elementa je enaka Θ_1 (saj je vrednost ekponentnega dela v enačbi (C.4) enaka 1, ker je oddaljenost med točkama enaka 0). Kovariance med učnimi točkami in napovedano točko so manjše od Θ_1 in se ekponentno zmanjšujejo z oddaljenostjo od napovedovane točke.

Zaradi enakosti vhoda, za katerega želimo napoved, z enim od učnih vhodov, postane tudi vrednost izraza $\mathbf{k}_{N+1}^T \mathbf{C}_N^{-1} \mathbf{k}_{N+1}$ enaka Θ_1 . Iz tega pa sledi, da je vrednost negotovosti napovedi, kot ga opisuje enačba (C.1), enaka 0.

To je povsem razumljivo, saj model lahko v točki s katero je bil učen, s popolno gotovostjo napove izhod.

Vrednost izraza $\mathbf{k}_{N+1}^T \mathbf{C}_N^{-1} \mathbf{k}_{N+1}$ se z oddaljevanjem od učnih točk ekponentno zmanjšuje in doseže v neki veliki oddaljenosti končno vrednost, ki znaša 0. Zaradi tega pa negotovost eksponentno narašča proti vrednosti Θ_1 .

C.2. Uporaba stacionarne kovariančne funkcije z upoštevanjem šuma

Predpostavljajmo, da je šum naključen. Iz tega sledi, da ne pričakujemo nobene korelacije med šumom in določenimi izhodi ter, da drugi izraz vpliva samo na diagonalne elemente kovariančne matrike. Varianco šuma označimo z Θ_2 , za katero se predpostavlja, da je v primeru stacionarne kovariančne funkcije neodvisna od vhodov.

Z δ_{mn} pa je označena Kronecker-jeva delta funkcija

$$\delta_{mn} = \begin{cases} 0; & m \neq n \\ 1; & m = n \end{cases} \quad (C.6)$$

Če uporabimo pri modeliranju z Gaussovimi procesi kovariančno funkcijo oblike

$$C_f(\mathbf{x}_m, \mathbf{x}_n; \Theta) = \Theta_1 e^{-\frac{1}{2} \sum_{l=1}^L \frac{(x_m^{(l)} - x_n^{(l)})^2}{2r_l^2}} + \delta_{mn} \Theta_2, \quad (C.7)$$

ki upošteva šum, se negotovost napovedi z oddaljevanjem iz področja učnih točk ekponentno povečuje od vrednosti malce večje od Θ_2 do neke končne vrednosti, ki je enaka $\Theta_1 + \Theta_2$.

V takšnem primeru stacionarne kovariančne funkcije je vrednost vseh diagonalnih elementov kovariančne matrike (elementi predstavljajo varianco posameznih učnih točk) enaka $\Theta_1 + \Theta_2$ (saj je vrednost ekponentnega dela v enačbi (C.7) enaka 1, ker je oddaljenost med točkama enaka 0, vrednost Kronecker-jeve delta funkcije pa je enaka 1).

Poleg tega pa je vrednost κ v tem primeru enaka $\Theta_1 + \Theta_2$ (drugačna kot v prejšnjem primeru).

Zopet v enem skrajnem primeru, ko je vhod v model enak enemu izmed učnih vhodov (na vhod modela pripeljemo vrednost vhoda s katero smo ga pred tem učili), postane en element kovariančnega vektorja $\mathbf{k}_N$ enak varianci učne točke, ostali elementi pa so enaki kovariancam med učnimi točkami in napovedano točko

$$\mathbf{k}_N = [C(\mathbf{x}_1, \mathbf{x}_{N+1}; \Theta), \dots, \Theta_1, \dots, C(\mathbf{x}_N, \mathbf{x}_{N+1}; \Theta)] \quad (9.1)$$

Vrednost omenjenega elementa je enaka Θ_1 (saj je vrednost ekponentnega dela v enačbi (C.7) enaka 1, ker je oddaljenost med točkama enaka 0, vrednost Kronecker-jeve delta funkcije pa je enaka 0). Zaradi enakosti vhoda, za katerega želimo napoved, z enim od učnih vhodov, postane tudi vrednost izraza $\mathbf{k}_{N+1}^T \mathbf{C}_N^{-1} \mathbf{k}_{N+1}$ blizu vrednosti Θ_1 . To pomeni,

- da je v skrajnem primeru, ko nimamo šuma, izraz $\mathbf{k}_{N+1}^T \mathbf{C}_N^{-1} \mathbf{k}_{N+1}$ enak Θ_1 ,
- da je v primeru majhne vrednosti šuma Θ_2 izraz $\mathbf{k}_{N+1}^T \mathbf{C}_N^{-1} \mathbf{k}_{N+1}$ malce manjši od Θ_1 ,
- v primeru večje vrednosti šuma Θ_2 , pa postaja vrednost izraza $\mathbf{k}_{N+1}^T \mathbf{C}_N^{-1} \mathbf{k}_{N+1}$ čedalje manjša od Θ_1 .

Glede na enačbo (C.1) in dejstvo, da je κ v tem primeru enaka $\Theta_1 + \Theta_2$, sledi, da je vrednost negotovosti napovedi, kot ga opisuje enačba (C.1), za faktor, ki je funkcija parametrov Θ_1 in Θ_2 ter kovarianc med učnimi točkami, večja od Θ_2 .

Tudi to je povsem razumljivo, saj model v točki s katero je bil učen, ne more izdelati napovedi z manjšo negotovostjo, kot je negotovost zaradi šuma.

Vrednost izraza $\mathbf{k}_{N+1}^T \mathbf{C}_N^{-1} \mathbf{k}_{N+1}$ se z oddaljevanjem od učnih točk ekponentno zmanjšuje in doseže v neki veliki oddaljenosti končno vrednost, ki znaša 0. Zaradi tega pa negotovost eksponentno narašča proti vrednosti $\Theta_1 + \Theta_2$.

Izjava

Izjavljam, da sem magistrsko nalogo izdelal samostojno pod vodstvom prof. dr. Marka Muniha in prof. dr. Draga Matka. Izkazano pomoč ostalih sodelavcev sem v celoti navedel v zahvali.

Boštjan Grašič