

*Napovedovanje meteorološkega stanja vozišča iz  
preteklih podatkov in vremenskih napovedi*

Rok Kršmanc

DOKTORSKA DISERTACIJA

PREDANA

FAKULTETI ZA RAČUNALNIŠTVO IN INFORMATIKO

KOT DEL IZPOLNJEVANJA POGOJEV ZA PRIDOBITEV NAZIVA

DOKTOR ZNANOSTI

S PODROČJA

RAČUNALNIŠTVA IN INFORMATIKE



Ljubljana, 2013





*Napovedovanje meteorološkega stanja vozišča iz  
preteklih podatkov in vremenskih napovedi*

Rok Kršmanc

DOKTORSKA DISERTACIJA

PREDNA

FAKULTETI ZA RAČUNALNIŠTVO IN INFORMATIKO

KOT DEL IZPOLNJEVANJA POGOJEV ZA PRIDOBITEV NAZIVA

DOKTOR ZNANOSTI

S PODROČJA

RAČUNALNIŠTVA IN INFORMATIKE



Ljubljana, 2013



## IZJAVA O AVTORSTVU DOKTORSKE DISERTACIJE

*Izjavljam, da sem avtor doktorske disertacije z naslovom *Napovedovanje meteorološkega stanja vozišča iz preteklih podatkov in vremenskih napovedi*, ki sem jo izdelal samostojno pod vodstvom mentorja in da slednje ne vsebuje materiala, ki bi ga kdorkoli predhodno že objavil ali oddal v obravnavo za pridobitev naziva na univerzi ali na drugem visokošolskem zavodu, razen v primerih, kjer so navedeni viri. Elektronska oblika doktorske disertacije je identična s tiskano obliko doktorske disertacije. Soglašam z javno objavo elektronske doktorske disertacije.*

— Rok Kršmanc —

junij 2013

ODDAJO SO ODOBRLI

dr. Janez Demšar

*izredni profesor za področje računalništva in informatike*

MENTOR IN ČLAN OCENJEVALNE KOMISIJE

dr. Polona Oblak

*docentka za področje matematike*

PREDSEDNICA OCENJEVALNE KOMISIJE

dr. Tomaž Kramberger

*docent za področje računsko intenzivnih metod in aplikacij v logistiki*

ZUNANJI ČLAN OCENJEVALNE KOMISIJE

Univerza v Mariboru, Fakulteta za logistiko



## PREDHODNA OBJAVA

Izjavljam, da so bili rezultati obravnavane raziskave predhodno objavljeni/sprejeti za objavo v recenzirani reviji ali javno predstavljeni v naslednjih primerih:

- [1] R. Kršmanc, A. Šajn Slak, S. Čarman in M. Ivačič. Next steps in forecasting road surface temperature and developing MDSS. *19<sup>th</sup> International Symposium on Electronics in Transport (ISEP 2011)*, Ljubljana (Slovenija), marec 2011.
- [2] R. Kršmanc, A. Šajn Slak in J. Demšar. Statistical approach for forecasting road surface temperature. *Meteorological Applications*. doi: [10.1002/met.1305](https://doi.org/10.1002/met.1305)

Potrjujem, da sem pridobil pisna dovoljenja vseh lastnikov avtorskih pravic, ki mi dovoljujejo vključitev zgoraj navedenega materiala v pričujočo disertacijo. Potrjujem, da zgoraj navedeni material opisuje rezultate raziskav, izvedenih v času mojega diplomskega študija na Univerzi v Ljubljani.







Miroslav Kršmanec

\*1953 – †2012



Videl sem misleca: pisal je zákone  
ljudstvu ne zemskemu — zvezdam je kazal pot,  
pa nesoglasje v vsemirju zasledil je,  
novih svetov je zahteval njegov račun,  
“Bodi!” je rekel — in noč mu je dala nov svet ...  
Tu, tu se žile življenja stekajo,  
pota vesoljstva tukaj se sekajo,  
ljubim jih s hrupom in šumom, ta vélika mesta —  
skoznje v svobodo gre, skoznje v bodočnost gre cesta ...

— O. Župančič: Dúma



## POVZETEK

V disertaciji je predstavljena razvita metoda za gradnjo hierarhičnih regresijskih modelov, ki temelji na združevanju vnaprej podanih množic podatkov na podlagi znanja o združljivosti in/ali podobnosti modelov. ¶ Regresijski modeli so danes široko razširjeni in se uporabljajo na mnogih področjih. Z vidika robustnosti je najbolj primerno uporabiti linearne regresijske modele. Uporaba enega samega, splošnega linearnega regresijskega modela ni primerna na podatkih, pri katerih se mehanizmi, ki bi jih želeli z regresijskimi modeli opisati, spreminjajo. Na takšnih podatkih je primerneje uporabiti več linearnih modelov. Ker ob tem zaradi pomanjkanja podatkov narašča varianca, je smiselno modele, pri katerih so mehanizmi sorodni, združiti. Združeni modeli morajo ustrezno ločiti te mehanizme in morajo kljub njihovi poenostavljeni obliki (linearizacija) doseči zadovoljive rezultate.

Delovanje metode smo teoretično utemeljili z razcepom napake na varianco in pristranskost in pri tem pokazali, da pri združevanju z naraščanjem napake narašča pristranskost, varianca pa pada. V kolikor bi uporabili en sam model, bi imel ta veliko pristranskost, po drugi strani pa bi pri velikem številu ločenih modelov imeli veliko varianco. Na podlagi teh spoznanj smo razvili metodo za gradnjo hierarhičnih regresijskih modelov, ki iterativno združuje modele vsake podmnožice podatkov, dokler ni vpliv zmanjševanja variance (zaradi večjih množic podatkov združenih modelov) manjši od povečevanja pristranskosti (zaradi obravnavanja nelinearnih povezav med spremenljivkami kot linearne). Metodo smo preizkusili na umetno pripravljenih podatkih in potrdili ustreznost takšnega pristopa. ¶ Metodo smo preizkusili tudi na podatkih s področja cestne meteorologije in z njo uspešno napovedovali temperaturo cestišča za več ur vnaprej, napovedi pa smo primerjali z napovedmi fizikalnega modela. Pokazali smo, da je mogoče sestavljati dovolj natančne modele za napovedovanje temperature cestišča tudi zgolj s statističnimi metodami. Prednost takšnega modeliranja pred fizikalnimi

modeli je, da ne zahteva nastavljanja parametrov modela, slabost pa, da zahteva vnaprej zbrane podatke iz cestnovremenskih postaj na lokaciji napovedovanja.

Temu problemu bi se izognili, če bi lahko poiskali ustrezen model za lokacije, kjer nam podatki cestnovremenskih postaj niso na voljo, na podlagi podobnosti s preostalimi lokacijami, za katere pa imamo te podatke na voljo. Najprej smo se prepričali, da so izbrane lastnosti, ki opisujejo lokacije (primer takšne lastnosti je vidnost neba), povezane s koeficienti regresijskih modelov, zatem pa smo predstavili metodo, ki poišče povezave med modeli in lastnostmi objektov, na katere se ti modeli nanašajo. ¶ Metodo smo uporabili tako na umetnih podatkih kot tudi na problemu napovedovanja temperature cestišča. Pri slednjem smo pokazali, kako lahko za lokacije, pri katerih izmerjeni podatki iz cestnovremenskih postaj niso na voljo, sestavimo napovedne modele na podlagi ostalih podatkov, ki ustrezno opisujejo lastnosti lokacij cestnovremenskih postaj. Poizkusi kažejo, da imajo ti modeli v primerjavi z modeli, ki bi jih pripravili, če bi imeli na voljo vse podatke, manjšo, vendar še vedno zadovoljivo napovedno točnost.

Rezultati disertacije so lahko uporabni v praksi. Napovedi meteorološkega stanja cestišča predstavljajo dragocene informacije voznikom in cestnim vzdrževalnim službam pri sprejemanju odločitev, posebej pomembne pa so za zimsko vzdrževalno službo. Omogočajo ji zanesljivo predvidevanje nastanka poledice in s tem izvajanje učinkovitega preventivnega posipavanja, kar pomeni prihranek pri organizaciji virov in posipnih materialih ter zmanjšanje negativnih vplivov na okolje.

*Ključne besede:* regresijska analiza, napovedni modeli, dekompozicija napake, podobnosti, cestna meteorologija

## ABSTRACT

The dissertation presents a new method for construction of hierarchical regression models by combining the predetermined groups of data based on the knowledge of compatibility and/or similarity of models. ¶ Regression models are widely used in many areas. The simplest, linear regression models are often chosen because of their robustness. In this work we consider a case in which the data can be split into multiple subsets on which the predicted outcome is linear. While such data cannot be modelled with a single linear model, inducing a separate linear model for each subset would yield inaccurate models due to small samples.

This problem can be nicely framed in the theory of bias-variance decomposition of error. A single model would have an unacceptably high bias, while separate models would have high variances. Based on this theoretical justification, we developed a method for hierarchical construction of linear models, which starts with separate models for each subgroup and then iteratively merges them until the effect of decreased variance due to larger data samples available for each model begins to be overcome by the increased bias due to treating a non-linear relation as linear. We tested the method on controlled synthetic data, which proved the correctness of our approach. ¶ The method was then tested on the data on road meteorology: we were able to successfully predict the road surface temperature for several hours ahead. This result is interesting for the field of road meteorology as it shows that it is possible to construct models with good forecasting accuracy with statistical methods alone. The advantage of such modelling compared to physical models based on energy balance equation is that they do not require any knowledge about the road construction properties. Their weakness is that they require past data from road weather station at a particular location.

To overcome this problem, we investigated a problem in which the data for several locations is available and the task is to find a predictive model for a new location

based on the known physical, but easily obtainable similarities between this and other locations, such as sky visibility and similar. We first checked that the attributes that we chose to describe locations are indeed correlated with coefficients from regression models. Based on positive findings of this study, we defined a modelling technique that can construct a linear model for the location based on linear models for other locations. ¶ We again first empirically tested the method on synthetic data constructed in such a way that it fulfills the assumptions of the method, and then on the actual data from the road weather stations. As expected, the accuracy of such models is below those constructed from the actual data, yet still quite in the acceptable range for its potential practical use.

Results of the dissertation may find practical use. Forecasts of road weather conditions are a valuable resource for drivers, as well as for road maintenance services, in particular in winter. More accurate forecasts can provide safer roads while cutting down the maintenance costs and minimizing the environmental damage from over-salting.

*Key words:* regression analysis, prediction models, bias-variance decomposition, similarities, road meteorology



## ZAHVALA

*Cesta zahvale se začne pri mentorju. Imel sem že napisano besedilo, ki je opisovalo najino zanimivo sodelovanje, a sem na koncu ugotovil, da ga lahko nadomestim s stavkom: če bi bilo potrebno pripraviti priročnik za odlične mentorje, bi si morali ogledati najin primer. Janez, hvala!*

*Zaposlitev v podjetju CGS plus je botrovala k odločitvi, da je moja življenjska pot krenila na podiplomski študij. Na oddelku Okolje sem imel zelo prijetne pogoje za delo. Hvaležen sem svoji raziskovalni mentorici, dr. Alenki Šajn Slak, in sodelavcu Samu Čarmanu, ki sta me navdušila in uvedla v to raziskovalno področje. Rad bi omenil tudi oddelek GIS, zlasti Andreja Bedena, ki je vedno poskrbel, da so bili sistemi s podatki na voljo.*

*Za izvedbo te raziskave je bilo potrebno pripeljati veliko podatkov. Za uporabo podatkov meteoroloških modelov se zahvaljujem Agenciji Republike Slovenije za okolje (ARSO). Tu bi še posebej izpostavil Janka Meršeta, ki me je uvedel v meteorološke napovedne modele in mi pojasnil marsikatero tehnično vprašanje glede INCA sistema. Veliko znanja o slednjem sem lahko pridobil tudi od partnerjev mednarodnega projekta INCA–CE. Za uporabo podatkov iz cestnovremenskih postaj pa gre zahvala Direkciji Republike Slovenije za ceste (DRSC) in Družbi za avtoceste v Republiki Sloveniji (DARS).*

*Zahvala ne more mimo preostalih članov ocenjevalne komisije, doc. dr. Polone Oblak in doc. dr. Tomaža Krambergerja, za trud, ki sta ga vložila v pregled tega dela in za vse pripombe v okviru predstavitve raziskovalnega dela pri predmetu Seminar V.*

*Cesta zahvale se konča doma. Hvala staršem; mami in pokojnemu očetu. Vajina podpora mi je omogočila marsikaj. Posebna zahvala gre Tadeji, ki mi je v času študija stala ob strani in me spodbujala, ko mi je bilo težko.*

— Rok Kršmanc, Ljubljana, junij 2013.



# KAZALO

<i>Povzetek</i>	<i>i</i>
<i>Abstract</i>	<i>iii</i>
<i>Zahvala</i>	<i>v</i>
<b>1</b> <i>Uvod</i>	<b>1</b>
1.1 Motivacija . . . . .	2
1.2 Pregled vsebine . . . . .	3
1.3 Prispevki k znanosti . . . . .	5
<b>2</b> <i>Opis področja in preteklega dela</i>	<b>7</b>
2.1 Metodološko ozadje . . . . .	8
2.1.1 Regresijski modeli . . . . .	8
2.1.2 Dekompozicija napake . . . . .	9
2.1.3 Hierarhični regresijski modeli . . . . .	10
2.2 Cestna meteorologija . . . . .	11
2.2.1 Cestnovremenske postaje . . . . .	12
2.2.2 Meteorološki modeli . . . . .	17
2.2.3 Temperatura cestišča . . . . .	18
2.2.4 Napoved stanja cestišča . . . . .	22
2.2.5 Delo in optimizacija zimske službe v Sloveniji in drugod . . . . .	30
<b>3</b> <i>Opis podatkov</i>	<b>37</b>
3.1 Izmerjeni podatki in podatki vremenskih napovedi . . . . .	38
3.1.1 Podatki za preizkus gradnje hierarhičnih modelov . . . . .	38

3.1.2	Podatki za analizo povezav med modeli in lastnostmi objektov	40
3.2	Priprava umetnih podatkov . . . . .	45
4	<i>Opis dela</i>	47
4.1	Metoda za gradnjo hierarhičnih regresijskih modelov . . . . .	48
4.1.1	Opis metode združevanja . . . . .	49
4.1.2	Preizkus na umetnih podatkih . . . . .	51
4.1.3	Preizkus na problemu napovedovanja temperature cestišča . . . . .	53
4.2	Dekompozicija napake hierarhičnih regresijskih modelov . . . . .	63
4.2.1	Opis postopka dekompozicije . . . . .	64
4.2.2	Rezultati postopka dekompozicije . . . . .	66
4.3	Analiza povezav med modeli in lastnostmi objektov . . . . .	67
4.3.1	Opazovanje podobnosti med modeli in lastnostmi . . . . .	68
4.3.2	Metoda za analizo povezav . . . . .	77
4.3.3	Preizkus na umetnih podatkih . . . . .	79
4.3.4	Preizkus na problemu napovedovanja temperature cestišča . . . . .	80
5	<i>Zaključek</i>	85
5.1	Povzetek praktičnih prispevkov . . . . .	87
5.2	Nadaljnje delo . . . . .	88
A	<i>Programska koda za pripravo umetnih podatkov</i>	91
	<i>Slike</i>	95
	<i>Tabele</i>	96
	<i>Literatura</i>	99

*Uvod*

Regresijski modeli se uporabljajo na mnogih področjih zlasti za namene zgoščene prikaza podatkov, napovedovanja trenda in opazovanja povezav med spremenljivkami. Z vidika robustnosti je najbolj primerno uporabiti linearne regresijske modele, četudi povezave med spremenljivkami niso linearne. Uporaba enega samega, splošnega linearnega regresijskega modela ni primerna na podatkih, pri katerih se mehanizmi, ki bi jih želeli z regresijskimi modeli opisati, spreminjajo. Na takšnih podatkih je primerneje uporabiti več linearnih modelov.

V skladu z dekompozicijo napake na pristranskost in varianco ima uporaba več modelov manjšo pristranskost, vendar večjo varianco, saj smo za njihovo učenje uporabili manjše učne množice. Da bi imeli na voljo več podatkov in s tem zmanjšali varianco brez znatnega povečevanja pristranskosti, je smiselno podobne modele, pri katerih so mehanizmi sorodni, združiti. Postavlja se vprašanje, kako takšne modele združevati in pod kakšnimi pogoji. Združeni modeli morajo namreč ustrezno ločiti ključne mehanizme in morajo kljub njihovi poenostavljeni obliki (linearizacija) doseči zadovoljive rezultate.

Poleg iskanja postopka združevanja je zanimivo tudi raziskovanje modelov, ki se nanašajo na določene objekte. Ti modeli so opisani z različnimi spremenljivkami (na primer z regresijskimi koeficienti), prav tako pa imajo objekti, ki se na te modele nanašajo, določene lastnosti (atribute). Takšne povezave med spremenljivkami modelov in lastnostmi objektov niso eksplicitne, marveč se pojavlja možnost po njihovi analizi. Pri njihovem iskanju nas zanima, katere lastnosti objektov vplivajo na podobnost med modeli in na kakšen način. Za tovrstno analizo je potrebno natančno določiti, katere spremenljivke modelov želimo uporabiti in katere so tiste lastnosti objektov, ki bi utegnile predstavljati takšne povezave.

## 1.1 Motivacija

V obdobju od pozne jeseni do zgodnje pomladi se v mnogih državah soočajo z izzivi težkih vremenskih razmer, ki povzročajo povečanje števila nesreč in prometne zastoje, ki predstavljajo ekološke, finančne in druge posledice. Zimska služba mora ob zmanjševanju stroškov pri vzdrževanju cest omogočati visoko mobilnost in varnost prometa. Stroški zimskega vzdrževanja cest neprestano rastejo, kar je razumljivo, saj se povečujejo tako prometne obremenitve kot zahteve uporabnikov. ¶ Zimska služba poleg pluženja izvaja posipavanje, ki povzroča taljenje ledu in snega ter izboljša oprijem na cestišču. Ugotovitve kažejo, da je za odstranitev ledu s površine ceste potrebno tudi štirikrat in

več posipnih materialov kot pri izvajanju preventivnega posipavanja. Vsako nepotrebno posipavanje prinaša vrsto težav, kot so korozija vozil in mostov, onesnaženost vode in tal s soljo, zaviranje rasti dreves in grmičevja itd.

Zanesljiva napoved meteorološkega stanja na cestah (npr. zmrzal, led, sneg) predstavlja dragocene informacije voznikom in cestnim vzdrževalnim službam pri sprejemanju odločitev (svetlobni napisi, cestne zapore ipd.), posebej pomembna pa je za zimsko vzdrževalno službo. Omogoča ji zanesljivo predvidevanje nastanka poledice in s tem izvajanje učinkovitega preventivnega posipavanja, kar pomeni prihranek pri organizaciji virov in posipnih materialih in zmanjšanje negativnih vplivov na okolje. Temperatura cestišča je ključna pri odločitvah, katere posipne materiale je potrebno uporabiti in v kakšni količini. Takšne napovedi so še posebej pomembne na kritičnih odsekih cest (klanci, gorski prelazi, ozke soteske, viadukti, vetru izpostavljene ceste), kjer prihaja do ekstremnih zimskih razmer. ♣ Napovedovanje stanja cestišča ni mogoče le s poznavanjem klasične vremenske napovedi, saj je veliko dejavnikov, ki vplivajo na temperaturo in stanje cestišča in so medsebojno nelinearno odvisni. Najpomembnejši so temperatura zraka, padavine in sevalni tokovi. V svetu so razviti številni modeli, ki na podlagi podatkov iz cestnovremenskih postaj in podatkov obstoječih meteoroloških modelov omogočajo napovedovanje temperature in stanja cestišča. Kljub nenehnemu razvoju je napaka teh modelov v nekaterih primerih še vedno prevelika, zlasti na področjih z razgibano topografijo in gostim prometom. Vsakršna izboljšanja obstoječih napovedi ali pa napovedovanje z drugačnimi metodami, ki bi dosegale boljše rezultate, predstavljajo potreben in pomemben doprinos k boljšim odločitvam na področju zimskega vzdrževanja cest.

## 1.2 Pregled vsebine

Uvodnemu poglavju doktorske disertacije sledi poglavje, ki podaja kratek opis metodološkega ozadja nekaterih področij statistike in strojnega učenja, ki so potrebna podlaga za razumevanje vsebine doktorske disertacije. Predstavljeni so široko razširjeni regresijski modeli v linearni in posplošeni (nelinearni) obliki, teorija dekompozicije napake na pristranskost in varianco, ki se uporablja za proučevanje algoritmov, in postopki za vpeljavo hierarhije v regresijske modele. Sledi bolj obširen opis področja cestne meteorologije, ki je tesno povezan z naslovom doktorske disertacije. Bralcu omogoča seznanitev z manj znanim področjem in opisuje cestnovremenske postaje (CVP), ki omogočajo spremljanje trenutnega stanja cestišča in meteoroloških spremenljivk na

lokaciji CVP, meteorološke modele, ki so za napovedovanje stanja cestišča ključni, dejavnike, ki vplivajo na temperaturo cestišča (npr. topografija, osončenost, vidnost neba, konstrukcija ceste), dosedanje pristope napovedovanja meteorološkega stanja cestišča (termalno kartiranje, fizikalni modeli, statistični modeli, prostorske napovedi), delo in optimizacijo zimske službe.

Tretje poglavje podaja podrobnejši opis podatkov, ki so v tej nalogi uporabljeni. Ti se v osnovi delijo na izmerjene oziroma izračunane podatke iz CVP in meteoroloških modelov, ter na podatke, ki smo jih pripravili umetno. Obe vrsti podatkov omogočata, da se postopke, predstavljene v tej nalogi, preizkusi tako iz teoretičnega kot iz aplikativnega vidika. Opisani so tudi podatki o lokacijah CVP, na katerih modeliramo temperaturo cestišča in so potrebni pri analizi povezav med modeli in lastnostmi objektov, na katere se ti modeli nanašajo.

Četrto poglavje predstavlja osrednji del doktorske disertacije in je razdeljeno na tri glavne razdelke. Prvi razdelek opisuje metodo za gradnjo hierarhičnih regresijskih modelov, ki temelji na združevanju vnaprej podanih množic podatkov in njeno uporabo na umetnih podatkih oziroma na problemu napovedovanja temperature cestišča. Pri slednjem je navedena napaka združenih modelov, s katerimi smo napovedovali temperaturo cestišča. To smo primerjali z napako povsem združenih modelov in z napako fizikalnega modela METRo. Drugi razdelek opisuje dekompozicijo napake na pristranskost in varianco v prejšnjem razdelku predstavljene metode. Opisana je prilagoditev postopka dekompozicije za predlagano metodo in njen preizkus na aplikativnih podatkih. Tretji razdelek analizira povezave med modeli in lastnostmi objektov, na katere se ti modeli nanašajo. Na problemu napovedovanja temperature cestišča smo z obstoječimi postopki poizkušali te povezave nakazati. Sledi opis metode za analizo povezav, ki predstavlja nadgradnjo metode iz prvega razdelka. Metodo smo uporabili na umetnih podatkih in poizkušali poiskati matriko podobnosti ter jo primerjati z matriko, ki je bila uporabljena pri generiranju podatkov. Metodo smo na drugačen način uporabili tudi na aplikativnih podatkih, kjer smo iskali koeficiente modelov iz lastnosti teh modelov in primerjali rezultate napovedi s poiskanimi koeficienti z napovedmi, kjer smo koeficiente izračunali iz podatkov teh modelov. To smo naredili z ločenimi dnevnimi in nočnimi modeli za vsako CVP, poleg tega pa smo te rezultate primerjali še z modeli, ki so bili skupni vsem CVP.

V petem poglavju zaključimo z opisom opravljenega dela, dobljenimi rezultati in sklepi. Predstavljeni so še predlogi za nadaljnje delo.



### 1.3 Prispevki k znanosti

V disertaciji so predstavljeni naslednji prispevki k znanosti:

- Razvoj metode za gradnjo hierarhičnih regresijskih modelov, ki temelji na združevanju vnaprej podanih množic podatkov na podlagi znanja o združljivosti modelov in/ali podobnosti modelov. Metoda je preizkušena tako na umetnih podatkih kot tudi na podatkih s področja cestne meteorologije (napovedovanje temperature cestišča).
- Prilagoditev postopka za empirično dekompozicijo in analizo razmerja napake na pristranskost in varianco. Postopek je mogoče uporabiti za teoretični preizkus delovanja metode, razvite v okviru prve točke.
- Razvoj oziroma prilagoditev metode za analizo povezav med regresijskimi modeli in lastnosti objektov, na katere se ti modeli nanašajo. Metoda je preizkušena tako na umetnih podatkih kot tudi na podatkih s področja cestne meteorologije (napovedovanje temperature cestišča).

Deli disertacije so bili predstavljeni na mednarodni konferenci *International Symposium on Electronics in Transport* [63] in v znanstveni reviji *Meteorological Applications* [65]. Poleg znanstvenega prispevka ima disertacija tudi praktično vrednost, saj so metode, ki smo jih razvili, lahko neposredno uporabne tako na področju cestne meteorologije pri napovedovanju temperature cestišča, kot tudi na ostalih problemih, ki so temu podobni. Na področju cestne meteorologije je uporaba razvitih metod v splošnem prenosljiva na različna geografska območja, saj so osnovni procesi, ki se odvijajo v ozadju, prostorsko neodvisni.



*Opis področja  
in preteklega dela*

Poglavje podaja kratek opis metodološkega ozadja nekaterih področij statistike in strojnega učenja, ki so potrebna podlaga za razumevanje vsebine doktorske disertacije. Sledi bolj obširen opis manj znanega področja cestne meteorologije.

## 2.1 Metodološko ozadje

V nadaljevanju so predstavljeni široko razširjeni regresijski modeli v linearni in posplošeni (nelinearni) obliki, teorija dekompozicije napake na pristranskost in varianco, ki se uporablja za proučevanje algoritmov, in postopki za vpeljavo hierarhije v regresijske modele.

### 2.1.1 Regresijski modeli

Regresijski modeli [36] se uporabljajo na mnogih področjih za namene zgoščenega prikaza podatkov, napovedovanja trenda v podatkih in opazovanja povezav med spremenljivkami. Klasični linearni regresijski model lahko v matričnem zapisu podamo kot:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad (2.1)$$

kjer je  $\mathbf{y}$  vektor realizacij odvisne spremenljivke velikosti  $n \times 1$ ,  $\mathbf{X}$  matrika neodvisnih spremenljivk velikosti  $n \times p$ ,  $\boldsymbol{\beta}$  vektor koeficientov velikosti  $p \times 1$  (ta vsebuje tudi koeficiente prostega člena  $\alpha$ , ki ga zaradi tega ne zapisujemo posebej) in  $\mathbf{e}$  vektor napak (za napake predvidevamo, da so neodvisne in normalno porazdeljene s srednjo vrednostjo 0 in standardno deviacijo  $\rho$ ) velikosti  $n \times 1$ ;  $p$  je število spremenljivk in  $n$  število podatkov.  $\boldsymbol{\beta}$ , ki minimizira vsoto kvadratov napake:

$$\sum_{i=1}^n (y_i - X_i \hat{\boldsymbol{\beta}})^2, \quad (2.2)$$

se izračuna kot:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}, \quad (2.3)$$

kjer ' pomeni operacijo transponiranja matrike. Doseganje minimalne vsote kvadratov napake je pomembna zaradi prilagoditve regresijskega modela na podatke [2].

Posplošeni linearni regresijski modeli [70] predstavljajo razširitev linearnih regresijskih modelov. Pri posplošenih linearnih regresijskih modelih ni potrebno, da je

$\mathbf{y}$  linearno odvisen od  $\mathbf{X}$ . Povezavo med neodvisnimi in odvisnimi spremenljivkami vzpostavlja inverzna funkcija  $g^{-1}$ :

$$\mathbf{y} = g^{-1}(\mathbf{X}\boldsymbol{\beta}) + \mathbf{e}. \quad (2.4)$$

Tudi linearnost v regresiji lahko zamenjamo z nelinearno funkcijo [3], kjer odziv neodvisnih spremenljivk določa poljubna funkcija  $f$ :

$$\mathbf{y} = f(\mathbf{X}, \boldsymbol{\beta}) + \mathbf{e}. \quad (2.5)$$

Primer nelinearnega regresijskega modela z velikim številom parametrov so tudi nevronske mreže [45], s katerimi lahko opišemo poljubno gladko funkcijo.

### 2.1.2 Dekompozicija napake

Dekompozicija napake na pristranskost in varianco je široko uporabljena metoda za proučevanje algoritmov v strojnem učenju. Kot najbolj razširjena funkcija izgube se uporablja srednja kvadratična napaka (v nadaljevanju  $MS$  napaka). Razvite so bile tudi metode za druge funkcije izgube, denimo 0-1 pri obravnavanju klasifikacijskih problemov [57, 59], medtem ko je Domingos [31] predlagal splošno dekompozicijo za poljubno funkcijo izgube. Praktične napotke pri uporabi dekompozicije navaja Bouckaert [14].

$MS$  napaka je definirana kot pričakovana vrednost kvadrata razlike med pravimi vrednostmi  $y$  in izračunanimi (napovedanimi) vrednostmi  $\hat{y}$ :

$$MS(\hat{y}) = E[(\hat{y} - y)^2]. \quad (2.6)$$

Za dekompozicijo  $MS(\hat{y})$  napišemo:

$$\begin{aligned} MS(\hat{y}) &= \\ &= E[(\hat{y} - y)^2] \\ &= E[\hat{y}^2 - 2\hat{y}y + y^2] \\ &= E[\hat{y}^2] - 2E[\hat{y}]E[y] + E[y^2]. \end{aligned} \quad (2.7)$$

Ker velja zveza  $E[x^2] = E[(x - E[x])^2] + E[x]^2$ , nadaljujemo:

$$\begin{aligned}
 MS(\hat{y}) &= \\
 &= E[(\hat{y} - E[\hat{y}])^2] + E[\hat{y}]^2 - 2E[\hat{y}]E[y] + E[(y - E[y])^2] + E[y]^2 \\
 &= E[(\hat{y} - E[\hat{y}])^2] + \underbrace{(E[\hat{y} - y])^2}_{\text{pristranskost}^2} + \underbrace{E[(y - E[y])^2]}_{\text{šum}} \\
 &= \text{varianca} + \text{pristranskost}^2 + \text{šum}.
 \end{aligned} \tag{2.8}$$

Varianca opisuje delež napake, ki ga lahko pripišemo naključnosti pri izboru učne množice. Pristranskost opisuje povprečno napako  $\hat{y}$ , ki izvira iz izbora metode modeliranja. Šum je lastnost podatkov; nanj nimamo vpliva in ga ne moremo odstraniti, medtem ko sta pristranskost in varianca odvisni od lastnosti algoritma, ki ga proučujemo.

### 2.1.3 Hierarhični regresijski modeli

Pri hierarhičnih regresijskih modelih [39, 40] se prosti člen in/ali koeficienti neodvisnih spremenljivk lahko spreminjajo (slika 2.1). Na hierarhično regresijo lahko gledamo kot na regresijo, kjer ima vsaka posamezna skupina podatkov svoj nabor neodvisnih spremenljivk in prostih členov, ki jih ravno tako modeliramo z regresijo. Hierarhični regresijski model zapišemo kot:

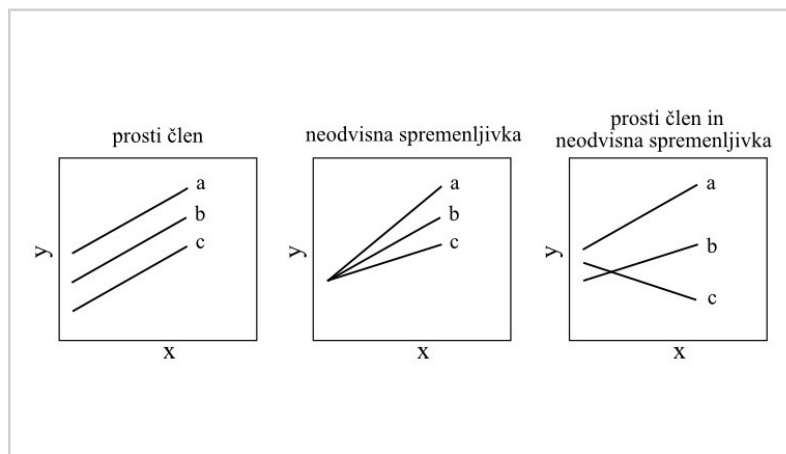
$$y = \alpha_i + X\beta_i + e. \tag{2.9}$$

Koeficiente prostega člena smo v razdelku 2.1.1 predvideli v vektorju koeficientov  $\beta$ , sedaj pa so zapisani posebej kot  $\alpha_i$ . Enačba 2.9 predstavlja prvi nivo regresijskega modeliranja, medtem ko se na drugem nivoju ukvarjamo z regresijskim modeliranjem koeficientov prostega člena ( $\alpha_i$ ) in/ali posameznih koeficientov neodvisnih spremenljivk ( $\beta_i$ ):

$$\alpha_i = X_\alpha \gamma + e_\alpha, \tag{2.10}$$

$$\beta_i = X_\beta \delta + e_\beta, \tag{2.11}$$

kjer sta  $\gamma$  in  $\delta$  vektorja koeficientov neodvisnih spremenljivk (vključno s prostim členom),  $X_\alpha$  in  $X_\beta$  matriki neodvisnih spremenljivk drugega nivoja regresije in  $e_\alpha$  in  $e_\beta$  vektorja napak tega nivoja. Na ta način ima lahko vsak nivo modeliranja svoje neodvisne spremenljivke s koeficienti. Hierarhični regresijski modeli sicer predstavljajo



Slika 2.1

Linearni regresijski modeli z eno neodvisno spremenljivko  $X$ . Shema na levi strani slike prikazuje tri regresijske modele z različnimi prostimi členi  $\alpha_i$  in konstantnim koeficientom  $\beta$  neodvisne spremenljivke  $X$  ( $y = \alpha_i + X\beta$ ). Shema v sredini slike prikazuje tri regresijske modele z različnimi koeficienti  $\beta_i$  neodvisne spremenljivke  $X$  in konstantnim prostim členom  $\alpha$  ( $y = \alpha + X\beta_i$ ). Shema na desni strani slike prikazuje tri regresijske modele z različnimi prostimi členi  $\alpha_i$  in z različnimi koeficienti  $\beta_i$  neodvisne spremenljivke  $X$  ( $y = \alpha_i + X\beta_i$ ).

datno kompleksnost, saj je potrebno regresijske modele vsakega nivoja izračunati posebej, vendar pa je veliko problemov takšnih, da je njihova uporaba smiselna.

Druga možnost za vzpostavitev hierarhije v regresijskih modelih so regresijska drevesa [18]. Linearni regresijski model nam predstavlja globalni model, s katerim zajamemo vse podatke. Ko imamo opravka z nelinearnimi podatki lahko bodisi uporabimo nelinearno regresijo bodisi podatke razdelimo v območja, ki jih z linearno regresijo opišemo bolj natančno. Postopek rekurzivno ponavljamo, dokler ne pridemo do razdelkov, kjer podatke že zelo dobro opišemo zgolj s preprostimi linearnimi modeli. Pri tem se za rekurzivna območja uporabi struktura drevesa, zato imenujemo postopek regresijska drevesa. Vsak list v drevesu predstavlja svojo skupino podatkov in mu pripada preprost linearen regresijski model, ki ustreza samo tej skupini podatkov.

## 2.2 Cestna meteorologija

Sledi bolj obširen opis področja cestne meteorologije, ki je tesno povezan z naslovom doktorske disertacije. Bralcu omogoča seznanitev z manj znanim področjem in opisuje cestnovremenske postaje, meteorološke modele, dejavnike, ki vplivajo na temperaturo cestišča, dosedanje rešitve napovedovanja meteorološkega stanja cestišča (s poudarkom na fizikalnem modelu METRo, ki ga bomo v poglavju 4 uporabili za primerjavo z rezultati), delo in optimizacijo zimske službe. Zlasti slednje poizkuša podati motivacijo

za reševanje problema napovedovanja stanja cestišča, s katerim se že desetletja sooča cestna stroka.

### 2.2.1 Cestnovremenske postaje

Cestnovremenske postaje (CVP) omogočajo spremljanje trenutnega stanja cestišča in meteoroloških spremenljivk na lokaciji CVP (slika 2.2). Na slovenskem cestnem omrežju je trenutno vgrajenih približno 100 CVP različnih proizvajalcev merilne opreme (zlasti Boschung, Vaisala in Lufft). Podatki iz CVP so namenjeni predvsem kot pomoč vzdrževalnemu osebju za pravočasno, učinkovito in racionalno izvajanje zimske službe, obenem pa se podatki uporabijo tudi kot pomoč pri vodenju prometa in opozarjanju voznikov o razmerah na cestah.

Najpomembnejši merilnik CVP je cestni senzor, vgrajen v cestno telo, ki meri temperaturo asfalta (na površini in različnih globinah), debelino vodnega filma in koncentracijo soli ter poda temperaturo zmrzišča. Cestni senzor predstavlja relativno majhen vzorec cestišča, zato je mikrolokacija cestnega merilnika zelo pomembna. Le s pravilno izbiro mikrolokacije namreč dobimo reprezentativni vzorec cestnega odseka in so izmerki, ki jih merilnik posreduje, verodostojni in uporabni za zimsko službo. Zadnjih nekaj let so na trgu dosegljivi tudi brezkontaktni merilniki in na voljo je več poročil o njihovem testiranju [58, 87, 100]. Glavna pomanjkljivost brezkontaktnih merilnikov je, da trenutno še ne merijo koncentracije posipnih materialov. CVP so navadno opremljene tudi z merilniki meteoroloških spremenljivk.

Glede na predlog standarda [35] naj bi CVP izvajale naslednje meritve:

- stanje cestišča (navedenih je 5 osnovnih stanj),
- debelina vodnega filma,
- temperatura zraka,
- temperatura rosišča,
- temperatura zmrzišča,
- temperatura cestišča,
- relativna vlažnost zraka,





*Slika 2.2*

Cestnovremenska postaja  
Moste z vgrajenimi in  
brezkontaktnimi merilniki.

- čas zaznavanja padavin,
- tip padavin (tekoče, trdne),
- intenziteta padavin,
- debelina snežne odeje,
- hitrost vetra,
- sunki vetra,
- smer vetra,
- vidljivost.

Slovenske CVP po opremljenosti deloma sledijo omenjenemu predlogu standarda, vendar se med seboj razlikujejo. Merilniki vetra, na primer, v večini primerov niso

nameščeni na CVP na državnih cestah. Po drugi strani pa predlog standarda ne navaja nekaterih koristnih meritev, kot so meritve kratkovalovnega in dolgovalovnega sevanja, koncentracije soli na cestišču ter zračnega tlaka. Slednja meritev je pomembna zlasti pri uporabi fizikalnih modelov za napovedovanje stanja cestišča.

Izbira lokacije CVP je odvisna od potrebe po zagotavljanju cestne varnosti na določenih odsekih ter objektih. Izkušnje kažejo, da se objekti kot so viadukti in mostovi hitreje ohlajajo in so zato veliko bolj izpostavljeni pojavu poledice od ostalih delov cestišč. Na podlagi tega dejstva se izvajajo postavitve in izbire lokacij CVP, pri tem pa je v dodatno pomoč mnenje lokalnih vzdrževalcev cest, ki na podlagi pogostih pojavov nevarnih cestnih razmer predlagajo lokacijo postavitve. Število postaj in razmiki med posameznimi postajami so odvisni od več različnih dejavnikov: topografija, struktura cestnega telesa, okolica, klasifikacija ceste. V splošnem lahko rečemo, da večja kot je raznolikost teh dejavnikov, večje je število potrebnih postaj. Zaradi visoke cene postaj in vzpostavitve pa je potrebno njihovo število minimizirati. Za določitev ustreznih lokacij CVP se v svetu uporablja tehnika termalnega kartiranja<sup>1</sup> in uporaba geografskih informacijskih sistemov [34].

Za zagotavljanje kakovosti in verodostojnosti izmerjenih podatkov iz CVP se izvajajo redna letna vzdrževanja v jesenskem času pred pričetkom zimske sezone ter v pomladnih mesecih po izteku zimske sezone. Taka vzdrževanja obsegajo predvsem redno čiščenje senzorjev in celotnega sistema, vključujejo pa tudi kalibracije senzorjev.

V nadaljevanju so opisani cestni merilniki, ki so v Sloveniji najbolj pogosti.

*Vaisala DRS511 (slika 2.3) [103].* Merilnik je namenjen vgradnji v površino cestišča. Izvaja naslednje meritve:

- temperatura cestišča,
- temperatura zmrzišča,
- stanje cestišča (suho, vlažno, mokro, mokro in soljeno, ivje, sneg, led),
- debelina vodnega filma,
- koncentracija slanosti cestišča (enota v  $g/L$  ter v  $g/m^2$ ).

---

<sup>1</sup>Prim. razdelek 2.2.4 na strani 22.



*Slika 2.3*

Cestni merilnik Vaisala DRS511.

*Boschung BOSO II (slika 2.4).* Merilnik je namenjen vgradnji v površino cestišča. Uporablja se skupaj z merilnikom ARCTIS, ki omogoča natančnejše določanje slanosti cestišča. Izvaja naslednje meritve:

- temperatura cestišča,
- stanje cestišča (suho, vlažno, mokro, mokro in soljeno, ivje, sneg, led),
- slanost cestišča (v %).



*Slika 2.4*

Cestni merilnik Boschung BOSO II.

*Lufft IRS21 (slika 2.5) [55].* Merilnik je namenjen vgradnji v površino cestišča. Izvaja naslednje meritve:

- temperatura cestišča,
- debelina vodnega filma,

- temperatura zmrzišča za NaCl,
- stanje cestišča (suho, vlažno, mokro, led, sneg, zmrzal).



Slika 2.5

Cestni merilnik Lufft  
IRS21.

*Vaisala DST111 (slika 2.6) [102].* Je merilnik za brezkontaktno merjenje temperature in vlažnosti in predstavlja alternativno možnost merjenja temperature cestišča. Merilnik odpravlja poseg v cestišče in omogoča montažo ob cesti z usmeritvijo na zeleno lokacijo na cestišču. Izvaja naslednje meritve:

- temperatura cestišča,
- temperatura zraka,
- relativna zračna vlažnost.

*Vaisala DSC111 (slika 2.6) [101].* Je merilnik za brezkontaktno merjenje stanja cestišča in oprijema. Predstavlja alternativno možnost merjenja stanja cestišča. Merilnik odpravlja poseg v cestišče in omogoča montažo ob cesti z usmeritvijo na zeleno lokacijo na cestišču. Ob namestitvi dodatne programske opreme omogoča določitev vidljivosti v razdalji do 2000 m.

*Slika 2.6*

Cestna merilnika Vaisala  
DST111 (levo) in DSC111  
(desno).

### 2.2.2 Meteorološki modeli

Ključni podatki za vse specialne modele za napovedovanje stanja cestišča so meteorološke napovedi. Za napovedovanje meteoroloških razmer se uporabljajo različni modeli glede na časovno skalo:

- Globalni meteorološki modeli za srednjeročno napovedovanje vremena (3–5 dni).
- Meteorološki modeli nad nekim območjem (npr. ALADIN/SI) za kratkoročne napovedi (1–3 dni). Ločljivost je v primerjavi z globalnimi modeli že bistveno izboljšana.
- Meteorološki modeli za zelo kratkoročne napovedi (do 12 ur vnaprej), ki so za modele napovedovanja stanja cestišča najpomembnejši.

Eden izmed sistemov za izračun zelo kratkoročne meteorološke napovedi v visoki prostorski in časovni ločljivosti je INCA<sup>2</sup> [50], ki ga uporablja tudi Državna meteorološka služba RS. INCA uporablja kot prvi približek stanja v atmosferi prostorska polja

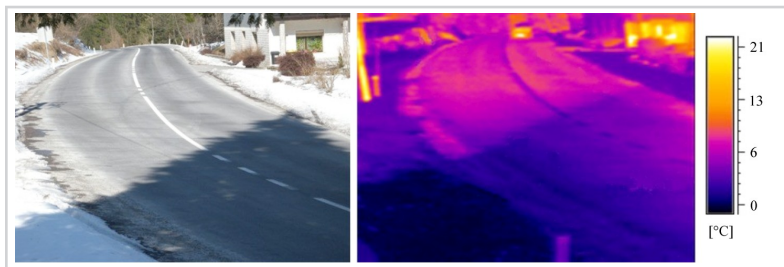
<sup>2</sup>INCA — Integrated Nowcasting through Comprehensive Analysis.

meteoroloških spremenljivk numeričnega meteorološkega modela (npr. ALADIN), nato pa s pomočjo interpolacijskih metod ob upoštevanju določenih fizikalnih zakonitosti izračunava 3–dimenzionalno fizikalno konsistentno analizo v visoki krajevni ločljivosti (1 km), v katero vključuje širok spekter različnih meritev (podatke s klasičnih in avtomatskih meteoroloških postaj, radarske in satelitske podatke in še druge razpoložljive podatke izven državne meteorološke mreže, kot so na primer podatki CVP). Na ta način dobimo prostorska polja temperature, vlage vetra, razporeditve padavin, oblačnosti, sevanja itd. Ta analiza je nato osnova za izračun določenih diagnostičnih polj kot tudi za kratkoročno napoved za 12 ur vnaprej. ☞ Aprila 2010 je pričel teči projekt INCA–CE, katerega eden izmed glavnih ciljev je INCA meteorološko napoved razviti do te mere, da bo uporabna v drugih specialnih modelih [8, 90].

### 2.2.3 Temperatura cestišča

Za določitev stanja cestišča je potrebno poznavanje temperature cestišča in količine vode/vlage na cestišču in v plasti zraka nekaj metrov nad cestiščem [81]. Ključni sta ugotovitvi, kdaj se bo temperatura cestišča toliko znižala, da bo le-to pričelo zmrzovati (to se zgodi, ko je temperatura cestišča nižja od temperature zmrzovanja) ter kdaj se bo sneg ali led na cesti pričel topiti. Poznavanje temperature cestišča je pomembno tudi za pravilno odločanje pri izbiri ustreznega veziva pri polaganju asfalta.

Izkaže se, da je dejavnikov, ki vplivajo na temperaturo cestišča, veliko in so medsebojno nelinearno povezani (tabela 2.1). Najpomembnejši so temperatura zraka, padavine in sevalni tokovi [99]. Ti lahko na posameznem cestnem odseku povzročijo temperaturne razlike celo do 10 °C (slika 2.7) [23, 84], pogosto pa prihaja tudi do pojava, ko je temperatura cestišča pod zmrziščem zgolj na nekaterih delih celotnega cestnega omrežja [85].



Slika 2.7

Termalni posnetek cestnega odseka v bližini CVP Ravbarkomanda. Vidne so večje razlike v temperaturi cestišča. Vir: Articon d.o.o.

Tabela 2.1

Tipični dejavniki, ki vplivajo na temperaturo cestišča. *Prerejeno po [25, 99].*

Meteorološki	Geografski	Cestni
temperatura zraka in tal	zemlj. širina in dolžina	termalne lastn. ceste
dolgo kratkovalovno sevanje	nadmorska višina	deb. materialov ceste
tip in delež oblačnosti	topografija	emisivnost
hitrost vetra	vidnost neba	odbojnost (albedo)
vlažnost / rosišče	toplotni viri	promet
tip in količina padavin	raba tal	
zračni tlak	nagib tal	

Z meteorološkimi spremenljivkami, ki vplivajo na temperaturo cestišča, se je ukvarjal Gustavsson [46], ki predlaga uporabo preprostega modela, ki opisuje spremembe temperature cestišča in vzpostavlja povezavo med temperaturo cestišča, topografijo in meteorološkimi spremenljivkami. Pokazal je, da obstaja linearna povezava med temperaturo zraka in temperaturo cestišča za posamezno CVP. Ta povezava velja po sončnem zahodu v mirnih in jasnih nočeh, splošne enačbe med temperaturo zraka in temperaturo cestišča pa zaradi kompleksnosti ni bilo mogoče vzpostaviti. ♣ Temperature cestišča so ob oblačnih in vetrovnih razmerah večje [13]. Temperatura cestišča se v zavetrnih lokacijah lahko razlikuje tudi do 3 °C v primerjavi z izpostavljenimi območji, zavetrna območja pa imajo manjšo temperaturo cestišča [96]. Raziskovalci so v pokrajini West Midlands na podlagi podatkov 15 CVP lahko pojasnili največ sprememb temperature cestišča s faktorjem vidnosti neba in naklonom cestišča [17]. Podobno so tudi v [25] s pomočjo regresijske analize prišli do spoznanja, da ima na temperaturo cestišča ob visoki stabilnosti ozračja največji vpliv faktor vidnosti neba, ko pa stabilnost pada, ima pomemben vpliv nadmorska višina.

V nadaljevanju so nekateri dejavniki, ki pomembno vplivajo na temperaturo cestišča podrobneje opisani.

**Zemljepisna širina.** Zemljepisna širina pomembno vpliva na splošno klimo območja, na primer na količino snega pozimi. Poleg tega vpliva na količino prejetega sončnega sevanja, kar je mogoče opisati z ustreznimi enačbami [81].

**Višina.** Vpliv višine na temperaturo cestišča je pogosto nelinearen [85] in očiten zlasti v času majhne stabilnosti ozračja [25]. Temperatura cestišča pada z naraščanjem višine (v primeru temperaturne inverzije je pojav obraten). Tipičen padec je  $6,5\text{ }^{\circ}\text{C}$  na 1000 metrov, v ekstremih pa tudi do  $10\text{ }^{\circ}\text{C}$  [93].

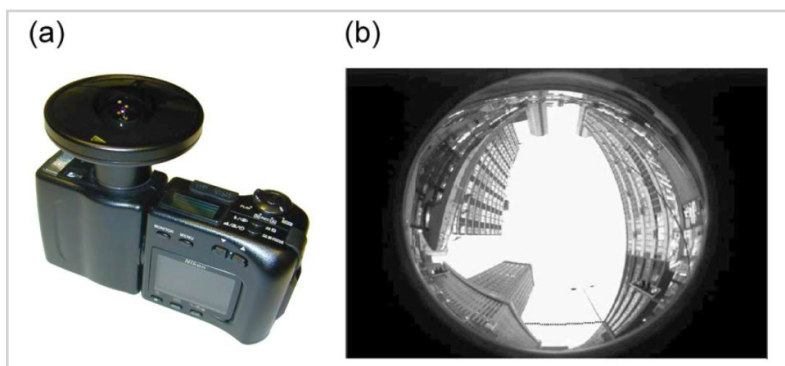
**Topografija.** V času jasnih in brezvetrnih noči so temperature zraka odvisne od številnih topografskih dejavnikov, kot so: globina in širina doline, okolica (vegetacija, stavbe), velikost območja hladnega zraka in podobno [46]. V splošnem so spremembe temperature zaradi topografije najmanjše na dnu dolin. Na razgibani topografiji se plast hladnega zraka začne premikati navzdol, na vrhu temperaturne inverzije pa se oblikuje toplejše območje zraka, kar imenujemo termalni pas in je odvisen od intenzitete zračnega premika in relativne velikosti obdajajoče topografije. ♣ Kondo in Okusa [60] sta vpliv topografije opisala s preprostim numeričnim modelom, medtem ko sta Bogren in Gustavsson [11] zaradi kompleksnosti uporabila empirični pristop. Določila sta dve najpomembnejši spremenljivki, ki vplivata na premikanje hladnega zraka v stabilnih pogojih, to sta globina doline  $D$  (v metrih) in širina doline  $W$  (v kilometrih). Z dvo-dimenzionalnim modelom doline:  $\Delta T = 0,35 + 0,9 \cdot W + 0,025 \cdot D$ , sta pojasnila okrog 55 % sprememb temperature cestišča  $T$ . Model predvideva, da se cesta nahaja na dnu doline in da je dolina navpične oblike.

**Osončenost.** Osončenost vpliva na količino prejetega kratkovalovnega sevanja. Zaradi manj prejetega kratkovalovnega sevanja imajo zasenčena mesta nižjo temperaturo cestišča [10]. Običajno je osončenost sistematična in se s časom ne spreminja (izjema so področja z listnatim drevjem). Termalno kartiranje je pokazalo, da imajo področja, ki niso zasenčena, relativno visoko temperaturo cestišča, razlike med zasenčenimi področji pa so bile med 2 in  $6\text{ }^{\circ}\text{C}$  [13]. Vpliv osončenosti je prisoten tudi po sončnem zahodu, kar še posebej velja v primeru nizke oblačnosti in med obdobjem zgodnje in pozne zime.

**Faktor vidnosti neba.** Označimo ga s  $\Psi_s$ . Opisuje vidnost neba in se uporablja na več področjih (gozdarstvo, zimsko vzdrževanje cest, onesnaženost neba). Vrednost ena pomeni povsem odprto nebo, vrednost nič pa povsem zaprto nebo. Na zmanjšanje  $\Psi_s$  vplivajo ovire (zlasti zgradbe in drevesa).  $\Psi_s$  se uporablja kot predstavnik osončenosti in je pomemben pri energijskih pretokih, saj vpliva na prejeto direktno kratkovalovno sevanje na cestno telo in na oddano dolgovalovno sevanje ceste. Tako imajo področja z



nizkim  $\Psi_s$  ponoči večjo temperaturo cestišča [73]. Po drugi strani pa imajo zasenčena mesta podnevi manjšo temperaturo cestišča zaradi manj prejetega direktnega kratkovalovnega sevanja [10]. Vpliv  $\Psi_s$  na temperature cestišča je manjši po opoldnevu, zlasti pa je pomemben med obdobjem zgodnje in pozne zime, ko je vpliv solarnega sevanja na napovedovanje temperature cestišča povečano [12]. Vpliv  $\Psi_s$  je največji v času visoke atmosferske stabilnosti, ko so sevalne izgube največje [25]. V raziskavi, opravljeni na Švedskem [77], je avtor kot eno izmed ključnih spremenljivk za določanje temperature cestišča navedel  $\Psi_s$ , saj je lahko zgolj z njegovim poznavanjem pojasnil kar 61 % variance temperature cestišča. ¶ Določanje  $\Psi_s$  običajno poteka s pomočjo fotografij neba [27, 44], pa tudi z opazovanjem GPS signala [21] ali dovolj podrobne topografije [38]. Izdelava fotografij neba poteka s kamero, opremljeno z lečo ribjega očesa (slika 2.8), nameščeno na strehi avtomobila. Za doseganje ustreznih rezultatov je potrebno meritve izvajati v mraku in ob homogeni oblačnosti [25]. Za obdelavo fotografij in računanje  $\Psi_s$  je na voljo več tehnik [9, 91].



*Slika 2.8*

Digitalna kamera, opremljena z lečo ribjega očesa (a) in primer črno bele fotografije neba, primerne za izračun faktorja vidnosti neba (b) [28].

**Raba tal.** Vpliv rabe tal je najbolj viden na stičiščih podeželskega in urbanega področja. V stabilnih pogojih atmosfere so pozidana območja običajno za nekaj stopinj toplejša. Na mestno klimo vplivajo tudi uporabljeni materiali za gradnjo in antropogena toplota (stavbe in promet) [73]. Eliasson [33] navaja, da je v času stabilnih pogojev temperaturna razlika med centrom mesta in parkom blizu mesta 4 °C. Podobno velja tudi za bližino večjih vodnih mas, kjer povečana vlažnost zmanjšuje temperaturne spremembe. Raba tal lahko pridobivamo na terenu ali pa jo odčitamo iz satelitskih

posnetkov [16].

*Konstrukcija ceste.* Spremembe temperature cestišča so odvisne tudi od fizikalnih lastnosti cestnega telesa [25]. Odbojnost (albedo) določa odboj kratkovalovnega sevanja od površine. Emisivnost površine vpliva na dolgovalovno sevanje površine v okolico. Hrapavost površine vpliva na vertikalno prehajanje toplotnih tokov med površino in zrakom. Poleg materiala je pomembna tudi globina konstrukcije, ki ima neposreden vpliv na pojem termalnega spomina (čas, ko ima na cesto toplota dnevnega solarnega sevanja še vpliv) [98]. Tako imajo mostovi in podobni objekti v primerjavi s cesto manjši termalni spomin. Podatke o konstrukciji ceste lahko pridobivamo z odčitavanji (tudi vrtanjem) na terenu ali pa iz banke cestnih podatkov.

*Gostota prometa.* Vozila vplivajo na temperaturo cestišča zaradi mešanja toplega in hladnega zraka (kar je najbolj izrazito v stabilnih atmosferskih pogojih) in zaradi povzročanja zasenčenosti cestišča [98]. Vpliv večjega števila počasi vozečih vozil lahko na temperaturo cestišča vpliva tudi do 2 °C. Podatke o gostoti prometa je mogoče statistično napovedovati [42] in jih nato vključevati v modele za napovedovanje stanja cestišča.

#### 2.2.4 Napoved stanja cestišča

Napoved stanja cestišča (denimo suho, mokro, zmrzal, led, sneg) predstavlja dragoceno informacijo voznikom, posebej pomembna pa je za zimsko vzdrževalno službo, saj ji omogoča zanesljivo predvidevanje nastanka poledice<sup>3</sup>. Napoved je še posebej pomembna na kritičnih odsekih cest (klanci, gorski prelazi, ozke soteske, viadukti, vetru izpostavljene ceste), kjer prihaja do ekstremnih zimskih razmer.

Napovedovanje meteorološkega stanja cestišča zgolj z upoštevanjem klasične meteorološke napovedi ni ustrezno. Zato je v sredini prejšnjega stoletja področje cestne meteorologije postalo predmet intenzivnega proučevanja. Raziskave so bile omejene zlasti na proučevanje vpliva posameznih dejavnikov na cesto<sup>4</sup>.

Za doseganje ustreznih napovedi je pomembno upoštevati meteorološke napovedi z veliko časovno in prostorsko ločljivostjo ter podatke meritev iz cestnovremenskih

<sup>3</sup>Prim. razdelek 2.2.5 na strani 30.

<sup>4</sup>Prim. razdelek 2.2.3 na strani 18.

postaj.<sup>5</sup> Že leta 1944 je David Brunt podal zelo splošno enačbo za napovedovanje minimalne temperature tal [19], v zadnjih desetletjih pa je bilo razvitih več specialnih modelov za napovedovanje stanja cestišča.

*Termalno kartiranje.* Prvi pristopi napovedovanja temperature cestišča so temeljili na metodi termalnega kartiranja. Gre za prikaz sprememb temperature cestišča na cestnem omrežju. Razvoj termalnega kartiranja sega v 80. leta prejšnjega stoletja [49, 92]. Rezultate prikažemo na grafu, ki ga imenujemo termalni odtis ceste. Abscisna os predstavlja razdaljo, ordinatna os pa temperaturni odmik od povprečne vrednosti na določeni razdalji (slika 2.9). ¶ Termalno kartiranje se izvaja tik pred zoro ob različnih vremenskih pogojih. Meritve potekajo z vozilom, opremljenim z infrardečo kamero, ki v določenih časovnih intervalih meri energijski tok ceste  $E$ . Energijo je mogoče neposredno pretvoriti v temperaturo s pomočjo Stefan-Boltzmannove enačbe, ki enači energijski tok  $E$  z zmnožkom četrte potence temperature ceste  $T$ , Stefan-Boltzmannove konstante  $\rho$  in emisivnosti  $e$  (običajna vrednost za asfalt je 0,95):

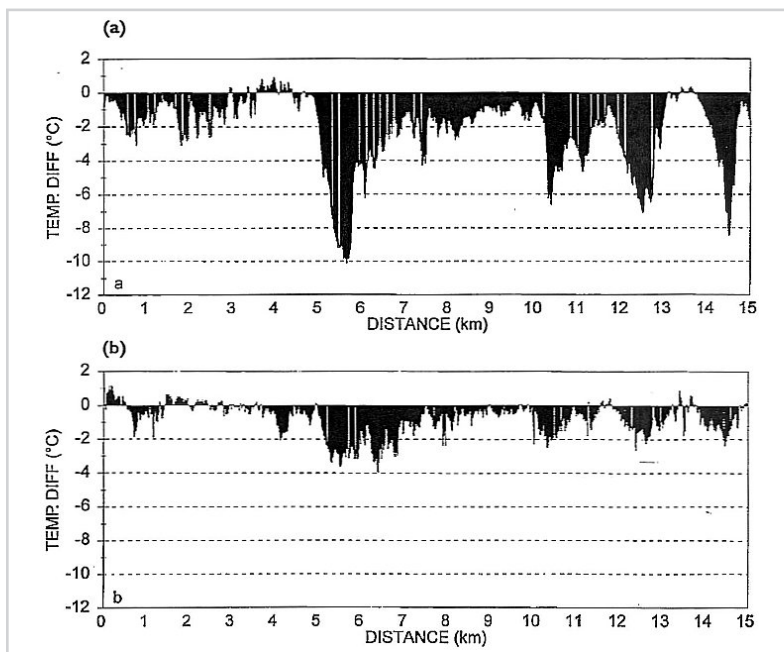
$$E = e \cdot \rho \cdot T^4. \quad (2.12)$$

Shao s sodelavci [84] je opisal natančnost, ponovljivost in zanesljivost termalnega kartiranja. Postopek je potrebno opraviti kvalitetno in z ustrezno opremo, saj lahko več dejavnikov povzroči skupno napako tudi do 5 °C:

- spreminjanje emisivnosti ceste (razlika v sestavi in starosti asfalta ter vremenskih pogojev),
- sevanje vozil,
- nečistoče na lečah merilnika,
- ekstremni temperaturni pogoji v času merjenja.

V splošnem so si termalni odtisi ceste v podobnih stabilnih pogojih podobni, zato lahko v želeni situaciji napovedovanja izberemo tisti termalni odtis, kjer so razmere najbolj podobne meteorološki napovedi. Poleg napak pri izvajanju meritev, ki se jim v polnosti ne da izogniti, je velika slabost pristopa termalnega kartiranja ta, da so meritve

<sup>5</sup>Obstajajo tudi modeli, ki za izdelavo zelo kratkoročnih napovedi stanja cestišča potrebujejo zgolj meritve iz cestnovremenskih postaj [83].



Slika 2.9

Meritve temperature zraka (a) in temperature cestišča (b) vzdolž cestnega omrežja [48]. Prikazana so odstopanja temperature od referenčne temperature.

veljavne zgolj v trenutku izvajanja meritev, zato je uporaba zgolj termalnih odtisov ceste za namene napovedovanja stanja cestišča manj primerna. Termalno kartiranje se sicer pogosto uporablja kot validacijska tehnika pri prostorskem napovedovanju temperature cestnega omrežja in za identifikacijo najbolj mrzlih področij na nekem cestnem omrežju, kar lahko služi kot objektivno merilo za določanje lokacij in števila CVP ter za izdelavo prioritet, katere cestne odseke je potrebno najprej posipati.

**Fizikalni modeli.** Napovedovanje stanja cestišča je mogoče s pomočjo fizikalnega energijsko–bilančnega modela. Gre za natančen opis tokov vode in vodne pare (izhlapevanje, odtokanje in pronicanje padavin) ter toplote (sevanje, kondukcija, turbulentni transport) s pomočjo izračuna energijskih enačb. Od pravilne ocene omenjenih tokov so odvisne lastnosti vozne površine. ¶ Od 80-ih let prejšnjega stoletja je bilo razvitih več različnih fizikalnih modelov z zadovoljivo natančnostjo napovedovanja [29, 61, 79–81, 83, 97]. Tovrstni modeli potrebujejo dobro poznavanje začetnega stanja in natanč-

ne vhodne podatke, ki jih lahko zagotavljajo meritve iz CVP in meteorološke napovedi. Dejavniki, ki jih navedeni modeli parametrizirajo, se od modela do modela razlikujejo, vendar pa jih je mogoče vseskozi dopolnjevati (npr. dodatna parametrizacija vpliva prometa [94]).

ISBA-Route/Crocus [15] je v Franciji razvit model za simulacijo temperature cestišča ter snega na površini ceste (količina, tip). ISBA-Route/Crocus združuje enodimenzionalni fizikalni model in snežni model. Za napovedovanje uporablja zgolj meteorološke napovedi, začetni pogoji (temperatura ceste, profil vlažnosti) pa se vzpostavijo z daljšimi opazovanji in iterativnim izvajanjem modela. Model so uporabili za prostorsko napovedovanje z ločljivostjo 8 km za celotno Francijo. Avtorji so za celotno območje uporabili enake pogoje: sestava ceste, albedo (0,07) in emisivnost (0,96). V modelu niso upoštevani dejavniki, kot so soljenje, pluženje, senčenje in promet. Globalna srednja kvadratična napaka je 2,8 °C pri uporabi podatkov iz opazovanj ter 4,1 °C pri uporabi podatkov meteoroloških analiz.

Icebreak [81] je model, ki uporablja enodimenzionalno difuzijsko enačbo toplote. Predvideva, da je struktura ceste horizontalno homogena, zato širjenje toplote v okolici zanemari. Icebreak v primeru padavin predvideva, da je voda enakomerno (tako časovno kot prostorsko) prisotna na cestnem telesu. Stanje cestišča določi glede na mejno debelino vodnega filma in ali je temperatura cestišča nad ali pod 0 °C. V energijski enačbi modela je sevanje razdeljeno na direktno in difuzno sevanje. Ko je lokacija zasenčena, je upoštevano zgolj difuzno sevanje, sicer sta upoštevani obe sevanji. Vhod v model so meteorološke napovedi temperature zraka, temperature rosišča, hitrosti vetra, oblačnosti in padavin. Vpliv soljenja in prometa Icebreak ne upošteva.

*Fizikalni model METRo.* METRo<sup>6</sup> [29, 66] je model za napovedovanje temperature in stanja cestišča, ki se je začel razvijati leta 1999 v Kanadi in se trenutno uporablja v več državah po svetu: Avstrija, Češka Republika, Finska, Francija, Hrvaška, Irska, Italija, Kanada, Litva, Nova Zelandija, Rusija, Slovaška, Slovenija, Švedska, Velika Britanija in ZDA. Uporablja meritve iz cestnovremenskih postaj in meteorološke napovedi ter na podlagi teh podatkov napoveduje temperaturo in stanje cestišča.

Izvajanje modela zajema tri stopnje:

1. Inicializacija temperaturnega profila ceste z uporabo zgodovinskih meritev na CVP. Dobljeni koeficienti se uporabijo pri napovedi.

---

<sup>6</sup>Model of Environment and Temperature of Roads.

2. Sklapanje meteorološke napovedi z opazovanji med obdobjem prekrivanja meritev na CVP in preteklih meteoroloških napovedi. To omogoča prilagoditev koeficientov lokalnim pogojem ter popravek odmikov.

3. Napoved.

Model uporablja eksplicitno shemo s 30 sekundnim časovnim intervalom. Mreža je neenakomerna in ima 29 nivojev s korakom od 0,01 do 0,05 m. Največja globina je 1,4 m. ¶ Crevier in Delage navajata [29], da je približno polovico časa napaka napovedane temperature cestišča znotraj območja  $\pm 2$  K.

Model METRo sestoji iz treh modulov, ki so v nadaljevanju podrobneje opisani.

ENERGIJSKO-BILANČNI MODUL POVRŠJA CESTIŠČA. Osnova za zapis bilance energijskih tokov na površju cestišča je enačba:

$$R = (1 - \alpha)S + \varepsilon I - \varepsilon \sigma T_s^4 - H - L_a E \pm L_f P + A, \quad (2.13)$$

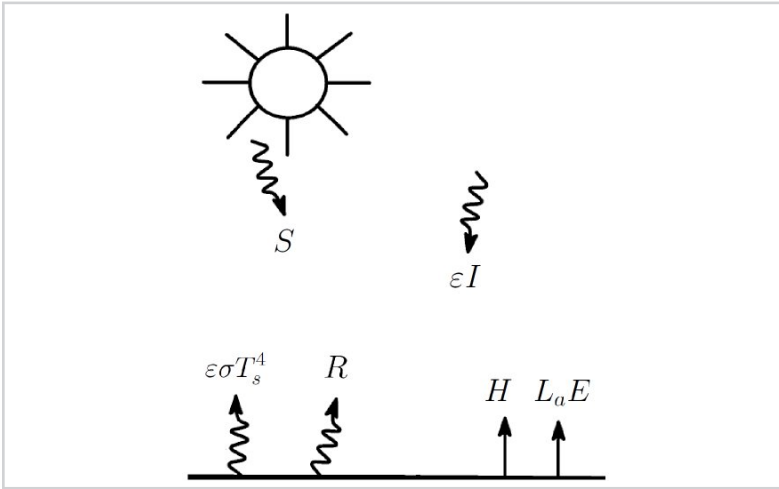
kjer je  $R$  skupno odbojno sevanje ceste,  $\alpha$  albedo,  $S$  vpadno sončno sevanje,  $\varepsilon$  emisivnost,  $I$  vpadno infrardeče sevanje,  $\sigma$  Stefan-Boltzmannova konstanta,  $T_s$  temperatura ceste,  $H$  tok zaznavne toplote,  $L_a$  izparilna toplota vode,  $E$  tok izhlapevanja,  $L_f$  talilna toplota vode,  $P$  intenziteta padavin,  $A$  antropogeni toplotni tok. Tako je  $(1 - \alpha)S$  prejeta sončno sevanje,  $\varepsilon I - \varepsilon \sigma T_s^4$  razlika med absorbiranim infrardečim sevalnim tokom in sevanjem ceste,  $L_a E$  tok latentne toplote,  $\pm L_f P$  tok faznih sprememb vode ( $\pm$  pomeni zmrzovanje ali taljenje). Grafična predstavitev energijskih tokov je za lažjo predstavo ponazorjena na sliki 2.10. ¶ Albedo se spreminja linearno od 0,1 do 0,5 (zasnežena cesta). Za  $\varepsilon$  je uporabljena vrednost 0,92, za  $A$  pa  $10 \frac{W}{m^2}$ . Turbulentni tokovi se računajo posebej:

$$H = -\rho c_p C_m \cdot \max(V, V_c) \cdot C_h \cdot (T_a - T_s) \quad (2.14)$$

in

$$E = -\rho C_m \cdot \max(V, V_c) \cdot C_h \cdot (q_a - q_s), \quad (2.15)$$

kjer je  $\rho$  gostota zraka,  $c_p$  specifična toplota zraka pri konstantnem pritisku,  $V$  hitrost vetra na 10 m,  $V_c$  minimalna hitrost vetra,  $T_a$  temperatura zraka,  $q_a$  specifična vlažnost zraka na višini,  $q_s$  specifična vlažnost zraka pri tleh.  $C_m$  in  $C_h$  sta brezdimenzijski koeficienta. ¶ Temperatura ceste  $T_s$  je med računanjem energijske bilance konstantna.



Slika 2.10

Prikaz nekaterih energijskih tokov na površini cestišča. Prirjeno po [7].

MODUL TROPLOTNE PREVODNOSTI CESTE. Računanje temperaturnega profila ceste  $T$  se računa z uporabo enodimenzionalne difuzijske enačbe toplote:

$$C(z) \frac{\partial T(z, t)}{\partial t} = - \frac{\partial G(z, t)}{\partial z} \tag{2.16}$$

kjer je  $C$  toplotna kapaciteta,  $z$  globina,  $t$  čas in  $G$  toplotni tok v tleh, ki ga parametriziramo z:

$$G(z, t) = -k(z) \frac{\partial T(z, t)}{\partial z}, \tag{2.17}$$

kjer je  $k$  toplotna prevodnost.  $C$  in  $k$  sta odvisna od sestave ceste (asfalt, beton, pesek, zemlja ...) in imata različne vrednosti glede na globino  $z$ . Mehanizem vsebnosti vode v cesti ni vključen. Robni pogoj je  $G_0 = R$ .

MODUL AKUMULACIJE VODE, SNEGA IN LEDU uporablja dva rezervoarja:  $W_l$  za vodo in  $W_s$  za sneg/led, ki simulirata količino vode oz. snega/ledu na cesti. V nekem času je lahko poln samo en od obeh rezervoarjev. Njuna vsebina je določena z:

$$\frac{dW_l}{dt} = P - E + \frac{R - G_1}{L_f} - r \tag{2.18}$$

in

$$\frac{dW_s}{dt} = P - E - \frac{R - G_1}{L_f} - r, \quad (2.19)$$

kjer je  $G_1$  toplotni tok med prvo in drugo plastjo modela v cesti in  $r$  odtok padavin.  $P$  in  $E$  vplivata zgolj na en rezervoar, kar je odvisno od temperature površja (nad ali pod  $0^\circ\text{C}$ ). Člen  $\frac{R-G_1}{L_f}$  je aktiven le pri  $0^\circ\text{C}$ . Odtok  $r$  opisuje preprosta eksponentialna funkcija. Tudi odstranitev snega je parametrizirana s podobno funkcijo kot  $r$ .

*Statistični modeli.* Na problematičnih mestih razni dejavniki (razgibana topografija, snežna odeja, gost promet itd.) povzročijo poslabšanje delovanja fizikalnih modelov [82]. Ker so takšni dejavniki navadno odvisni od lastnosti ceste in njene bližnje okolice, je pogosto težko izmeriti njihov vpliv in jih ustrezno parametrizirati. Zaradi tega bi bilo smiselno fizikalne modele zamenjati ali vsaj dopolniti z modeli, sestavljenimi na podlagi preteklih podatkov, ki so bili zbrani za posamezno lokacijo. Boljše rezultate fizikalnega modela lahko dosežemo tako, da se vhodne ali izhodne spremenljivke fizikalnega modela izboljša s statističnimi pristopi [74, 82]. Obstajajo pa tudi primeri, kjer so bili za napovedovanje uporabljeni zgolj statistični pristopi [6, 86].

Model nevronske mreže je že leta 1998 uporabil Shao [82], kasneje pa sta Pasero in Moniaci [74] z uporabo nevronske mreže s pomočjo podatkov iz CVP modelirala količino padavin, ledu in snega na cesti do tri ure vnaprej. Model sta testirala na petletnih podatkih na izbrani meteorološki postaji v Italiji.

Fizikalno-statistični model za napovedovanje prijema na cestišču so razvili na Finskem meteorološkem inštitutu [53, 72]. Fizikalni del modela upošteva tudi promet in deluje od leta 2000 naprej. Statistični del modela temelji na korelacijski analizi med meritvami prijema brezkontaktnega merilnika DSC111<sup>7</sup> ter med meteorološkimi meritvami na cesti in izbranimi meteorološkimi spremenljivkami iz fizikalnega dela modela, ki so povezani s trenjem (količina vode, ledu in snega na cestišču ter temperatura cestišča). Glede na pogoje na cesti (suho, voda, led/sneg) so uporabili tri preproste regresijske modele. V primeru zasneženega ali poledenelega cestišča se uporabi za napovedovanje višino snežne odeje oziroma debelino ledu in temperaturo cestišča. V primeru, ko je na cestišču voda, je oprijem odvisen od debeline vodnega filma. V primeru suhe ceste je oprijem konstanten (0,82). Regresijski koeficienti se med enačbami za različne CVP razlikujejo, vendar v realni aplikaciji zaradi praktičnosti uporabljajo

<sup>7</sup>Prim. razdelek 2.2.1 na strani 12.



prostorsko neodvisen model. Za učenje so uporabili podatke iz zime 2007/2008, validacija pa je potekala v prihodnji zimi. Rezultati kažejo ustrezna ujemanja v vseh stanjih ceste (dobljene korelacije med napovedanimi in izmerjenimi vrednostmi oprijema se gibljejo okrog 0,9).

Berrocal s sodelavci [6] je uporabila dve statistični metodi, ki temeljita na linearni oz. logistični regresiji, za napovedovanje verjetnosti nastanka ledu na cestišču, ni pa se ukvarjala z napovedovanjem temperature cestišča. Privzela je, da je potrebno ukrepanje, če obstaja vsaj ena točka na nekem cestnem odseku, ki ima temperaturo zraka nižjo od 0 °C in se tam pojavijo padavine. Metodi je preizkusila na izbrani avtocesti v zvezni državi Washington (ZDA) v letih 2002–2005.

Temperaturo cestišča v odvisnosti od meteoroloških pogojev sta statistično modelirala Sherif in Hassan [86]. Uporabila sta podatke iz obdobja zime 2001/2002 iz 6 CVP v Ottawi (Kanada), ki velja za eno najbolj zasneženih mest na svetu. Za napovedovanje trenutne temperature cestišča sta za vsako CVP uporabila svoj linearni regresijski model. V končnih modelih sta obdržala zgolj spremenljivke z ustreznim doprinosom glede na nivo zaupanja in korelacijski koeficient  $R^2$ . Tipično so to bile temperatura zraka, temperatura rosišča, stanje cestišča, vlažnost, hitrost in smer vetra. Zgolj s temperaturo zraka in rosišča sta lahko pojasnila več kot 80 % sprememb temperature cestišča. Ker so se vse CVP nahajale na istem geografskem območju, sta preizkusila uporabo zgolj enega modela, ki je imel  $R^2 = 0,84$ . Napoved sta znatno izboljšala, če sta v modele vključila tudi temperaturo cestišča za eno ali dve uri pred pričetkom napovedi in tako dosegla  $R^2 \approx 0,98$ . Avtorja zaključujeta, da njuno modeliranje sicer ne more nadomestiti CVP, v prihodnje pa si želita te modele uporabiti za napovedovanje temperature cestišča.

*Prostorska napoved stanja cestišča.* Napovedi stanja cestišča so običajno točkovne ravne in eden izmed pomembnih izzivov, s katerimi se soočajo raziskovalci cestne meteorologije, je ekstrapolacija točkovnih napovedi na prostorsko mrežo. Bogren, Gustavsson in Lindqvist [13, 47] opisujejo uporabo preprostega modela za prostorsko napovedovanje temperature cestišča. Gre za uporabo empiričnih formul za ekstrapolacijo točkovnih napovedi temperature cestišča na podobna območja. Avtorji so prišli do zaključka, da je takšno modeliranje uporabno zgolj na enem območju in da prenosljivost v splošnem ni mogoča.

Tradicionalno se je problematiko prostorske napovedi reševalo s tehniko termalnega

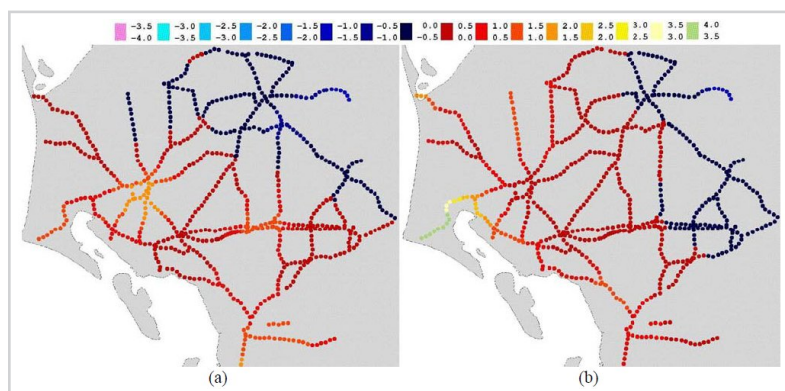
kartiranja<sup>8</sup>, ki nudi vpogled v temperaturna odstopanja cestnega omrežja ob določenih meteoroloških pogojih. Shao [85] navaja, da je termalno kartiranje zanesljiva in učinkovita metoda za prikaz prostorskih in časovnih sprememb temperature cestišča. Predlaga, da se prostorska napoved izvaja s točkovnim fizikalnim modelom v povezavi s termalnim kartiranjem. Po drugi strani pa nekaj let kasneje drugi avtorji [25] opozarjajo na vrsto slabosti pri uporabi pristopa termalnega kartiranja in predlagajo nov pristop: zbiranje prostorskih podatkov v visoki ločljivosti, ki so primerni za ustrezno opisovanje sprememb temperature cestišča. V točkovnem fizikalnem modelu so uvedli prostorsko komponento tako, da so prostorske konstante nadomestili s spremenljivkami. Na ta način so dobili veliko število točkovnih napovedi, ki so skupaj tvorile prostorsko napoved na cestnem omrežju. Podoben pristop 24-urne prostorske napovedi na podlagi fizikalnega modela, prostorskih spremenljivk in meteorološke napovedi opisujeta tudi Chapman in Thornes [22].

Za dobro napoved stanja posameznega cestnega odseka je v modeliranje potrebno vključiti kar največ lokalnih podatkov o cestnem omrežju [75]. V zadnjem času tako vse več raziskovalcev predlaga uporabo geografske podatkovne baze [17, 22, 25, 26, 38], ki naj bi vsebovala naslednje spremenljivke: zemljepisna dolžina in širina, višina, faktor vidnosti neba, konstrukcija ceste, naklon ceste, gostota prometa, raba tal, topografija (tudi z upoštevanjem zgradb in dreves) itd. Zbiranje takšnih prostorskih podatkov lahko poteka zgolj s pomočjo GIS orodij [38]. Poleg tega je pomembna tudi uporaba meteorološke napovedi visoke prostorske in časovne ločljivosti [24]. Na ta način je mogoče učinkovito povezati prostorske spremenljivke s točkovnim fizikalnim modelom in uspešno modelirati stanje cestnega odseka (slika 2.11), za njeno verifikacijo pa zadostuje že manjše število točk [51].

### 2.2.5 Delo in optimizacija zimske službe v Sloveniji in drugod

Zimska služba oz. zimsko vzdrževanje cest je zelo obsežen, zahteven in drag segment v sklopu rednega vzdrževanja cest in obsega sklop dejavnosti in opravil (pregledovanje, posipanje in pluzenje cest), potrebnih za omogočanje prevoznosti cest in varnega prometa. Pogoste dnevne in sezonske spremembe v temperaturah in padavinah povzročijo velike spremembe vozniških pogojev in terjajo pogoste in učinkovite odzive zimske službe. ¶ Največji izziv predstavlja zimski službi nastanek poledice (ledu na cestišču). Do

<sup>8</sup>Prim. razdelek 2.2.4 na strani 22.



Slika 2.11

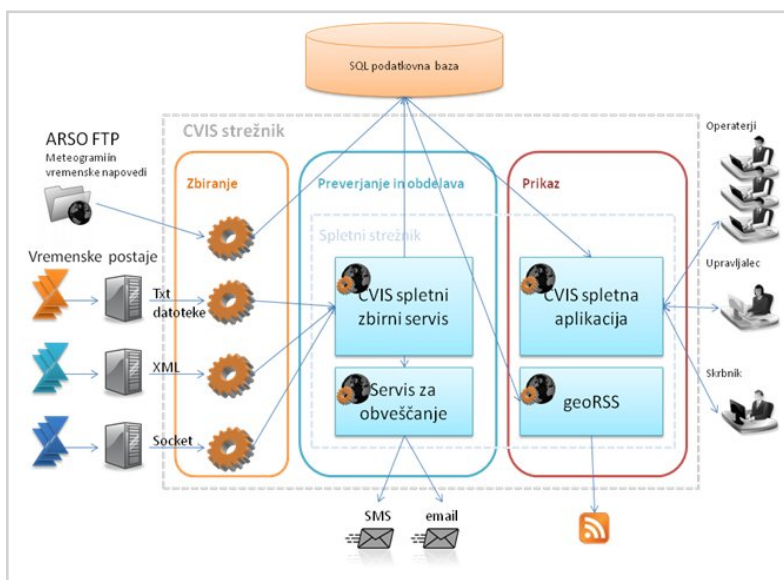
Primer triurne prostorske napovedi temperature cestišča (a) in napovedi eno uro kasneje (b) na Danskem [68].

pojava poledice lahko pride zaradi podhladitve cestišča, nenadne vlage pri nizkih temperaturah, slane, snežne plohe ali ledenega dežja (žled). Največja pogostost poledice nastopi ob pogojih, ko je podnevi toplo (tali se sneg), ponoči pa zmrzuje. Nastanek poledice je tipičen zlasti za kritične cestne odseke, ki so znani na podlagi izkustev: ostre krivine, večje strmine, mostovi, senčni odseki (posebej v gozdovih in ob vodotokih), cestni prehodi preko železnice, križišča in podobno.

Odstranjevanje snega z voznih površin se prične takrat, ko višina snega na cestah preseže določeno mejno vrednost, ki je odvisna od prednostnega razreda ceste. Zimska služba poleg pluženja izvaja posipavanje, ki povzroča taljenje ledu in snega ter izboljša oprijem na cestišču. Za posipanje cest se najpogosteje uporablja morska ali kamena sol ( $NaCl$ ) sama ali kot mešanica soli in raztopine  $CaCl_2$  oz.  $MgCl_2$  in drobljenec. Izbira načina posipanja in ustreznih materialov je odvisna od strokovne presoje glede na vsakokratne okoliščine in prednostni razred ceste. Posipanje se lahko začne izvajati takoj, ko se na cestišču zazna pojav poledice. Bistveno boljše pa je ukrepanje pred nastankom poledice, kar ima za posledico znatne prihranke.

Slovenija leži v meteorološko zelo raznolikem območju. Povprečne temperature v najhladnejših mesecih ne padejo pod  $-3\text{ }^{\circ}C$ , vsaj štirje meseci pa imajo povprečno temperaturo zraka nad  $10\text{ }^{\circ}C$ . Povprečne letne najnižje dnevne temperature so pod  $2\text{ }^{\circ}C$  v visokih predelih, okrog  $8\text{ }^{\circ}C$  v predelih blizu morja in okrog  $5\text{ }^{\circ}C$  v ostalih delih. Slovenija ima več kot 38 tisoč kilometrov javnih in 13 tisoč kilometrov gozdnih cest, na katerih se zimska služba izvaja v skladu z Zakonom o cestah [32].

*Cestnovremenski informacijski in odločitveni sistemi.* Včasih so informacije o stanju na cestnem omrežju pridobivali zgolj z izvajanjem izvidniških pregledov, danes pa lahko velik del te naloge opravijo cestnovremenski informacijski sistemi (CVIS). CVIS so postali nepogrešljivo orodje cestni vzdrževalni službi, saj omogočajo pregled nad trenutnim stanjem na cestišču na lokacijah CVP in pomembno prispevajo k boljši varnosti v cestnem prometu, k optimiziranju dejavnosti zimskih služb in k varovanju okolja [4, 41, 52]. Podatki CVP se preko komunikacijske opreme zapisujejo v skupno bazo podatkov, kjer so na voljo za prikazovanje in analiziranje. Ker je na CVP pogosto nameščena merilna oprema različnih proizvajalcev, ki posreduje različne veličine podatkov, mora CVIS skrbeti tudi za njihovo poenotenje.



Slika 2.12

Shema slovenskega cestnovremenskega informacijskega sistema.

V Sloveniji je CVIS vzpostavljen na Direkciji Republike Slovenije za ceste ter na Družbi za avtoceste v Republiki Sloveniji [56] in je zasnovan kot spletna aplikacija na *Microsoft .NET* tehnologiji, za prostorski del prikaza pa uporablja tehnologijo *MapGuide Open Source* (slika 2.12). Aplikacija omogoča pregled vseh CVP, pregled vremenskih podatkov na posameznih postajah (slika 2.13), izdelavo mesečnih poročil in *GeoRSS* servis za posredovanje podatkov. Aplikacija omogoča naslednje funkcional-

nosti:

- pregled vseh CVP postaj na različnih kartografskih podlagah,
- celosten pregled vremenskih podatkov za vse postaje v sistemu,
- pregled vremenskih podatkov za posamezno CVP,
- tabelarni in grafični pregled merilnih podatkov za poljuben dan,
- pregled osnovnih podatkov o posamezni CVP,
- izdelava mesečnih poročil,
- administracija CVP, senzorjev in merjenih količin,
- GeoRSS servis za posredovanje podatkov.



Slika 2.13

DARS cestnovremenski informacijski sistem. Pregled podatkov meritev na izbrani cestnovremenski postaji.

Nadaljnji razvoj CVIS poteka v smeri celovitih sistemov za podporo odločanju v zimski vzdrževalni službi (Maintenance Decision Support System — MDSS) [20]. Takšni sistemi rešujejo problematiko zahtevnega in odgovornega sprejemanja odločitev

vzdrževalcev v zimski službi, ki potrebujejo čim boljše odgovore na vprašanja: kaj, kdaj, kje, kako pomembno in na kakšen način je potrebno izvesti ukrep na cestnem omrežju in omogoča optimizacijo zimske službe, nadzor nad izvajanjem zimske službe, povečanje varnosti voznikov, racionalizacijo rabe posipnih materialov in zmanjšanje škodljivih vplivov na okolje [108]. MDSS omogoča prikaz vseh pomembnih informacij, kratkoročno napoved stanja cestišča, nadzor nad zimsko službo in integracijo z medsebojno odvisnimi ali sorodnimi sistemi. Razvoj MDSS poteka v ZDA [76], na Japonskem [95], v Avstriji [30], na Slovaškem [5], pa tudi v Sloveniji [88] in drugod.

*Vidiki varnosti, prihrankov in okolja.* Zimska služba mora ob zmanjševanju stroškov pri vzdrževanju cest omogočati visoko mobilnost in varnost prometa. Stroški zimskega vzdrževanja cest neprestano rastejo, kar je razumljivo, saj se povečujejo tako prometne obremenitve kot zahteve uporabnikov. Poleg tega povzročajo prometni zastoji ekološke, finančne in druge posledice posamezniku in nacionalnemu gospodarstvu. Učinkovita zimska služba pomembno vpliva na:

- zmanjševanje števila prometnih nesreč [37, 71] (v Veliki Britaniji poročajo [82], da napovedni sistemi vsako leto zmanjšajo število smrtnih žrtev v prometnih nesrečah za 25 do 50),
- zmanjševanje zdrsov vozil in cestnih zastojev,
- zmanjševanje porabe goriva,
- zmanjševanje izpusta nevarnih snovi v okolje, kot so: ogljikov dioksid ( $CO_2$ ), ogljikov monoksid ( $CO$ ), dušikov dioksid ( $NO_2$ ), prizemni ozon, svinec in prašni delci ( $PM_{10}$ ), ki uničujejo rastline, drevesa in zgradbe; pri ljudeh, ki so jim najbolj izpostavljeni, pa redno prihaja do resnih poškodb in bolezn.

$NaCl$  je tako kot vse soli škodljiv za okolje, saj ga povprečno za 1 km avtoceste potrebujemo več kot 10 ton. Posledice soljenja so [78]:

- korozija vozil,
- škodljiv vpliv na cestno infrastrukturo,
- onesnaževanje bližnjih kmetijskih in drugih zemljišč,

- onesnaženost površinskih voda in podtalnice,
- zaviranje rasti rastlin, dreves in grmičevja.

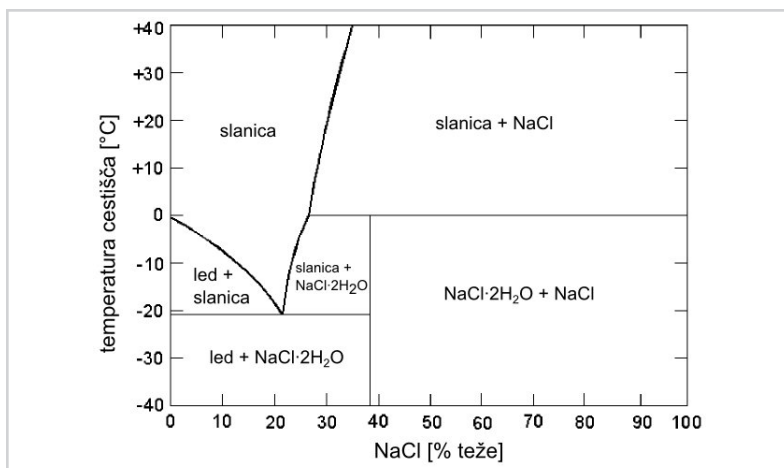
Poraba soli je tako omejena na najmanjšo možno mejo, ki še zagotavlja učinkovito odpravo poledice. S pravilno izbiro materialov za posip in z ustreznimi tehnologijami je potrebno v kar največji meri zmanjšati škodljive vplive porabljenih materialov [89]. Pri posameznih posipalcih za mokro ali suho posipanje, ki so opremljeni z napravami za nastavitev doziranja, se okvirne količine soli gibljejo med 5–50 g/m<sup>2</sup>, količine gramoznega materiala pa med 25–200 g/m<sup>2</sup>.

Za odstranitev ledu s površine ceste je potrebno tudi štirikrat in več posipnih materialov kot pri izvajanju preventivnega posipavanja [83], zato za preventivno posipavanje zadostuje že količina 5–10 g/m<sup>2</sup> [54]. Preventivni posip samo s soljo (NaCl) ali mešanico gramoz in soli ni primeren, ker zaradi prometa na cesti pride do prevelikega izmeta posipnega materiala s cestišča (do 90 %). Pri posipu z vlažno soljo (v razmerju 70 % MgCl<sub>2</sub> ali CaCl<sub>2</sub> in 30 % NaCl) pa je izmeta bistveno manj, saj se posipni material prilepi na cestišče in tako učinkuje daljši čas. ¶ V primeru pomanjkanja podatkov o prihodnjem stanju ceste se preventivno posipavanje navadno ne izvaja, saj bi prekomerno posipanje povzročalo le negativne učinke. Napoved stanja cestišča omogoča poleg preventivnega posipavanja še organizacijske prihranke s pravočasnim planiranjem aktivnosti zimske službe, organizacijo človeških virov, mehanizacije in posipnih materialov. Za določitev ustrezne količine posipnega materiala je ključno merjenje ostanka soli na cestišču od zadnjega posipa, poleg tega pa tudi poznavanje napovedi temperature cestišča, ki vpliva na izbiro posipnega materiala in njegovo koncentracijo (slika 2.14).

V Veliki Britaniji letni stroški zimskega vzdrževanja presegajo 140 milijonov funtov vsako leto, korozija soli pa povzroča nadaljnjih 100 milijonov funtov škode na vozilih in zgradbah [26]. V zimi 1997/1998 so v Kanadi porabili kar 5 · 10<sup>6</sup> ton soli [86], v Sloveniji pa se vsako zimo porabi okrog 150 tisoč ton soli za odstranjevanje ledu [62]. ¶ Shao in Lister [83] navajata, da so v Veliki Britaniji (pokrajine Cheshire, Hereford, Worcester) po uvedbi napovednih sistemov prihranili 15–20 % soli. V zvezni državi Wisconsin v ZDA ocenjujejo prihranke zaradi napovedi na okrog 75 tisoč dolarjev in 2500 ton soli na tipično snežno nevihto [82]. V zvezni državi Indiana v ZDA so v zimi 2008/2009 z uvedbo MDSS v primerjavi s prejšnjo sezono prihranili ≈188 tisoč ton soli (36 % prihranek) [69].

Slika 2.14

Fazni diagram vode in  $\text{NaCl}$  v odvisnosti od koncentracije  $\text{NaCl}$  in temperature cestišča. Za preprečevanje tvorjenja ledu so pri nižjih temperaturah cestišča potrebne bistveno višje koncentracije  $\text{NaCl}$ , pri evtektrični točki (približno pri  $-21\text{ }^{\circ}\text{C}$ ) pa se tvorjenju ledu ni mogoče izogniti. *Prirjeno po [105].*





# *Opis podatkov*

Poglavje podaja podrobnejši opis podatkov, ki so v tej nalogi uporabljeni. Ti se v osnovi delijo na izmerjene oziroma izračunane podatke iz cestnovremenskih postaj (CVP) in meteoroloških modelov, ter na podatke, ki smo jih pripravili umetno. Obe vrsti podatkov omogočata, da se metode, predstavljene v tej nalogi, preizkusi tako iz teoretičnega kot iz aplikativnega vidika. Branje bo olajšalo poznavanje področja cestne meteorologije, ki je opisano v razdelku 2.2.

### *3.1 Izmerjeni podatki in podatki vremenskih napovedi*

Razdelek opisuje podatke, pridobljene iz CVP in podatke meteoroloških modelov (vremenske napovedi). Navedene so CVP, katerih podatki so nam bili na voljo, njihov opis, obdobje meritev in vhodne spremenljivke. Pri tem so ločeno opisani podatki, ki so bili uporabljeni pri metodi za gradnjo hierarhičnih regresijskih modelov (razdelek 4.1), in podatki, ki so bili uporabljeni za analizo povezav med modeli in lastnostmi objektov, na katere se ti modeli nanašajo (razdelek 4.3). Pri slednjih je navedenih več lastnosti lokacij CVP, poleg tega pa so navedeni tudi regresijski koeficienti modelov, ki so bili pripravljene z metodo za gradnjo hierarhičnih regresijskih modelov (razdelek 4.1) in so pri analizi povezav med modeli in lastnostmi objektov uporabljeni kot vhodni podatki.

#### *3.1.1 Podatki za preizkus gradnje hierarhičnih modelov*

Podatke, ki so bili uporabljeni pri preizkusu metode za gradnjo hierarhičnih regresijskih modelov, smo pridobili iz treh izbranih cestnovremenskih postaj (CVP Jeprca, Mislinja in Črmošnjice), ki se med seboj, zlasti v smislu topografskih značilnosti, zelo razlikujejo. Vse tri CVP so opremljene z merilnikom, vgrajenim v cestno telo, in z meteorološkimi merilniki.

CVP Jeprca se nahaja na ravni cesti med osrednjim in severozahodnim delom Slovenije. Teren okrog postaje je topografsko nerazgiban, najbližji objekti in gozdovi pa so oddaljeni približno 200 m. CVP Mislinja (slika 3.1) se nahaja v manjši dolini ob zaviti in vzpenjajoči se regionalni cesti na severnem delu Slovenije. Postaja je obdana z drevesi in bližnjimi objekti. CVP Črmošnjice se nahaja na začetku manjše vasi v jugovzhodnem delu Slovenije. Topografija terena je po razgibanosti med obema prej opisanimi postajama.

Vse tri postaje so redno vzdrževane in zagotavljajo podatke meritev brez večjih izgub. Za ilustracijo so na sliki 3.2 prikazane meritve temperature cestišča za 8 zaporednih dni

*Slika 3.1*Cestnovremenska postaja  
Mislinja.

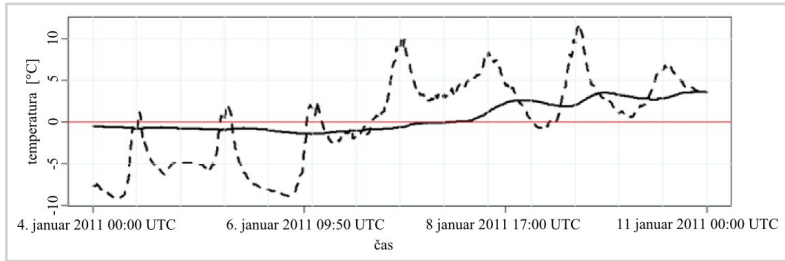
(od 4. do 11. januarja 2011) na CVP Mislinja.

Za vremenske napovedi smo uporabili 6-urne napovedi meteoroloških modelov nad slovenskim območjem. Uporabili smo podatke sistema za izračun zelo kratkoročnih vremenskih napovedi v visoki prostorski (1 km) in časovni (do 12 ur) ločljivosti INCA [50]. INCA analize in kratkoročne napovedi vključujejo temperaturo, zračno vlažnost, veter in količino ter tip padavin. Poleg INCA podatkov pa smo iz numeričnega napovednega modela ALADIN z večjo prostorsko ločljivostjo (9,5 km) uporabili napovedi dolgovalovnega in kratkovalovnega sevanja, zračnega tlaka in oblačnosti. ♣ Vsi vhodni modelirni podatki so predstavljeni v tabeli 3.1.

Za učenje smo uporabili podatke iz zime 2009/2010 (od 1. decembra 2009 do 1. aprila 2010), za testiranje pa smo uporabili podatke iz naslednje zime (od 1. oktobra 2010 do 1. februarja 2011). Uporabili smo vse podatke urnih meritev na izbranih CVP in meteorološke podatke INCA/ALADIN napovednih modelov na lokaciji CVP. Podatke smo standardizirali tako, da smo jim odšteli njihovo povprečje in jih delili z dvakratnikom njihove standardne deviacije.

Slika 3.2

Temperatura cestišča na površini (prekinjena krivulja) in temperatura cestišča v globini (neprekinjena krivulja) na CVP Mišinja med 4. in 11. januarjem 2011.



### 3.1.2 Podatki za analizo povezav med modeli in lastnostmi objektov

Za analizo povezav med modeli in lastnostmi objektov, na katere se ti modeli nanašajo, smo pripravili podatke od 18. 1. 2011 do konca zimskega obdobja leta 2012. Te smo razdelili na učno množico, ki je vsebovala podatke za januar, februar, marec, oktober in november 2011, in testno množico, ki je vsebovala preostale podatke do vključno februarja 2012. Ker je lahko modelov, ki se nanašajo na določeno lokacijo, večje število in je lahko za posamezne modele na voljo malo podatkov, je smiselno modele medsebojno združiti, kot je to opisano v razdelku 4.1, in s tem zmanjšati njihovo število. V našem primeru smo za vsako lokacijo pripravili dva modela, enega nočnega za opisovanje temperature cestišča ponoči in enega dnevnega za opisovanje temperature cestišča podnevi. Podrobnosti takšnega združevanja so opisane v razdelku 4.1.3. V tabelah 3.2 in 3.3 so navedeni regresijski koeficienti dnevnih in nočnih modelov, ki se nanašajo na posamezne lokacije.

Pripravili smo podatke desetih CVP na državnih cestah (za te postaje velja, da se ne nahajajo na viaduktu) in desetih CVP na avtocestah (za te postaje velja, da se nahajajo na viaduktu). Za vsako CVP smo opredelili, ali se nahaja na viaduktu,<sup>1</sup> pridobili njeno višino, opredelili vodne vire v bližini in ocenili vidnost neba na lokaciji CVP. Navedene lastnosti smo izbrali na podlagi navedb v literaturi, kjer so opisane kot lastnosti z največjim vplivom na temperaturo cestišča<sup>2</sup>. Podatke smo pridobili s pomočjo sistema VIPOS [43], iz fotografij okolice CVP, iz satelitskih posnetkov in z uporabo različnih vrst zemljevidov. Podatki so zbrani v tabeli 3.4.

<sup>1</sup>V podatku, ali se CVP nahaja na viaduktu ali ne, se skriva glavna ločnica v strukturi cestišča na lokaciji CVP, čeprav se znotraj te razdelitve seveda pojavljajo (manjše) razlike v razporeditvi, debelini in fizikalnih lastnosti materialov cestišča.

<sup>2</sup>Prim. razdelek 2.2.3 na strani 18.

Tabela 3.1

Okrajšave vhodnih spremenljivk in njihov pomen.

Okrajšava	Pomen
Meritve CVP	
MA	temperatura zraka
MR	temperatura cestišča (na površini)
MD	temperatura v globini cestišča (na 30 cm)
MH	vlažnost zraka
MF	debelina vodnega filma
Meteorološke napovedi za 6 ur vnaprej	
FA <sub>3</sub>	temperatura zraka (za 3 ure vnaprej)
FA <sub>6</sub>	temperatura zraka (za 6 ur vnaprej)
FH	vlažnost zraka
FW	hitrost vetra
FP	količina padavin
FL	dolgovalovno sevanje
FS	kratkovalovno sevanje
FC	oblačnost

Tabela 3.2

Regresijski koeficienti dnevni modelov, ki se nanašajo na posamezne lokacije ( $M_{it}$ ). IN pomeni prosti člen regresije, ostale oznake so obrazložene v tabeli 3.1.

CVP	IN	MA	MR	MD	FA <sub>6</sub>	FS
Kranj	7,26	3,09	-2,73	1,10	2,81	3,03
Jeprca	7,57	1,78	-1,23	0,31	3,47	3,16
Trojane	5,75	1,24	1,26	0,40	2,64	3,25
Rimske	6,68	0,86	1,61	-0,38	2,93	2,47
Črnova	5,61	1,93	-1,55	1,61	2,91	2,32
Črni vrh	-1,12	0,39	2,90	-0,13	0,75	-0,19
Soteska	5,20	0,38	0,61	1,21	2,31	3,09
Dolsko	8,42	1,90	-1,76	0,69	3,93	3,42
Mislinja	7,92	1,13	-2,09	2,21	3,59	3,11
Črmošnjice	7,48	1,94	-1,75	1,79	3,10	3,14
Kresnica	9,84	0,43	-1,26	0,00	4,95	3,84
Črni Kal	14,02	1,74	-0,61	0,00	4,52	3,53
Devina	9,78	3,53	-3,15	0,00	5,16	3,82
Ravbark.	7,73	0,67	-0,13	0,00	5,30	4,08
Malence	7,90	1,75	-1,45	0,00	4,82	3,42
Slatina	9,08	0,80	-0,61	0,00	5,84	3,35
Verd	7,15	2,77	-1,07	0,00	3,59	3,03
Sava	8,17	2,90	-2,57	0,00	4,32	2,77
Peračica	5,04	-0,65	1,23	0,00	3,58	2,31
Moste	9,53	3,30	-2,06	0,00	3,34	4,17

Tabela 3.3

Regresijski koeficienti nočnih modelov, ki se nanašajo na posamezne lokacije ( $M_{ij}$ ). IN pomeni prosti člen regresije, ostale oznake so obrazložene v tabeli 3.1.

CVP	IN	MA	MR	MD	FA <sub>6</sub>	FS
Kranj	0,74	-0,73	1,31	0,88	2,60	-
Jeprca	0,06	-1,81	2,93	0,20	3,08	-
Trojane	-0,41	-1,66	3,62	0,17	2,31	-
Rimske	1,33	-2,08	4,17	0,67	1,30	-
Črnova	-0,30	0,31	1,66	1,22	1,19	-
Črni vrh	-3,06	-0,53	3,98	0,00	0,69	-
Soteska	-0,77	0,52	1,86	1,01	0,65	-
Dolsko	0,96	-1,04	1,26	1,40	2,89	-
Mislinja	-0,28	0,18	2,20	0,75	1,09	-
Črmošnjice	-0,22	0,15	1,90	1,00	1,37	-
Kresnica	0,32	-1,74	3,59	0,00	2,23	-
Črni Kal	3,40	1,09	1,14	0,00	2,66	-
Devina	-0,68	-1,61	2,96	0,00	3,51	-
Ravbark.	-1,70	-0,18	2,34	0,00	3,10	-
Malence	0,07	0,59	1,07	0,00	3,09	-
Slatina	0,25	-1,03	3,52	0,00	2,77	-
Verd	0,47	-1,48	3,08	0,00	3,18	-
Sava	0,78	-2,46	3,30	0,00	3,65	-
Peračica	-3,14	-2,44	4,10	0,00	2,16	-
Moste	-0,60	-0,08	2,10	0,00	2,61	-

Tabela 3.4

Podatki o lastnostih lokacij CVP ( $L$ ), na katere se nanašajo modeli temperature cestišča.  $\phi$  predstavlja zemljepisno širino,  $\lambda$  zemljepisno dolžino,  $h$  nadmorsko višino in  $\Psi_s$  faktor vidnosti neba (bodisi za celotno nebo bodisi za posamezne strani neba: S–sever, J–jug, V–vzhod, Z–zahod). V primeru binarnih vrednosti pomeni 1 *da* in 0 *ne*. Vpliv posameznih lastnosti je opisan v razdelku 2.2.3. Lastnosti do vertikalne črte so informativne in jih v analizo nismo vključevali.

#	Ime	Lastnik	$\phi$	$\lambda$	$h$	Kopno	Vodni vir	$\Psi_s$	$\Psi_s(S)$	$\Psi_s(J)$	$\Psi_s(V)$	$\Psi_s(Z)$
1	Kranj	DRSC	46,229	14,361	380	1	1,00	0,80	0,75	0,85	0,85	0,75
2	Jeprca	DRSC	46,184	14,379	360	1	0,10	0,94	0,95	0,90	1,00	0,90
3	Trojane	DRSC	46,206	14,901	480	1	0,30	0,50	0,40	0,60	0,55	0,45
4	Rimske Toplice	DRSC	46,121	15,202	210	1	0,90	0,50	0,60	0,40	0,50	0,40
5	Črnova	DRSC	46,314	15,174	304	1	0,60	0,44	0,45	0,50	0,40	0,40
6	Črni vrh	DRSC	45,914	14,032	852	1	0,40	0,33	0,40	0,30	0,30	0,30
7	Soteska	DRSC	46,312	14,063	515	1	1,00	0,35	0,50	0,40	0,30	0,30
8	Dolsko	DRSC	46,096	14,680	230	1	0,10	0,60	0,50	0,60	0,50	0,80
9	Mislinja	DRSC	46,436	15,190	540	1	0,50	0,45	0,40	0,50	0,40	0,50
10	Črmošnjice	DRSC	45,671	15,102	415	1	0,30	0,51	0,55	0,45	0,55	0,50
11	Kresnica	DARS	46,662	15,646	318	0	0,10	0,78	0,75	0,75	0,75	0,70
12	Črni Kal	DARS	45,553	13,864	210	0	0,10	0,86	0,80	0,90	0,85	0,90
13	Devina	DARS	46,399	15,604	264	0	0,35	0,88	0,80	0,95	0,95	0,80
14	Ravbark.	DARS	45,791	14,228	602	0	0,20	0,61	0,70	0,50	0,60	0,65
15	Malence	DARS	46,014	14,557	320	0	0,20	0,69	0,60	0,70	0,70	0,75
16	Slatina	DARS	46,284	15,427	345	0	0,10	0,65	0,52	0,65	0,72	0,72
17	Verd	DARS	45,955	14,305	309	0	0,20	0,75	0,85	0,60	0,80	0,75
18	Sava Šentjakob	DARS	46,085	14,584	277	0	1,00	0,88	0,90	0,90	0,80	0,90
19	Peračica	DARS	46,320	14,241	450	0	0,30	0,71	0,75	0,80	0,60	0,70
20	Moste	DARS	46,408	14,126	532	0	0,90	0,50	0,40	0,60	0,50	0,50



### 3.2 Priprava umetnih podatkov

Pri pripravi umetnih podatkov smo poizkusili doseči, da bi ti imeli podobne lastnosti kot podatki za napovedovanje temperature cestišča. Naša osnovna predpostavka je, da med modeli (te lahko opišemo s spremenljivkami, na primer regresijskimi koeficienti) in lastnostmi modelov (te opišemo z atributi), obstajajo povezave. Tovrstne povezave so podrobneje predstavljene v razdelku 4.3. Umetne podatke smo uporabili tako pri preizkusu metode za gradnjo hierarhičnih regresijskih modelov (razdelek 4.1.2) kot pri preizkusu metode za analizo povezav (razdelek 4.3.2). Za lažje razumevanje postopka priprave umetnih podatkov je v dodatku A podana programska koda v jeziku *Python*.

Pri pripravi umetnih podatkov smo najprej generirali attribute — naključne vrednosti, ki predstavljajo lastnosti objektov, katere se nanašajo na modele. Naključne vrednosti smo generirali na podlagi normalne porazdelitve  $N(0, 1)$  in jih predstavili v matriki atributov  $A$ . Število vrstic v matriki  $A$  je odvisno od števila modelov, število stolpcev pa od števila lastnosti objektov. Pri tem so lahko lastnosti objektov zvezne ali pa preslikane na diskretne vrednosti.

Sledila je priprava matrike  $P$ , ki predstavlja povezave med matriko  $A$  in med koeficienti spremenljivk modelov (vsak model ima svoje koeficiente). Število vrstic v matriki  $P$  je odvisno od števila spremenljivk modelov, število stolpcev pa od števila lastnosti objektov. Rezultat množenja matrike  $P$  z matriko  $A$  nam poda vrednosti koeficientov. Tako je vsak koeficient predstavljen z linearno kombinacijo lastnosti objektov, na koncu pa mu dodamo še šum — naključno število z normalno porazdelitvijo  $N(0, \sigma^2)$ .

Matrika  $P$  je lahko samo ena in je tako enaka za vse modele. Ker pa imajo različni modeli različne lastnosti, so koeficienti med modeli med seboj različni. Mogoče pa je uporabiti več matrik  $P$ , vsako za svoj nabor modelov. To bi pomenilo, da so povezave med modeli in lastnostmi teh modelov med različnimi modeli različne. Predvidevamo, da do takšnega pojava prihaja tudi pri podatkih za napovedovanje temperature cestišča, kar je predstavljeno v poglavju 4.

Sledila je priprava vrednosti spremenljivk vsakega modela — naključnega števila z normalno porazdelitvijo  $N(0, 1)$ . S pomočjo koeficientov modelov in vrednosti njihovih odvisnih spremenljivk smo lahko s skalarnim produktom izračunali neodvisno spremenljivko  $y$  in ji na koncu dodali še šum — naključno število z normalno porazdelitvijo  $N(0, \sigma^2)$ .

Na ta način lahko pripravimo podatke, kjer je mogoče vplivati tako na poljubno

število modelov, koeficientov oziroma spremenljivk in lastnosti objektov (vključno s številom zveznih in diskretnih vrednosti), kot tudi na varianco šuma pri izračunavanju koeficientov in varianco šuma pri izračunavanju odvisne spremenljivke  $y$ . Takšni podatki imajo pomembno lastnost: koeficienti modelov so, glede na matriko  $P$ , povezani z lastnostmi objektov, ki se nanašajo na modele.

# *Opis dela*

Poglavje predstavlja osrednji del doktorske disertacije in je razdeljeno na tri glavne razdelke. Prvi razdelek opisuje metodo za gradnjo hierarhičnih regresijskih modelov, ki temelji na združevanju vnaprej podanih množic podatkov in njeno uporabo na umetnih podatkih oziroma na problemu napovedovanja temperature cestišča. Pri slednjem je navedena napaka združenih modelov, s katerimi smo napovedovali temperaturo cestišča. To smo primerjali z napako povsem združenih modelov in z napako najbolj razširjenega fizikalnega modela METRo in pokazali, da so rezultati predlagane metode primerljivi ali boljši. Drugi razdelek opisuje dekompozicijo napake na pristranskost in varianco v prejšnjem razdelku predstavljene metode. Opisana je prilagoditev postopka dekompozicije za predlagano metodo in njen preizkus na aplikativnih podatkih. Tretji razdelek analizira povezave med modeli in lastnostmi objektov, na katere se ti modeli nanašajo. Na problemu napovedovanja temperature cestišča smo z obstoječimi postopki poizkušali te povezave nakazati. Sledi opis metode za analizo povezav, ki predstavlja nadgradnjo metode iz prvega razdelka. Metodo smo uporabili na umetnih podatkih in poizkušali poiskati matriko  $P$ , ter jo primerjati z matriko, ki je bila uporabljena pri generiranju podatkov. Metodo smo na drugačen način uporabili tudi na aplikativnih podatkih, kjer smo iskali koeficiente modelov iz lastnosti teh modelov in primerjali rezultate napovedi s poiskanimi koeficienti z napovedmi, kjer smo koeficiente izračunali na podlagi podatkov modelov. To smo naredili z ločenimi dnevnimi in nočnimi modeli za vsako CVP; poleg tega pa smo te rezultate primerjali še z modeli, ki so bili skupni vsem CVP (en dnevni model, en nočni model, en dnevno-nočni model).

#### *4.1 Metoda za gradnjo hierarhičnih regresijskih modelov*

V nadaljevanju je predstavljena metoda za gradnjo hierarhičnih regresijskih modelov, ki temelji na združevanju vnaprej podanih množic podatkov na podlagi manjšanja (oziroma nesignifikantnega večanja) njihovih ocenjenih napak. Predlagana metoda temelji na statističnih postopkih, saj uporablja klasične regresijske modele več spremenljivk za opisovanje množic podatkov. ¶ Z vidika robustnosti je najbolj primerno uporabiti linearne regresijske modele, četudi povezave med dejavniki (spremenljivkami) niso linearne. Metoda združi posamezne modele in s tem zmanjšuje število končnih modelov ter posledično tudi napovedno napako na učni množici. Število končnih modelov in njihove meje so odvisne zgolj od podatkov in niso vnaprej določene. Naši raziskovalni hipotezi sta, da lahko končni modeli ustrezno ločijo ključne mehanizme, ki so tipični v določeni okolici sosednjih modelov ter da lahko s takšnim postopkom kljub poenosta-

vljeni obliki (linearizacija) dosežemo zadovoljive rezultate. ¶ Metodo smo preizkusili tako na umetno generiranih podatkih kot tudi na problemu napovedovanja temperaturnega stanja cestišča in pridobili več končnih modelov, ki so ustrezno opisovali procese spreminjanja temperature cestišča. Po primerjavi z že uveljavljenim načinom napovedovanja temperature cestišča smo ugotovili, da je predlagana metoda združevanja uspešna in predstavlja ustrezno alternativo fizikalnim napovednim modelom. Metoda je na področju cestne meteorologije v splošnem prenosljiva na različna geografska območja, saj so osnovni procesi, ki se odvijajo v ozadju, prostorsko neodvisni. ¶ Tako opis metode kot njen praktični preizkus je bil predstavljen na mednarodni konferenci *International Symposium on Electronics in Transport* [63] in v znanstveni reviji *Meteorological Applications* [65].

#### 4.1.1 Opis metode združevanja

Klasični linearni regresijski model lahko v matričnem zapisu podamo kot:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad (4.1)$$

kjer je  $\mathbf{y}$  vektor odvisne spremenljivke velikosti  $n \times 1$ ,  $\mathbf{X}$  matrika neodvisnih spremenljivk velikosti  $n \times p$ ,  $\boldsymbol{\beta}$  vektor koeficientov velikosti  $p \times 1$  in  $\mathbf{e}$  vektor napak (za napake predvidevamo, da so neodvisne in normalno porazdeljene s srednjo vrednostjo 0 in standardno deviacijo  $\rho$ ) velikosti  $n \times 1$ ;  $p$  je število spremenljivk in  $n$  število podatkov.

Ustrezno izbiro vhodnih spremenljivk lahko dosežemo z uporabo postopne (angleško *stepwise*) regresijske analize, ki izbere podmnožico podanih spremenljivk na podlagi izbranega informacijskega kriterija, npr. Akaikov informacijski kriterij [1]:

$$AIC = 2v - 2\log(L), \quad (4.2)$$

kjer  $v$  predstavlja število spremenljivk v modelu in  $L$  verjetje modela. Postopna regresija poišče podmnožico spremenljivk, ki minimizira 4.2.

Ko imamo za posamezne modele na voljo premalo podatkov, poleg tega pa ti vsebujejo kar nekaj naključnosti, je smiselno izvesti združevanje modelov. Običajni pristopi združevanja temeljijo na gradnji od zgoraj navzdol, torej postopnem deljenju množice primerov na podlagi vrednosti posameznih spremenljivk z namenom izboljšanja napovedne točnosti. Gradnja od spodaj navzgor, z združevanjem, je pri regresijskih drevesih redka in navadno ni primerna, saj je težko določiti primerne začetne množice.

V nekaterih primerih pa imamo začetne množice podane oziroma so podatki, katere želimo modelirati, naravno razdeljeni v več skupin. Za združevanje lahko določimo, kateri modeli so med seboj združljivi, z drugimi besedami, določiti želimo njihovo sosednost. Pri tem lahko uporabljamo ekspertno znanje o podobnostih modelov, na podlagi katerega izdelamo matriko sosednosti, ki za vsak par modelov opredeljuje njuno sosednost. Na ta način lahko, na primer, vpeljemo časovno lokalnost pri združevanju modelov (dovolimo le združevanje časovno sosednjih modelov) in periodičnost procesa (dovolimo združevanje zadnjega in prvega modela). Matriko sosednosti bi lahko pridobili tudi z opazovanjem podobnosti posameznih skupin podatkov, denimo njihovih lastnosti<sup>1</sup>, v skrajnem primeru pa se ji celo izognili, kar bi pomenilo, da bi metoda v vsakem koraku združevanja preverila vse možne kombinacije združevanj.

Naj bodo  $M_0, M_1, \dots, M_t$  množice modelov (število vseh množic je  $t + 1$ ) in  $P_0, P_1, \dots, P_t$  množice učnih podatkov teh modelov. Sosednost podatkov in s tem tudi modelov je določena tako, da sta si modela  $M_i$  in  $M_{i+1}$ ,  $i = 0, \dots, t$ , med seboj sosednja, zaradi upoštevanja periodičnosti pa velja  $t + 1 = 0$  oziroma seštevanje po modulu  $t + 1$ . Naj bo  $E_i$  izbrana mera napake modela  $M_i$ , na primer korenjena srednja kvadratična (v nadaljevanju RMS<sup>2</sup>) napaka, izračunana po metodi navzkrižnega preverjanja na podatkih učne množice.

V vsakem koraku združevanja izdelamo množice modelov  $M_{0,1}, M_{1,2}, \dots, M_{t-1,t}, M_{t,0}$  na unijah sosednjih množic  $P_0 \cup P_1, P_1 \cup P_2, \dots, P_t \cup P_0$ , na podlagi katerih bo kasneje temeljila odločitev, katere modele bomo združili. RMS napako posameznih združenih modelov označimo z  $E_{i,i+1}$  in jo primerjamo z uteženo povprečno napako teh dveh modelov pred združevanjem:

$$E'_{i,i+1} = \frac{E_i \cdot |S_i| + E_{i+1} \cdot |S_{i+1}|}{|S_i| + |S_{i+1}|}, \quad (4.3)$$

kjer je  $|S_i|$  moč učne množice  $S_i$  za model  $M_i$ . Takšna primerjava nam omogoča opazovanje, koliko je združevanje vplivalo na spremembo napake  $E$  v primerjavi s predvideno napako  $E'$ .<sup>3</sup>

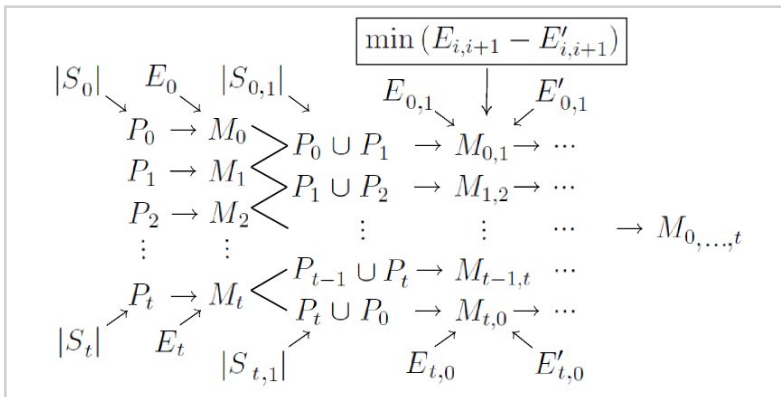
<sup>1</sup>Prim. razdelek 4.3 na strani 67.

<sup>2</sup>Prim. enačbo 2.6 na strani 9.

<sup>3</sup>Uporaba razlike med pričakovano napako pred in po združevanju za poenostavljanje modela je podobno naknadnemu rezanju dreves z vzvratno oceno napake (npr. rezanje z m-oceno napake). Razlika je v tem, da pri rezanju dreves poenostavljamo že inducirano drevo, v predlaganem postopku pa z združevanjem drevo šele gradimo; rezanje dreves spreminja notranja vozlišča v liste, z združevanjem pa izbiramo med različnimi možnimi notranjimi vozlišči.

V vsakem koraku določimo tista dva sosednja modela, kjer se je napaka  $E_{i,i+1} - E'_{i,i+1}$  najmanj povečala<sup>4</sup> (slika 4.1). Ko določimo  $\min(E_{i,i+1} - E'_{i,i+1})$ , združimo modela, pri katerih se ta minimum pojavi in ponavljamo postopek, ki sedaj zaradi združevanja vsebuje en model manj. Razlika  $E_{i,i+1} - E'_{i,i+1}$  je za vsak izbran par v prvih korakih združevanja po pričakovanju negativna (združevanje povečuje natančnost), medtem ko se bo pri kasnejših korakih predvidoma povečevala. Postopek se zaključi, ko združimo vse modele.

Opisan postopek gradnje in združevanja modelov je v osnovi podoben hierarhičnemu gručenju, vendar s specifično mero za izbor gruč, ki ju združimo v posameznem koraku (ocenjeno zmanjšanje napake). Le-ta služi tudi kot kriterij za ustavljanje postopka; končno število modelov določimo z opazovanjem povprečne napake  $E$  vseh modelov v posameznem koraku združevanja. Z združevanjem prenehamo, ko ta napaka preseže vrednost vnaprej določene dopustne napake. Opazujemo lahko tudi število modelov v posameznem koraku združevanja. V tem primeru z združevanjem prenehamo, ko je doseženo vnaprej določeno najmanjše število modelov.



Slika 4.1

Shematski prikaz prvega koraka postopka združevanja. Znak  $>$  predstavlja združevanje. Ko določimo par  $M_i$  in  $M_{i+1}$ , za katerega velja  $\min(E_{i,i+1} - E'_{i,i+1})$ , združimo množice učnih podatkov za modela  $M_i$  in  $M_{i+1}$  in začnemo z naslednjim korakom združevanja. Postopek se zaključi najkasneje, ko pridemo do končnega modela  $M_{0,\dots,t}$ .

#### 4.1.2 Preizkus na umetnih podatkih

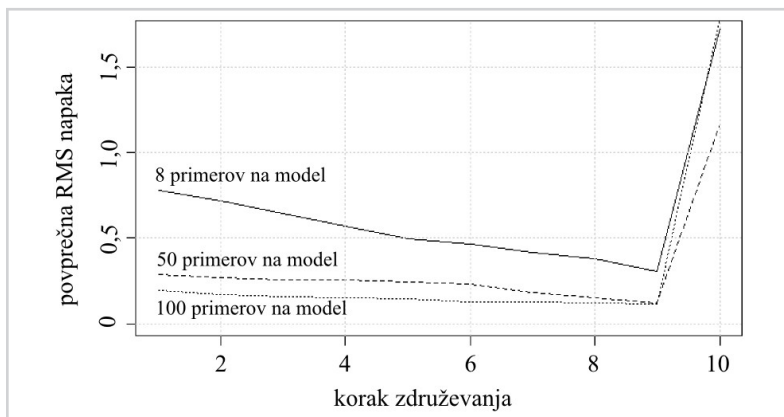
Metodo za gradnjo hierarhičnih modelov smo preizkusili na umetnih podatkih<sup>5</sup> z dve-ma skupinama modelov, kjer ima vsaka skupina modelov svojo matriko  $P$ . Pripravili

<sup>4</sup>Napaka  $E_{i,i+1} - E'_{i,i+1}$  je lahko tudi negativna.

<sup>5</sup>Prim. razdelek 3.2 na strani 45.

smo 10 modelov, izmed katerih jih 5 generira matrika  $P_a$ , 5 pa matrika  $P_b$ .  $P_a$  in  $P_b$  sta si med seboj različni. Vsak model ima pet spremenljivk, lastnosti modelov (atributi) pa so štirje. varianca pri šumu za koeficiente je 0,2, varianca pri šumu za odvisno spremenljivko pa 1,0. Podobno kot pri podatkih za napovedovanje temperature cestišča, katere bomo uporabili pozneje, so modeli med seboj ciklično povezani (1 z 2, 2 s 3, ..., 10 z 1). Zaradi tega metoda združuje zgolj sosednje modele, čeprav ta lastnost za sam postopek združevanja ni obvezna. V vsakem koraku združevanja se izračunava RMS napaka glede na trenutno število modelov.

Rezultati so po korakih združevanja prikazani na sliki 4.2, kjer je razvidno, da se povprečna RMS napaka z združevanjem manjša. Izjema so zadnji koraki, ko začne metoda združevati modele z različnimi matrikami  $P$ , kar povzroči znatno povečanje napake. V našem primeru je opazen porast napake zgolj v zadnjem koraku, ko metoda združi modela, generirana z matrikama  $P_a$  in  $P_b$ . Na sliki 4.2 so prikazani tri krivulje, ki ponazarjajo napako glede na število primerov, s katerimi smo pripravili umetne podatke (8, 50 ali 100 primerov na model).



Slika 4.2

Preizkus metode za gradnjo hierarhičnih modelov na umetnih podatkih. Abscisa os predstavlja korak združevanja, ordinata pa povprečno RMS napako. Prikazan je primer združevanja z različnim številom primerov na posamezni model.

Izkaže se, da pri večjem številu primerov združevanje ni tako koristno, ko je primerov dovolj, pa je lahko celo nepotrebno. Motivacija združevanja se kaže zlasti v tem, da ko imamo na voljo premalo podatkov za ločene modele, lahko z njihovim združevanjem dosežemo boljše rezultate. Izkaže se, da pri večjem številu primerov združevanje ni tako koristno, ko je primerov dovolj, pa je lahko celo nepotrebno. Motivacija združevanja



se kaže zlasti v tem, da ko imamo na voljo premalo podatkov za ločene modele, lahko z njihovim združevanjem dosežemo boljše rezultate.

#### 4.1.3 *Preizkus na problemu napovedovanja temperature cestišča*

Metodo za gradnjo hierarhičnih regresijskih modelov, ki temelji na združevanju vnaprej podanih množic, lahko uporabimo na problemu napovedovanja temperaturnega stanja cestišča, saj je proces spreminjanja temperature cestišča v splošnem odvisen od časa in prostora (denimo drastično povečanje temperature cestišča v primeru prisotnega dopoldanskega solarnega sevanja in manjše spremembe temperature cestišča ponoči).

V kolikor bi uporabili en sam, splošen linearni regresijski model, ki bi opisoval vse postaje in vse časovne intervale, bi bili koeficienti spremenljivk ( $\beta$ ) neodvisni od časa in prostora. Da bi se temu izognili, smo linearne modele izdelali ločeno glede na znane spremenljivke, po katerih je smiselno združevanje: lokacija CVP in ura dneva. Podatkov o podrobnostih lokacije namerno nismo vključili v modeliranje. Tovrstni podatki povzročajo v fizikalnih modelih pomemben vir netočnosti [82], saj je njihovo pridobivanje težavno in povezano s sistematičnimi napakami. V našem primeru so vse lastnosti, ki jih v modeliranje nismo vključili, implicitno vključene preko sprememb v koeficientih posameznih modelov.

V skladu z dekompozicijo napake na pristranskost in varianco imajo združeni modeli manjšo pristranskost, vendar večjo varianco<sup>6</sup>. Da bi imeli na voljo več podatkov in s tem zmanjšali varianco brez znatnega povečevanja pristranskosti, smo podobne modele, ki so si med seboj časovno sosednji, združili. Pri tem smo upoštevali, da so lahko rešitve, ki bi vsebovale več modelov, ki pokrivajo zgolj kratke časovne intervale, nezanesljive in se jim je v splošnem bolje izogniti.

Za določitev končnega števila modelov smo uporabili kriterij, ki ob čim manjši napaki poišče čim manjše končno število modelov. Poiskali smo zadnji minimum oziroma lokalni minimum z najmanjšim številom modelov, kjer RMS napaka ne preseže najmanjše povprečne RMS napake za več kot 0,5 °C (na primer, za CVP Jepca se tak minimum pojavi pri štirih modelih, kot je razvidno iz spodnjega dela slike 4.3).

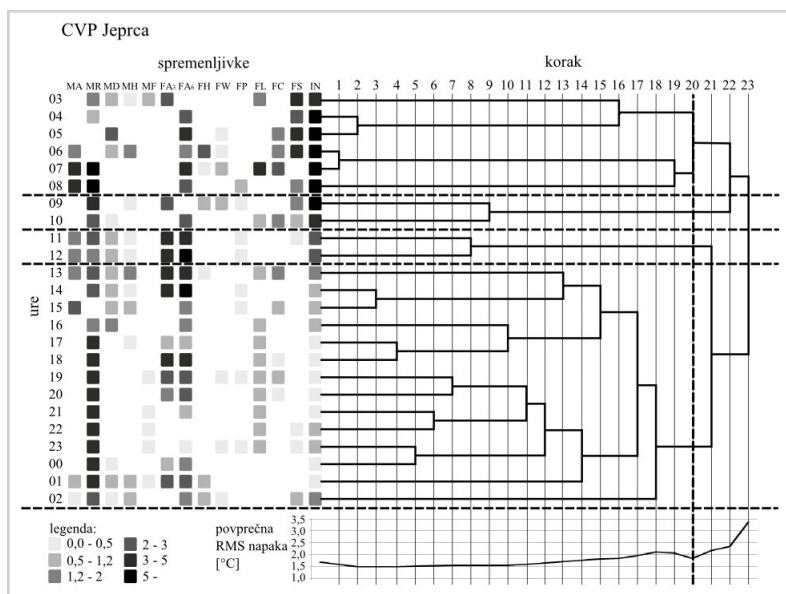
Pričakujemo, da bomo z uporabo predlagane metode dobili dva večja modela, enega za napovedovanje podnevi in drugega za ponoči, po možnosti z nekaj vmesnimi modeli za ublažitev prehodov. Poleg tega tudi predvidevamo, da bodo modeli ustrezno opisovali temperaturne spremembe na lokaciji posamezne postaje.

<sup>6</sup>Prim. razdelek 4.2 na strani 63.

*Napovedovanje temperature cestišča s pridobljenimi modeli.* Z metodo za gradnjo hierarhičnih regresijskih modelov smo želeli napovedovati temperaturo cestišča za 6 ur vnaprej. Sestavljanje modelov je potekalo v skladu z opisom predlagane metode<sup>7</sup>. Potek združevanja je razviden iz dendrogramov na slikah 4.3, 4.4 in 4.5, kjer so na levem delu prikazane pomembnosti posameznih spremenljivk (kolone) za posamezne urne modele (vrstice). Temnejše barve pomenijo večje absolutne vrednosti koeficientov linearnih modelov. Bela barva pomeni, da spremenljivka ni bila uporabljena, črna pa, da je absolutna vrednost koeficienta večja ali enaka 5. Okrajšave spremenljivk so navedene v tabeli 3.1. Pod dendrogramom se nahaja graf RMS napak pri poteku združevanja. Vertikalna črta določa zadnji korak, kjer so končni modeli ločeni s horizontalnimi črtami.

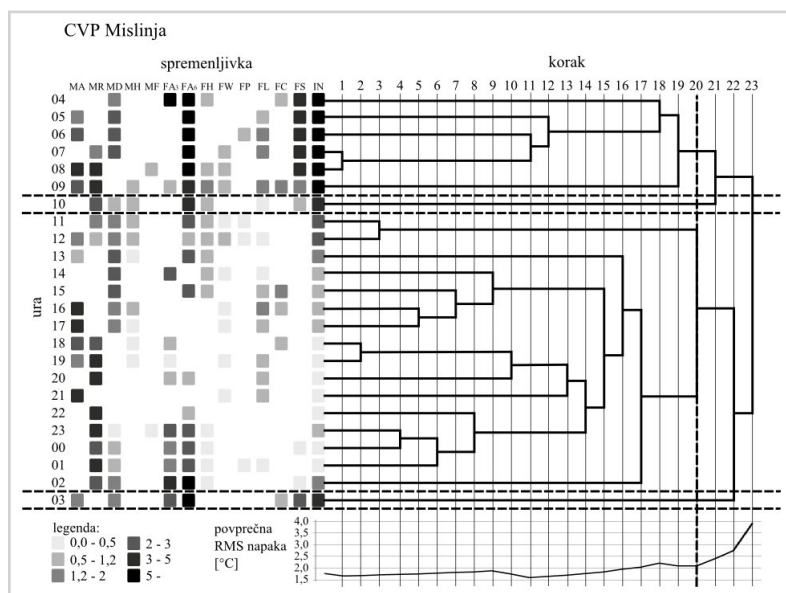
Slika 4.3

Toplotna karta [107], dendrogram s prikazom postopka združevanja in graf povprečne RMS napake na CVP Jeprca. Kvadrati na toplotni karti predstavljajo pomembnost posameznih spremenljivk. Bela barva pomeni, da spremenljivka ni bila uporabljena, črna pa, da je absolutna vrednost koeficienta večja ali enaka 5. Okrajšave spremenljivk so navedene v tabeli 3.1. IN pomeni prosti člen regresije. Vertikalne črte v tabeli predstavljajo posamezen korak združevanja, vertikalna črtkana črta pa končne modele (zadnji minimum povprečne RMS napake), ki so ločeni s horizontalnimi črtkanimi črtami. Zapis ur je v UTC in predstavlja čas izdelave napovedi za 6 ur vnaprej (na primer, "09" pomeni napoved temperature cestišča, ki je bila generirana s podatki ob 09:00 UTC za 15:00 UTC).



Metoda združevanja je za vsako CVP izdelala dva večja modela; enega za dnevne napovedne ure, ko je vpliv kratkovalovnega sevanja pomemben in ko je povprečna temperatura cestišča (prosti člen z oznako IN) visoka, in enega za nočne napovedne ure,

<sup>7</sup>Prim. razdelek 4.1.1 na strani 49.



Slika 4.4

Toplotna karta, dendrogram s prikazom postopka združevanja in graf povprečne RMS napake na CVP Mislinja. Pojasnila so enaka, kot na sliki 4.3.

ko je kratkovalovno sevanje nepomembno in je vrednost prostega člena nizka zaradi nižjih povprečnih temperatur cestišča. Med tema dvema modeloma sta dva modela, ki pokrivata časovni pas od ene do treh ur, ki blažita prehod med večjima modeloma.

V tabeli 4.1 so navedene numerične vrednosti koeficientov končnih modelov. Nekatere vhodne spremenljivke so pomembne večino časa, zlasti izmerjene in napovedane temperature zraka (spremenljivke MA, FA<sub>3</sub> in FA<sub>6</sub>), pa tudi meritve temperature cestišča (MR) in temperature cestišča v globini (MD). Slednja je bolj pomembna za CVP Mislinja in Črmošnjice, kot pa za CVP Jeprca, kar je možno pojasniti z razlikami v konstrukciji cestišča, topografiji, antropogenih vplivov ipd. v bližini CVP. Takšne razlike se odražajo tako v spremembah vrednosti koeficientov modelov za sleherno CVP kot tudi v spremembah med mejami časovnih intervalov. Na primer, na lokaciji CVP Mislinja je vpliv kratkovalovnega sončnega sevanja zaradi visokih dreves na zahodni strani blizu CVP šibkejši. To je tudi razlog, da se nočni model končne rešitve metode začne že dve uri prej kot pri ostalih dveh postajah.



*Tabela 4.1*

Izbrane vhodne spremenljivke z vrednostmi koeficientov končnih modelov. Okrajšave spremenljivk so obrazložene v tabeli 3.1. Prazne celice pomenijo neuporabljene spremenljivke.

CVP	Spremenljivka Model	MA	MR	MD	MH	MF	FA <sub>3</sub>	FA <sub>6</sub>	FH	FW	FP	FL	FC	FS	IN
Jeprca	od 03:00 do 08:00	-2,1	2,6	0,7	0,7	0,1	2,8	1,5	-1,2	-0,6	-0,4	-1,2	1,4	2,1	5,9
	od 09:00 do 10:00	1,2	3,1	0,8	0,4	-	-3,4	2,2	-0,4	-0,4	-0,3	0,3	-0,3	1,4	4,8
	od 11:00 do 12:00	0,9	2,1	0,7	0,4	-	-3,1	5,0	-	-0,2	-0,3	-	-	0,3	2,6
	od 13:00 do 02:00	-	2,4	-	0,2	-	-	4,2	-	-	-0,1	-	-	-	0,4
Mislinja	03:00	1,4	2,6	1,3	-0,6	-	-3,0	6,0	0,7	-	-0,6	0,6	1,1	2,4	3,5
	od 04:00 do 09:00	-3,6	2,1	1,4	-	-	-2,3	5,7	0,4	-	-	-0,3	-	2,7	6,8
	10:00	1,4	1,0	0,9	-	-	-	3,1	-	-	-	-	-	0,5	4,0
	od 11:00 do 02:00	-	-	1,6	-	-	-	3,3	-	-	-	-	-	-	0,6
Črmošnjice	od 03:00 do 09:00	-2,9	1,2	1,6	0,3	0,2	3,0	3,2	0,2	-0,9	-0,5	0,6	0,5	2,4	6,0
	od 10:00 do 12:00	1,8	1,4	1,2	-	0,2	-2,7	5,7	0,2	0,1	-0,4	-	0,2	0,9	2,6
	od 13:00 do 02:00	-	1,1	1,1	-	0,2	-2,1	3,0	-	-	-0,2	-	-	0,3	0,3

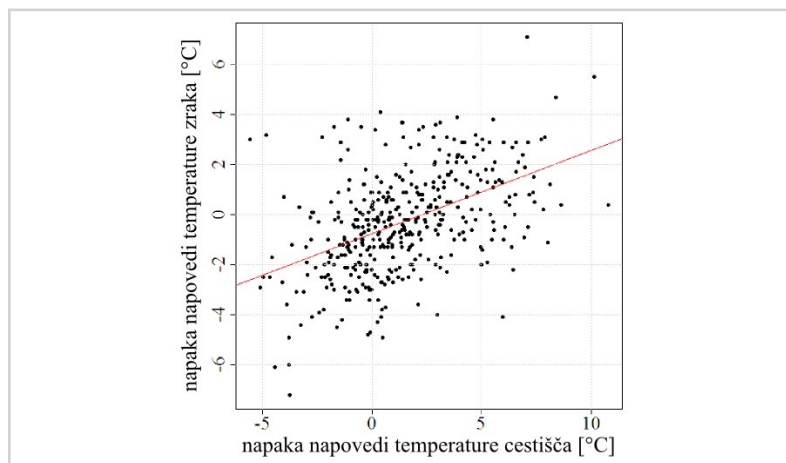
Tabela 4.2 prikazuje RMS napake končnih modelov za vse tri lokacije. Rezultat metode združevanja na CVP Jeprca so štirje modeli. Zgolj model za ure od 03:00 do 08:00 ima znaten delež napak večjih od 4 °C (15 %) kar bi pri aplikativni uporabi lahko pomenilo nesprejemljivo visok delež napak. Te napovedi sicer ustrezajo dnevnemu času, ko se temperature ceste običajno povečujejo, kar navadno ne pomeni poslabšanja varnosti na cestah. Model, ki napoveduje temperature v bolj kritičnem času (začetne ure napovedi od 13:00 do 02:00), ima visoko natančnost napovedovanja. Podobna ugotovitev je razvidna tudi na ostalih dveh CVP: napaka dnevnih modelov je večja, medtem ko je napaka večernih in nočnih modelov dokaj majhna. Zadovoljive napovedi dobimo tudi pri prehodnih modelih, celo pri CVP Mislinja, kjer eden izmed prehodnih modelov pokriva zgolj enourni časovni pas, kar pomeni, da smo za njegovo učenje imeli na voljo majhno učno množico.

Podrobnejša analiza dnevnih modelov za CVP Mislinja je pokazala, da običajno prihaja do večjih napovednih napak v primerih, ko je izmerjena temperatura cestišča večja od napovedanih temperatur. Takšne napake sicer lahko obravnavamo kot manj kritične, saj v večini primerov nastopijo, ko je temperatura cestišča izven kritičnega območja okrog 0 °C, še vedno pa lahko predstavljajo potencialno nevarnost, saj zaradi pojava termalnega spomina<sup>8</sup> vplivajo tudi na napake pri napovedovanju večernih temperatur cestišča.

Večje napovedne napake so običajno povezane z večjo spremenljivostjo kratkovalovnega sevanja in oblačnosti. Ustrezne vremenske napovedi, zlasti kratkovalovnega sevanja, so zelo pomembne za modele, ki imajo začetek napovedi od 04:00 do 09:00. Žal CVP niso opremljene z merilniki kratkovalovnega sevanja, zato ni bilo mogoče neposredno preveriti korelacije med napako vremenskih napovedi in razliko med napovedanimi in izmerjenimi vrednostmi kratkovalovnega sevanja. Kljub temu podatki nakazujejo na majhno pozitivno korelacijo ( $R^2 = 0,22$ ) med napako modela in napako napovedane temperature zraka (slika 4.6), kar lahko pomeni, da je napaka pri napovedi temperature cestišča povezana z napako vremenskih napovedi. Vsekakor bi za potrditev naše predpostavke potrebovali več podatkov, vendar pa je naša domneva v skladu z ugotovitvami raziskovalcev [64], ki so ob uporabi fizikalnega modela METRo opazili bistveno izboljšanje napovedi v primerih, ko so uporabili vremenske napovedi z manjšo napako temperature zraka.

Rezultate predlagane metode združevanja smo primerjali z rezultati linearnih regre-

<sup>8</sup>Prim. razdelek 2.2.3 na strani 18.



Slika 4.6

Korelacija med napako modela in napako napovedane temperature zraka.

sijskih modelov za celotni dan za vsako posamezno postajo. RMS napaka teh modelov na testni množici je bila 3,37 za CVP Jeprca, 3,92 za CVP Mislinja in 3,92 za CVP Črmošnjice. Te napake so približno dvakrat večje kot napake predlagane metode.


*Primerjava rezultatov s fizikalnim modelom.* Naša predpostavka je, da je predlagana metoda bolj točna od obstoječih fizikalnih modelov, za katere je znano, da je parametrizacija nekaterih dejavnikov, ki pomembno vplivajo na rezultate, zapletena. Predpostavko smo preverili s primerjavo predlagane metode in fizikalnega modela METRo [29] pri enakih vhodnih podatkih (gre za podatke v zimskem času 2009/2010 in 2010/2011). Model METRo smo vzpostavili na osebнем računalniku z operacijskim sistemom *Ubuntu*, na katerem smo namestili dodatne programske pakete (python, gfortran, numpy, swig, java6-jre). V programskem jeziku *Java* smo implementirali vmesnik za pripravo ustreznih podatkov (za lokacijo postaje: časovni pas, zemljepisna širina in dolžina, sestava cestnega telesa (plasti in debelina); INCA/ALADIN vremenske napovedi; izmerjeni vremenski podatki) glede na vhodne sheme modela METRo. Primer XML vhodne datoteke s podatki o CVP Jeprca je podan v nadaljevanju.

```
<?xml version="1.0"?>
<station>
  <header>
```

```

<version>1.0</version>
<time-zone>UTC+1</time-zone>
<name>Jeprca</name>
<station-type>road</station-type>
<coordinate>
  <latitude>46.1846</latitude>
  <longitude>14.3789</longitude>
</coordinate>
</header>
<roadlayer-list>
  <roadlayer>
    <position>1</position>
    <type>asphalt</type>
    <thickness>0.14</thickness>
  </roadlayer>
  <roadlayer>
    <position>2</position>
    <type>sand</type>
    <thickness>0.30</thickness>
  </roadlayer>
  <roadlayer>
    <position>3</position>
    <type>crushed rock</type>
    <thickness>0.50</thickness>
  </roadlayer>
</roadlayer-list>
</station>

```

Natančne podatke o strukturi in fizikalnih lastnostih materialov cestnega telesa na lokaciji CVP je mogoče pridobiti zgolj z izvajanjem vrtnja v cestno telo in laboratorijsko analizo. Namesto tega smo iz banke cestnih podatkov pridobili standardno sestavo cestnega telesa na lokaciji CVP. Takšno ravnanje je sicer običajno, vendar pa lahko poslabša rezultate modela METRo. Žal nam študije, ki bi opisovale vpliv netočnih vhodnih podatkov na delovanje modela METRo, niso poznane.  Napake modela METRo, ki smo jih pridobili na vseh treh CVP ustrezajo ugotovitvam drugih razisko-



valcev [29], ki so testirali model METRo na CVP v Kanadi: približno polovica dnevnih napovedi temperatur se nahaja v območju napake  $\pm 2$  K. Raziskovalci navajajo tudi, da se nočna RMS napaka giblje okrog 2 K, medtem ko so rezultati modela METRo na slovenskih CVP nekoliko slabši: 2,46 za CVP Jeprca, 3,26 za CVP Mislinja in 2,37 za CVP Črmošnjice.

Rezultati modela METRo so razdeljeni v časovne intervale, ki ustrezajo razdelitvi predlagane metode za posamezno CVP in so prikazani v tabeli 4.2. Njegovi rezultati so v primerjavi s predlagano metodo slabši, razen pri nekaterih krajših prehodnih časovnih intervalih ter pri dnevnem modelu za CVP Mislinja. ¶ Opravili smo Wilcoxonov [106] test na absolutnih vrednostih napak za oba pristopa in dobili zelo nizke  $p$  vrednosti na vseh treh postajah, iz česar lahko sklepamo, da so rezultati predlagane metode primerljivi ali boljši od rezultatov modela METRo, ki velja za najbolj razširjen in ustrezen fizikalni model.

*Tabela 4.2*

Napake 6-urnih napovedi temperature cestišča z uporabo predlagane metode (oznake S) in napake fizikalnega modela METRo (oznake F). Povprečne napake (oznake P) za vsako CVP so utežene glede na velikost vseh časovnih intervalov.

CVP	Časovni interval modela (UTC ure začetka napovedi)	RMS napake na testni množici		Odstotek napovedi na testni množici z napako večjo od 2 °C		Odstotek napovedi na testni množici z napako večjo od 4 °C	
		S	F	S	F	S	F
Jeprca	od 03:00 do 08:00	2,86	3,42	43	69	15	32
	od 09:00 do 10:00	1,63	1,53	22	16	2	4
	od 11:00 do 12:00	1,41	2,46	16	48	0	12
	od 13:00 do 02:00	1,55	2,86	20	56	0	20
	od 00:00 do 23:00 P	1,88	2,86	26	55	4	21
Mislinja	03:00	1,99	3,00	30	64	5	23
	od 04:00 do 09:00	3,21	3,09	48	56	21	26
	10:00	1,70	2,23	24	48	1	6
	od 11:00 do 02:00	1,33	3,26	13	73	0	26
	od 00:00 do 23:00 P	1,84	3,16	23	67	6	25
Črmošnjice	od 03:00 do 09:00	2,65	3,29	42	66	15	31
	od 10:00 do 12:00	1,49	1,46	17	17	1	3
	od 13:00 do 02:00	1,30	2,37	12	41	0	10
	od 00:00 do 23:00 P	1,72	2,52	21	45	5	15

*Interpretacija rezultatov.* Predlagana metoda združevanja se je izkazala kot uspešna. Za vsako postajo smo dobili več končnih modelov, ki ustrezno opisujejo procese spreminjanja temperature cestišča. Medtem ko so bile dnevne napovedi opisane z več modeli, je nočne napovedi odlično opisoval le en model. Še vedno pa je prisotna dokaj velika napaka modela za napovedi jutranjih temperatur cestišča, ko so procesi spreminjanja temperature najizrazitejši.

Uporaba fizikalnih modelov za napovedovanje stanja cestišča je pogojena s poznavanjem podatkov o lokaciji napovedovanja, katere pridobivanje je pogosto težavno in povezano z napakami. Naša predpostavka je, da se je tovrstnemu problemu mogoče izogniti z uporabo preteklih izmerjenih podatkov ter z izdelavo več modelov za vsako lokacijo napovedovanja. Eksperimenti to raziskovalno hipotezo potrjujejo in kažejo, da lahko s predlagano metodo v večini primerov pridemo do boljnih rezultatov v primerjavi s fizikalnim modelom.

Omejitev uporabljene metode je, da za ustrezno izdelavo modelov potrebuje podatke nekaj zimskih mesecev, vendar pa lahko svoje rezultate vseskozi izboljšuje s pomočjo novih podatkov. Po drugi strani lahko fizikalni model uporabimo že z zgolj nekaj urami zgodovinskih podatkov, ki so potrebni za postopek sklapljanja vremenskih napovedi z meritvami na postaji (postopek sklapljanja sicer ni obvezen, a ima pomemben prispevek pri doseganju boljnih napovedi), kljub temu pa ostali pretekli podatki fizikalnega modela ne morejo izboljšati.

Čeprav smo uporabili metodo na razmeroma majhni količini aplikativnih podatkov, ki so nam bili na voljo, so dobljeni rezultati zelo spodbudni. Ugotavljamo, da lahko predlagana metoda predstavlja ustrezno alternativo fizikalnim modelom za napovedovanje temperature in stanja cestišča. Z uporabo metode na umetnih podatkih (razdelek 4.1.2) pa smo pokazali, da je le-ta lahko uporabna tudi na podobnih, bolj splošnih podatkih.

## 4.2 Dekompozicija napake hierarhičnih regresijskih modelov

V tem razdelku bomo opisali prilagoditev postopka dekompozicije napake na pristranskost in varianco za njeno uporabo na metodi, predstavljeni v prejšnjem razdelku. Prilagojen postopek dekompozicije bomo preizkusili na aplikativnih podatkih in predstavili njene rezultate.

$MS$  napaka je definirana kot pričakovana vrednost kvadrata razlike med pravimi

vrednostmi  $y$  in izračunanimi (napovedanimi) vrednostmi  $\hat{y}$ :

$$MS(\hat{y}) = E[(\hat{y} - y)^2]. \quad (4.4)$$

Kot smo pokazali v razdelku 2.1.2, lahko dekompozicijo napake  $MS(\hat{y})$  zapišemo kot:

$$\begin{aligned} MS(\hat{y}) &= \\ &= \underbrace{E[(\hat{y} - E[\hat{y}])^2]}_{\text{varianca}} + \underbrace{(E[\hat{y} - y])^2}_{\text{pristranskost}^2} + \underbrace{E[(y - E[y])^2]}_{\text{šum}} \\ &= \text{varianca} + \text{pristranskost}^2 + \text{šum}. \end{aligned} \quad (4.5)$$

Varianca opisuje, kolikšna bo sprememba  $\hat{y}$  v primeru drugačne učne množice. Pristranskost opisuje povprečno napako  $\hat{y}$ . Na šum, ki predstavlja varianco  $y$ , običajno nimamo vpliva, medtem ko sta pristranskost in varianca odvisni od lastnosti algoritma, ki ga proučujemo.

#### 4.2.1 Opis postopka dekompozicije

Pri proučevanju teoretičnega ozadja delovanja metode, opisane v razdelku 4.1, je naša hipoteza, da naj bi se z združevanjem manjšala varianca in večala pristranskost. Vsak korak združevanja namreč zmanjšuje končno število modelov in k obstoječim dodaja nove podatke. Padanje variance je tako povezano z večjo množico učnih podatkov, zaradi česar se linearni modeli ne morejo ustrezno prilagoditi podatkom in so, gledano kot celota, manj kompleksni. Za empirični preizkus hipoteze predlagamo uporabo spodnjega postopka za dekompozicijo napake na pristranskost in varianco za metodo, predstavljeno v razdelku 4.1.1.

*Vhod:* modeli posameznega koraka združevanja metode za gradnjo hierarhičnih regresijskih modelov, skupaj z učnimi podatki.

*Izhod:* dekompozicija napake posameznega koraka združevanja metode za gradnjo hierarhičnih regresijskih modelov.

for vsak korak v postopku združevanja do

for vsak model trenutnega koraka do

repeat

naključna razdelitev učnih podatkov na dve enako veliki množici;  
učenje trenutnega modela na eni množici in izračun *MS* na drugi množici razdeljenih podatkov;

until ni dosežena zelena natančnost pri izračunu *MS*;

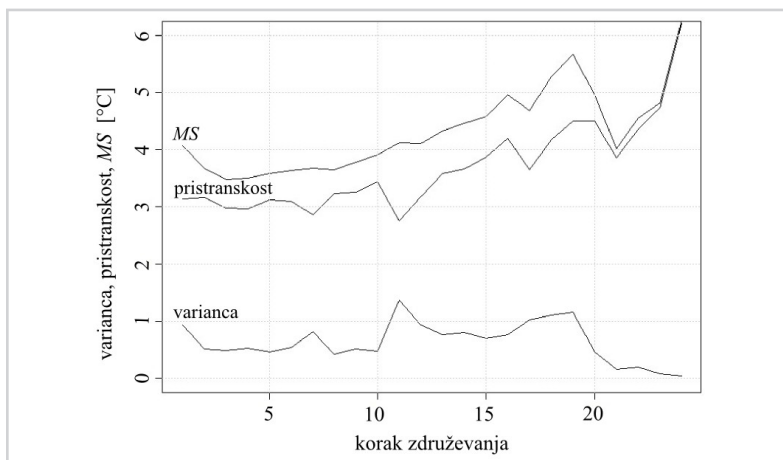
izračun povprečja vseh *MS* trenutnega modela in njegove variance;

end

izračun povprečja *MS* in varianc vseh modelov trenutnega koraka;

izračun pristranskosti po enačbi 2.8, pri čemer se privzame, da je  $\sum = 0$ ;

end



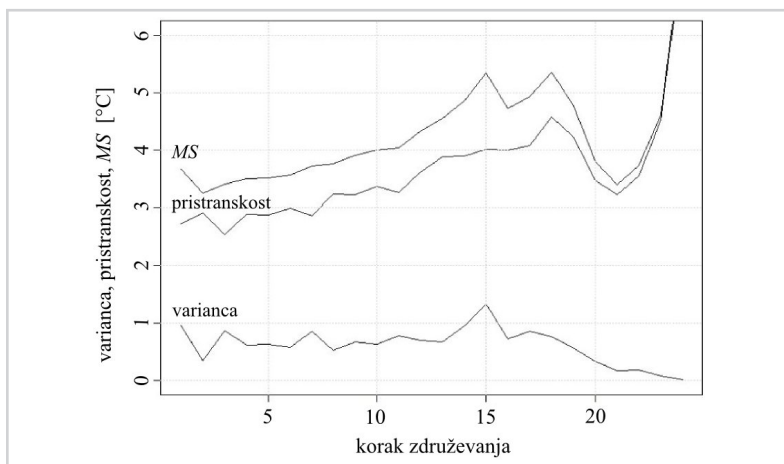
Slika 4.7

Dekompozicija napake metode za gradnjo hierarhičnih regresijskih modelov na problemu napovedovanja meteorološkega stanja cestišča na podatkih CVP Jepca. Na sliki je prikazana varianca (spodnja krivulja), kvadrirana pristranskost (srednja krivulja) in *MS* (zgornja krivulja).

#### 4.2.2 Rezultati postopka dekompozicije

Analiza razmerja med pristranskostjo in varianco pri združevanju modelov je pokazala, da z naraščanjem napake narašča pristranskost, varianca pa pada. Opisano je razvidno iz slike 4.7 in 4.8, kjer je prikazana dekompozicija napake metode za gradnjo hierarhičnih regresijskih modelov na problemu napovedovanja meteorološkega stanja cestišča na podatkih CVP Jeprca in CVP Črmošnjice. Na obeh slikah je pristranskost kvadrirana, zaradi česar je lažje razvidno, da je  $MS$  vsota kvadratov pristranskosti in variance.

S tem lahko hipotezo potrdimo. Iz slike 4.8 je razvidno, da je postopek empiričnega merjenja dekompozicije sicer nekoliko naključen (delež pristranskosti in variance naključno niha), medtem ko kasnejši koraki združevanja jasno kažejo padanje variance in naraščanje pristranskosti (zlasti v zadnjih nekaj korakih je naraščanje pristranskosti najbolj izrazito). Razlog znatnega padca variance tik pred njenim strmim naraščanjem je ta, da postopek v kasnejših korakih začne združevati modele, ki imajo majhno napako in veliko množico, skupaj z modeli, ki imajo veliko napako in majhno množico (takšni modeli se navadno do teh korakov še niso veliko združevali).



Slika 4.8

Dekompozicija napake metode za gradnjo hierarhičnih regresijskih modelov na problemu napovedovanja meteorološkega stanja cestišča na podatkih CVP Črmošnjice. Na sliki je prikazana varianca (spodnja krivulja), kvadrirana pristranskost (srednja krivulja) in  $MS$  (zgornja krivulja).

### 4.3 *Analiza povezav med modeli in lastnostmi objektov*

Pri iskanju povezav med podobnostjo modelov in podobnostjo lastnosti objektov, na katere se ti modeli nanašajo, želimo uporabiti primerne postopke za opazovanje in analizo takšnih povezav. Zanima nas, katere lastnosti objektov vplivajo na podobnost med modeli in na kakšen način.

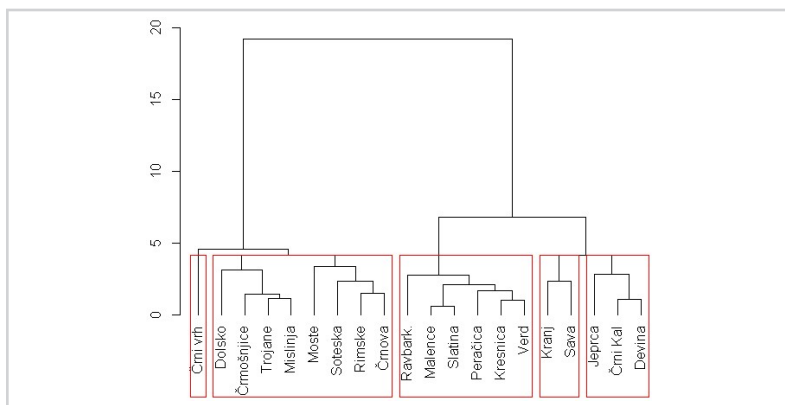
V splošnem imamo lahko podanih več modelov, ki se nanašajo na posamezne objekte. Ti modeli so opisani z različnimi spremenljivkami (na primer z regresijskimi koeficienti). Po drugi strani pa imajo objekti, ki se na te modele nanašajo, določene lastnosti. Za razliko od klasičnih fizikalnih modelov, kjer lastnosti objektov neposredno vstopajo v model, pa takšne povezave med spremenljivkami modelov in lastnostmi objektov niso eksplicitne, marveč se pojavlja možnost po njihovi analizi. Za tovrstno analizo je potrebno natančno določiti, katere spremenljivke modelov želimo uporabiti in katere so tiste lastnosti objektov, ki bi utegnile predstavljati takšne povezave. Poleg tega želimo takšne povezave kvantitativno ovrednotiti in določiti mero razdalje med spremenljivkami modelov in med lastnostmi objektov.

V nadaljevanju besedila se bomo osredotočili na primer modeliranja temperature cestišča, kjer lahko posamezne napovedne modele opišemo z regresijskimi koeficienti, ki jih izračunamo na podlagi preteklih podatkov. Objekti, ki se na te modele nanašajo, pa v našem primeru predstavljajo različne lokacije, kjer se nahajajo cestnovremenske postaje (CVP), in jih lahko opišemo z različnimi lastnostmi (atributi). Te opise lahko predstavimo v matrični obliki in nato z opazovanjem podobnosti med njimi z obstoječimi metodami poiščemo različne povezave (razdelek 4.3.1). Ko se bomo prepričali, da takšne povezave obstajajo, sledi opis metode za analizo povezav (razdelek 4.3.2), ki predstavlja nadgradnjo že predstavljene metode združevanja (razdelek 4.1.1). To metodo bomo uporabili na umetnih podatkih (razdelek 4.3.3), poizkušali poiskati matriko podobnosti  $P$  in jo primerjati z matriko, ki je bila uporabljena pri generiranju podatkov. Metodo bomo uporabili tudi na aplikativnih podatkih (razdelek 4.3.4), poiskali koeficiente modelov iz lastnosti teh modelov in primerjali rezultate napovedi s poiskanimi koeficienti z napovedmi, kjer smo koeficiente izračunali na podlagi podatkov modelov.

### 4.3.1 Opazovanje podobnosti med modeli in lastnostmi

V tem razdelku bomo opazovali podobnosti med modeli in lastnostmi objektov (atributi) pri modeliranju temperature cestišča. Pri tem bomo za nakazovanje takšnih povezav uporabili obstoječe metode na podatkih, ki so predstavljeni v razdelku 3.1.2.

**Uporaba gručenja.** Pri ločenem opazovanju podobnosti med posameznimi modeli in podobnosti med posameznimi lokacijami smo uporabili postopek gručenja. Izračunali smo razdalje med regresijskimi koeficienti vseh modelov  $\Delta M$ , poleg tega pa tudi razdalje med lastnostmi vseh lokacij  $\Delta L$  (tabele 4.3, 4.4 in 4.5). Na podlagi teh razdalj smo izdelali gručenje modelov glede na podobnost njihovih regresijskih koeficientov in gručenje lokacij glede na podobnost njihovih lastnosti, pri tem pa za gručenje uporabili metodo Warda [104]. Gručenje lokacij je prikazano na sliki 4.9. Pri gručenju modelov smo uporabili tako regresijske koeficiente dnevnih modelov (slika 4.10), kot tudi regresijske koeficiente nočnih modelov (slika 4.11).

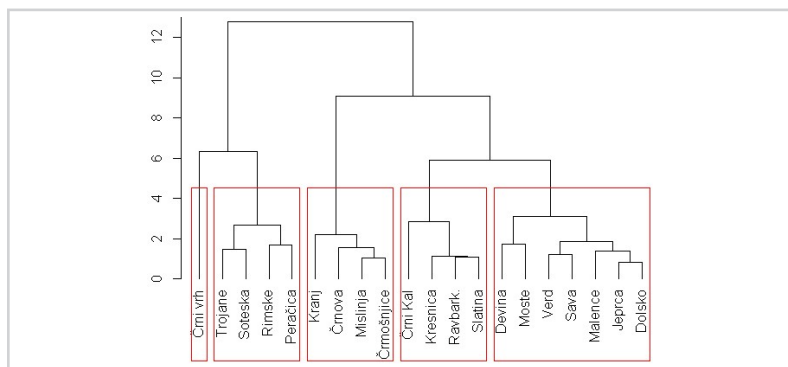


Slika 4.9

Gručenje lokacij glede na njihove lastnosti. Uporabljena je evklidska razdalja, gručenje pa izdelano s pomočjo Wardove metode.  $\Psi_s$  je predstavljena s štirimi lastnostmi (S, J, V, Z). Z rdečimi kvadrati je razmejenih pet gru.

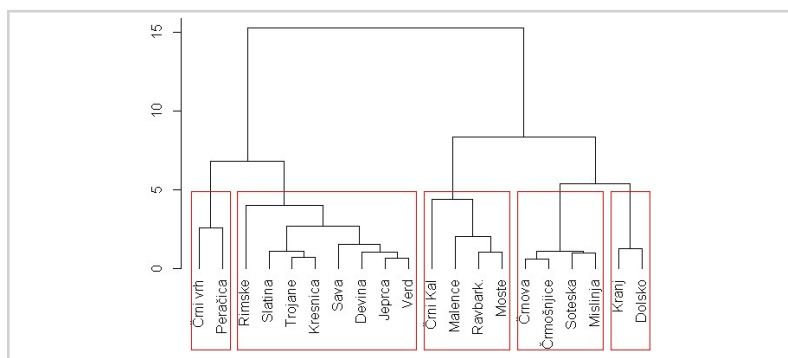
Pri izračunu razdalj smo se odločili za evklidsko razdaljo. V tem primeru lahko na naš rezultat vpliva dejstvo, da se lastnost vidnost neba pojavi kar petkrat in s tem po pomembnosti prevladuje nad ostalimi lastnostmi. Zaradi tega razloga smo poleg evklidske razdalje računali tudi Mahalanobisovo razdaljo [67], vendar se je pri vseh gručenjih izkazalo, da različni izračuni razdalj niso bistveno vplivali na razlike pri gručenju.





Slika 4.10

Gručenje dnevnih modelov glede na njihove regresijske koeficiente. Uporabljena je evklidska razdalja, gručenje pa izdelano s pomočjo Wardove metode. Z rdečimi kvadrati je razmejenih pet gruč.



Slika 4.11

Gručenje nočnih modelov glede na njihove regresijske koeficiente. Uporabljena je evklidska razdalja, gručenje pa izdelano s pomočjo Wardove metode. Z rdečimi kvadrati je razmejenih pet gruč.

Pri gručenju lokacij dobimo pri razmejitvi petih gruč: v prvi CVP, ki izstopa po višini in zelo nizki vrednosti vidnosti neba; v drugi CVP, ki se nahajajo pretežno na kopnem in imajo nizko vrednost vidnosti neba in nizko vrednost vodnih virov v bližini; v tretji CVP, ki se nahajajo na viaduktih in imajo visoko vrednost vidnosti neba in nizko vrednost vodnih virov v bližini; v četrti CVP, ki imajo v bližini vodne vire in visoko vrednost vidnosti neba; v peti CVP, ki imajo zelo visoko vrednost vidnosti neba in nizko vrednost vodnih virov. ❖ Pri gručenju dnevnih modelov dobimo pri razmejitvi petih gruč: v prvi model, ki ima zelo nizko povprečno temperaturo cestišča in visoko pomembnost temperature cestišča; v tretji modele s pomembnimi vsemi koeficienti; v četrti modele z zelo visokimi povprečnimi temperaturami cestišča, z manj pomembnim koeficientom temperature zraka in bolj pomembnim koeficientom krat-

kovalovnega sevanja; v peti modele z visokimi povprečnimi temperaturami cestišča, pomembnimi koeficienti temperature zraka, kratkovalovnega sevanja in temperature cestišča; preostali modeli so v drugi gruči. ¶ Pri gručenju nočnih modelov dobimo pri razmejitvi petih gruči: v prvi modele z nizkimi povprečnimi temperaturami cestišča in z visoko pomembnostjo temperature cestišča; v drugi modele z visoko pomembnostjo temperature zraka in temperature cestišča; v četrti modeli z visoko pomembnostjo temperature cestišča v globini in nizko pomembnostjo napovedane temperature zraka; v peti modela, ki imata visoko povprečno temperaturo cestišča, visoko pomembnost napovedane temperature zraka in nizko pomembnost temperature cestišča; preostali modeli so v tretji gruči.

*Tabela 4.3*

Matriki evklidskih razdalj za dnevne ( $\Delta M_d$ ) modele. Imena nekaterih CVP so skrajšana.

CVP	Kranj	Jep.	Tro.	Rim.	Črn.	Črni	Sot.	Dol.	Mis.	Črm.	Kre.	Kal	Dev.	Rav.	Mal.	Sl.	Verd	Sava	Per.
Jepca	1,9																		
Trojane	3,2	1,9																	
Rimske t.	4,0	2,3	1,5																
Črnova	1,7	2,2	2,7	3,5															
Črni vrh	6,7	6,0	4,9	4,5	5,6														
Soteska	3,3	2,4	1,4	2,5	2,2	5,0													
Dolsko	1,7	0,8	2,5	3,1	2,2	6,7	2,8												
Mislinja	2,4	2,7	3,4	4,3	1,7	7,0	2,7	2,2											
Črmošnjice	1,5	2,1	2,8	3,8	1,1	6,6	2,4	1,7	1,0										
Kresnica	3,7	2,1	3,1	3,1	3,8	7,3	3,5	2,0	3,5	3,4									
Črni Kal	3,6	2,5	3,6	3,5	4,2	7,8	4,2	2,3	3,9	3,6	2,0								
Devina	2,8	2,7	4,3	4,6	4,0	8,3	5,0	2,3	4,1	3,6	3,0	2,8							
Ravbark.	3,9	2,2	2,7	2,9	3,9	7,0	3,3	2,2	3,7	3,5	1,1	2,6	3,3						
Malence	2,7	1,2	2,7	2,9	3,1	6,7	3,3	1,2	3,3	2,8	1,4	2,2	2,1	1,5					
Slatina	3,9	2,3	3,2	3,1	3,8	7,1	3,7	2,2	3,7	3,6	1,1	2,2	3,0	1,1	1,4				
Verd	1,9	1,0	2,3	2,6	2,6	6,1	3,1	1,5	3,4	2,6	2,7	2,8	2,3	2,7	1,5	2,7			
Sava	2,0	1,6	3,3	3,5	2,9	6,8	3,9	1,6	3,5	2,8	2,7	2,7	1,6	3,0	1,5	2,7	1,2		
Peračica	4,6	3,0	2,2	1,7	3,7	4,4	2,3	3,5	4,2	4,0	3,2	4,3	5,3	2,9	3,3	3,1	3,5	4,2	
Moste	2,2	2,0	3,3	3,8	3,5	7,6	4,1	1,9	3,8	3,0	2,9	2,6	1,7	3,2	2,1	3,3	1,7	1,8	4,8

*Tabela 4.4*

Matriki evklidskih razdalj za nočne ( $\Delta M_n$ ) modele. Imena nekaterih CVP so skrajšana.

CVP	Kranj	Jep.	Tro.	Rim.	Črn.	Črni	Sot.	Dol.	Mis.	Črm.	Kre.	Kal	Dev.	Rav.	Mal.	Sla.	Verd	Sava	Per.
Jeprca	2,6																		
Trojane	3,2	1,2																	
Rimske t.	3,7	2,8	2,2																
Črnova	2,3	4,1	3,9	4,0															
Črni vrh	5,0	4,1	3,0	4,0	4,3														
Soreska	2,9	4,3	3,9	4,1	0,9	3,8													
Dolsko	1,3	3,4	3,9	4,1	2,7	5,8	3,5												
Mislinja	2,3	3,4	3,1	3,3	1,2	3,4	1,0	3,1											
Črmošnjice	2,0	3,5	3,3	3,6	0,6	3,9	1,0	2,6	0,7										
Kresnica	3,3	1,3	0,7	2,1	4,2	3,4	4,2	4,1	3,3	3,6									
Črni Kal	3,4	4,4	4,8	5,3	4,4	6,4	4,6	4,2	4,0	4,0	4,5								
Devina	3,1	0,9	1,6	3,6	4,5	4,1	4,7	3,9	3,8	4,0	1,8	4,7							
Ravbark.	3,0	2,3	2,4	4,4	3,8	3,5	3,8	4,0	3,1	3,3	2,8	4,3	1,8						
Malence	2,5	3,1	3,7	5,0	3,6	4,9	3,8	3,6	3,1	3,2	3,7	2,6	3,1	2,0					
Slatina	3,1	1,2	1,0	2,7	4,1	3,6	4,1	3,9	3,2	3,5	1,0	4,1	1,4	2,1	3,1				
Verd	2,9	0,7	1,4	3,0	4,3	4,2	4,5	3,7	3,6	3,7	1,3	4,0	1,0	2,3	3,0	0,8			
Sava	3,6	1,2	2,0	3,3	5,2	5,0	5,5	4,1	4,6	4,7	1,9	4,8	1,5	3,2	4,0	1,8	1,2		
Peračica	4,9	3,0	2,3	3,9	5,2	2,6	5,1	5,6	4,5	4,8	2,8	6,9	2,9	3,3	5,2	3,1	3,3	3,6	
Moste	2,4	2,1	2,3	3,9	3,2	3,6	3,3	3,6	2,5	2,7	2,4	3,4	2,1	1,0	1,5	1,9	2,0	3,2	3,8

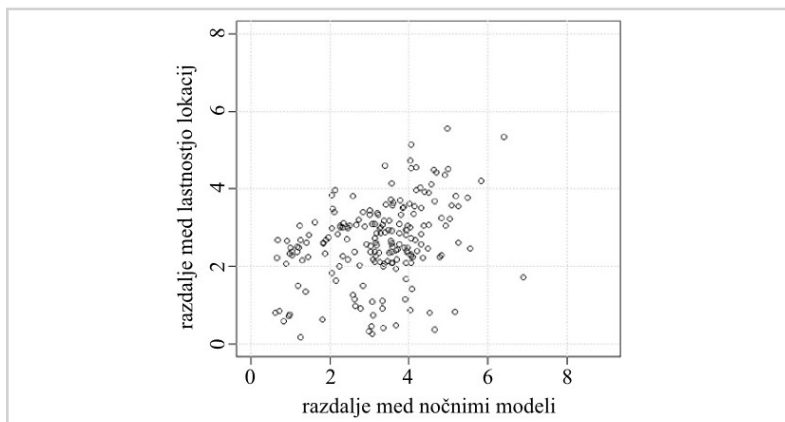
*Tabela 4.5*

Matrika evklidskih razdalj za lokacije ( $\Delta L$ ). Imena nekaterih CVP so skrajšana.

CVP	Kranj	Jep.	Tro.	Rim.	Črn.	Črni	Sot.	Dol.	Mis.	Črm.	Kre.	Kal	Dev.	Rav.	Mal.	Sl.	Verd	Sava	Per.
Jeprca	3,1																		
Trojane	3,8	4,7																	
Rimske t.	3,6	5,4	2,9																
Črnova	3,9	5,4	1,7	1,5															
Črni vrh	6,0	7,1	3,2	4,6	3,7														
Soteska	4,5	6,4	2,8	2,3	2,0	2,9													
Dolsko	3,8	3,9	2,5	3,3	2,7	5,1	4,3												
Mislinja	4,1	5,3	1,2	2,8	1,6	2,5	2,0	2,9											
Črmošnjice	3,7	4,4	1,2	2,3	1,6	3,4	2,7	2,2	1,5										
Kresnica	3,4	2,9	3,5	4,2	3,9	5,8	5,1	2,9	4,1	3,2									
Črni Kal	3,6	2,4	4,6	5,1	5,0	7,2	6,3	3,5	5,2	4,4	1,5								
Devina	2,9	2,4	4,5	5,0	5,0	7,1	6,1	3,9	5,2	4,4	1,7	1,1							
Ravbark.	4,0	4,2	2,9	4,1	3,6	4,0	4,0	3,4	3,0	2,6	2,4	3,7	3,7						
Malence	3,4	3,4	3,0	3,9	3,5	5,5	4,7	2,4	3,5	2,8	1,0	2,0	2,2	2,2					
Slatina	3,7	3,7	2,8	4,0	3,4	5,2	4,7	2,4	3,4	2,7	1,4	2,4	2,6	2,2	0,6				
Verd	3,4	2,9	3,9	4,1	4,1	6,0	5,2	3,2	4,3	3,3	1,1	1,9	2,0	2,4	1,5	1,9			
Sava	2,4	3,5	5,1	4,7	5,5	7,2	5,7	4,5	5,3	4,8	3,1	2,7	2,2	4,2	3,2	3,8	2,9		
Peračica	3,1	3,3	3,2	4,1	3,7	5,2	4,5	3,1	3,5	3,1	1,3	2,3	2,3	1,9	1,4	1,8	1,8	2,9	
Moste	3,8	5,7	2,7	3,2	2,7	3,8	2,7	3,9	2,4	3,0	3,8	4,9	4,5	2,9	3,2	3,1	4,0	4,3	3,1

*Analiziranje razdalj.* Pri gručenju dnevnih in nočnih modelov lahko najdemo več podobnosti, a tudi primere, ko je razvrstitev povsem drugačna (npr. pri CVP Verd). Da so modeli in lokacije med seboj povezani, nam nakazujejo že slike gručenja, zato smo te povezave podrobneje analizirali. Uporabili smo že izdelane matrike:  $\Delta M_d$ , ki opisuje razdalje med regresijskimi koeficienti dnevnih modelov, matriko  $\Delta M_n$ , ki opisuje razdalje med regresijskimi koeficienti nočnih modelov in matriko  $\Delta L$ , ki opisuje razdalje glede na lastnosti lokacij (tabele 4.3, 4.4 in 4.5).

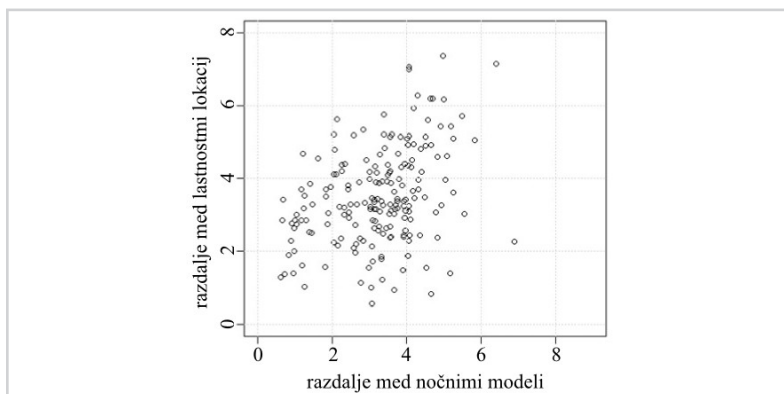
Ob primerjavi slik 4.12 in 4.13 lahko vidimo, da se z večanjem razdalje med modeli večja tudi razdalja glede na njihove lastnosti. Takšen pojav nakazuje, da so modeli povezani z lastnostmi teh modelov, čeprav je povezava zaradi razpršenosti razdalj šibka. V primeru, ko smo enako primerjavo naredili z več lastnostmi za vidnost neba (za vse štiri strani neba), slika 4.13, pa je ta povezava zaradi razpršenosti razdalj še šibkejša. S primerjavo vplivov nočnih (slika 4.13) in dnevnih (slika 4.14) modelov lahko ugotovimo, da je vpliv parametrov ponoči podoben kot podnevi.



Slika 4.12

Graf razdalj med nočnimi modeli (abscisna os) in lastnostjo pripadajočih lokacij  $\Psi_4$  (ordinatna os). Vsaka točka predstavlja par lokacij (modelov).

*Izračun korelacij.* Med matrikama dnevnih in nočnih modelov in matriko lokacij smo izračunali korelacije ( $cor(M_d, L)$  in  $cor(M_n, L)$ ). V tabelah 4.6 in 4.7 so podani korelacijski koeficienti med dnevnimi oziroma nočnimi modeli in lastnostmi lokacij, na katere se ti modeli nanašajo.



Slika 4.13

Graf razdalj med nočnimi modeli (abscisna os) in lastnostmi pripadajočih lokacij  $\Psi_s(S)$ ,  $\Psi_s(I)$ ,  $\Psi_s(V)$  in  $\Psi_s(Z)$  (ordinatna os). Vsaka točka predstavlja par lokacij (modelov).

Tabela 4.6

Korelacije med lastnostmi dnevnih modelov in lastnostmi lokacij ( $cor(M_d, L)$ ). IN pomeni prosti člen regresije, ostale oznake so obrazložene v tabeli 3.1, oznake lastnosti pa v tabeli 3.4.

Lastnosti	$h$	Kopno	Vodni vir	$\Psi_s$
IN	-0,68	-0,49	-0,22	0,59
MA	-0,37	-0,12	0,26	0,39
MR	0,48	0,22	-0,07	-0,50
MD	0,11	0,61	0,20	-0,39
FA <sub>6</sub>	-0,51	-0,72	-0,44	0,60
FS	-0,46	-0,42	-0,15	0,40

Tabela 4.7

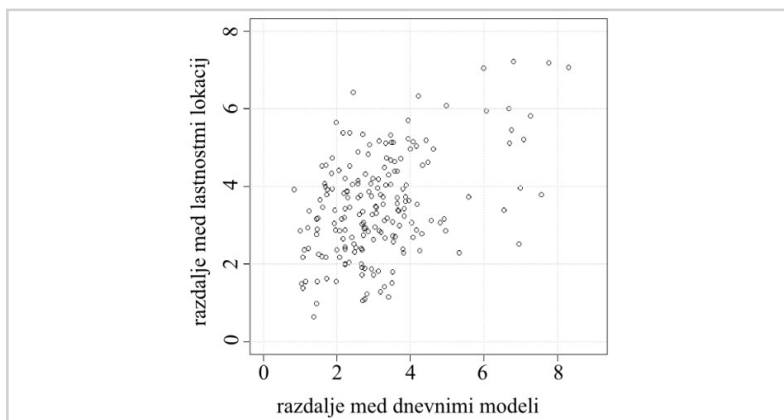
Korelacije med lastnostmi nočnih modelov in lastnostmi lokacij, na katere se ti modeli nanašajo ( $cor(M_n, L)$ ). IN pomeni prosti člen regresije, ostale oznake so obrazložene v tabeli 3.1, oznake lastnosti pa v tabeli 3.4.

Lastnost	$h$	Kopno	Vodni vir	$\Psi_s$
IN	-0,76	-0,04	0,00	0,41
MA	0,22	0,13	0,00	-0,40
MR	0,20	-0,11	-0,04	-0,01
MD	-0,15	0,76	0,30	-0,43
FA <sub>6</sub>	-0,46	-0,64	-0,29	0,82

**Opazovanje korelacij posameznih lastnosti.** Pregled obeh tabel korelacij nam pokaže, da so le-te pri določenih lastnostih dovolj visoke, da lahko sklepamo o vplivu določene

Slika 4.14

Graf razdalj med dnevnimi modeli (abscisna os) in lastnostmi pripadajočih lokacij  $\Psi_s(S)$ ,  $\Psi_s(I)$ ,  $\Psi_s(V)$  in  $\Psi_s(Z)$  (ordinatna os). Vsaka točka predstavlja par lokacij (modelov).

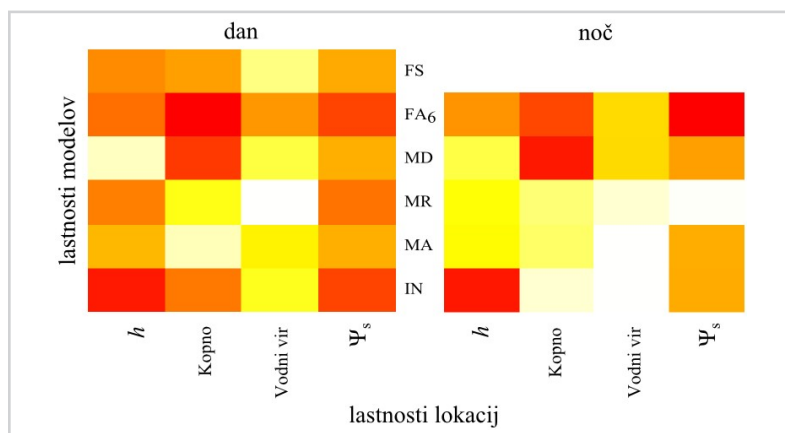


lastnosti na razdalje med posameznimi modeli. Opazovanje korelacijskih koeficientov nam olajša izdelava barvne karte korelacij. V kolikor opazujemo absolutne vrednosti korelacij, nam temnejša polja predstavljajo večjo (pozitivno ali negativno) korelacijo, svetlejša pa manjšo korelacijo.

Na levi strani slike 4.15 lahko vidimo, da ima podnevi lastnost  $h$  (višina CVP) znanen vpliv na prosti člen dnevnih modelov. Kot je razvidno iz tabele 4.6, je ta korelacija negativna. Opisan vpliv lahko pojasnimo z dejstvom, da imajo višje CVP manjšo povprečno temperaturo cestišča (manjše vrednosti prostega člana regresije), kar se sklada z ugotovitvami iz tega področja<sup>9</sup>. Podobno imajo tudi CVP, ki se nahajajo na kopnem, manjšo povprečno temperaturo cestišča (to lahko razložimo s pojmom termalnega spomina ceste, opisanim v razdelku 2.2.3), medtem ko so za CVP, ki se nahajajo na odprtih lokacijah (večja vrednost  $\Psi_s$ ), značilne večje povprečne temperature cestišča (vpliv sončnega sevanja podnevi). Za CVP, ki se nahajajo na odprtih lokacijah, ali CVP, ki se nahajajo na viaduktih, je značilen bolj izrazit vpliv napovedi temperature zraka ( $FA_6$ ). ¶ Pri opazovanju korelacij med nočnimi modeli in lastnostmi lokacij na desni strani slike 4.15 lahko opazimo več podobnosti z opazovanjem povezav dnevnih modelov; nekatere povezave so še bolj izrazite (npr. vpliv napovedi temperature zraka in vidnost neba).

<sup>9</sup>Prim. razdelek 2.2.3 na strani 18.





Slika 4.15

Opazovanje povezav med lastnostmi dnevnih in nočnih modelov in lastnostmi lokacij s pomočjo barvne karte korelacij. Temnejša polja predstavljajo večjo (pozitivno ali negativno) korelacijo, svetlejša pa manjšo korelacijo. IN pomeni prosti člen regresije, ostale oznake so obrazložene v tabeli 3.1, oznake lastnosti pa v tabeli 3.4.

*Združevanje modelov glede na lastnosti lokacij.* Želeli smo ugotoviti, ali je ob pridobljenem znanju o lastnostih lokacij in njihovem vplivu na modele te modele smotno združevati. Modele smo združevali po lastnostih lokacije, za katere smo ugotovili, da med njimi in lastnostmi modelov obstajajo povezave. Izračunali smo RMS napako enega splošnega združenega modela vseh lokacij. Nato smo modele lokacij združili po lastnosti *Kopno* in izračunali napako dobljenih dveh gruč. Na koncu smo združili modele lokacij po lastnosti *Kopno* in znotraj tega po lastnosti  $\Psi_s$  (tabela 4.8), kot tudi po lastnosti *Vodni vir* (tabela 4.9). V obeh primerih smo dobili štiri gruče modelov. Rezultati kažejo, da je združevanje modelov glede na lastnosti lokacij smotno, saj je v primerjavi z enim splošnim modelom vseh lokacij RMS napaka združenih modelov manjša ali podobna. To nakazuje, da bi metoda za analizo povezav, predstavljena v nadaljevanju, lahko dajala zadovoljive rezultate tudi na problemu napovedovanja temperature cestišča.

#### 4.3.2 Metoda za analizo povezav

Metoda za analizo povezav temelji na predpostavki, da med modeli in lastnostmi objektov, na katere se ti modeli nanašajo, obstajajo povezave. Za te povezave predpostavljamo, da jih je mogoče opisati kot linearno kombinacijo koeficientov modelov in koeficientov lastnosti objektov (atributov).

Metoda predstavlja nadgradnjo metode za gradnjo hierarhičnih regresijskih mode-

Tabela 4.8

RMS napaka modelov glede na združevanje lokacij po lastnosti *Kopno* in znotraj tega po lastnosti  $\Psi_s$ . Podatki in oznake lastnosti lokacij so navedeni v tabeli 3.4.

# CVP	<i>Kopno</i>	$\Psi_s$	RMS napaka
<i>Dnevni modeli</i>			
1, 2	1	$\geq 0,70$	3,08
3, 4, 5, 6, 7, 8, 9, 10	1	$< 0,70$	3,77
11, 12, 13, 17, 18, 19	0	$\geq 0,70$	4,29
14, 15, 16, 20	0	$< 0,70$	3,41
I – 10	1	-	3,66
II – 20	0	-	3,99
I – 20	-	-	4,06
<i>Nočni modeli</i>			
1, 2	1	$\geq 0,70$	1,55
3, 4, 5, 6, 7, 8, 9, 10	1	$< 0,70$	1,82
11, 12, 13, 17, 18, 19	0	$\geq 0,70$	2,56
14, 15, 16, 20	0	$< 0,70$	1,66
I – 10	1	-	1,77
II – 20	0	-	2,29
I – 20	-	-	2,12

lov, opisane v razdelku 4.1, ki nam z združevanjem pripravi tiste modele, ki so si med seboj podobni in za katere predpostavljamo, da imajo podobno matriko  $P^{10}$ . Ideja nadgradnje je, da z regresijo na podlagi podatkov odvisnih in neodvisnih spremenljivk vsakega modela poiščemo njihove koeficiente. Dalje, prav tako z regresijo, poiščemo še matriko  $P$  na podlagi prej izračunanih vrednosti koeficientov in vrednosti atributov. Iz matrike  $P$  lahko razberemo, kako pomembne so katere povezave med atributi in koeficienti, poleg tega pa nam omogoča, da iz atributov izračunamo koeficiente modelov, ki jih ne poznamo, ali obratno.

Preizkus metode na umetnih podatkih je predstavljen v razdelku 4.3.3, izračun koeficientov iz atributov na aplikativnih podatkih pa v razdelku 4.3.4.

<sup>10</sup>Prim. razdelek 3.2 na strani 45.

Tabela 4.9

RMS napaka modelov glede na združevanje lokacij po lastnosti *Kopno* in znotraj tega po lastnosti *Vodni vir*. Podatki in oznake lastnosti lokacij so navedeni v tabeli 3.4.

# CVP	<i>Kopno</i>	<i>Vodni vir</i>	RMS napaka
<i>Dnevni modeli</i>			
1, 4, 5, 7, 9	1	$\geq 0,50$	3,23
2, 3, 6, 8, 10	1	$< 0,50$	3,91
11, 12, 13, 14, 15, 16, 17, 19	0	$< 0,50$	4,06
18, 20	0	$\geq 0,50$	3,42
1 – 10	1	-	3,67
11 – 20	0	-	4,00
1 – 20	-	-	4,06
<i>Nočni modeli</i>			
1, 4, 5, 7, 9	1	$\geq 0,50$	1,59
2, 3, 6, 8, 10	1	$< 0,50$	1,87
11, 12, 13, 14, 15, 16, 17, 19	0	$< 0,50$	2,40
18, 20	0	$\geq 0,50$	1,66
1 – 10	1	-	1,78
11 – 20	0	-	2,28
1 – 20	-	-	2,11

#### 4.3.3 Preizkus na umetnih podatkih

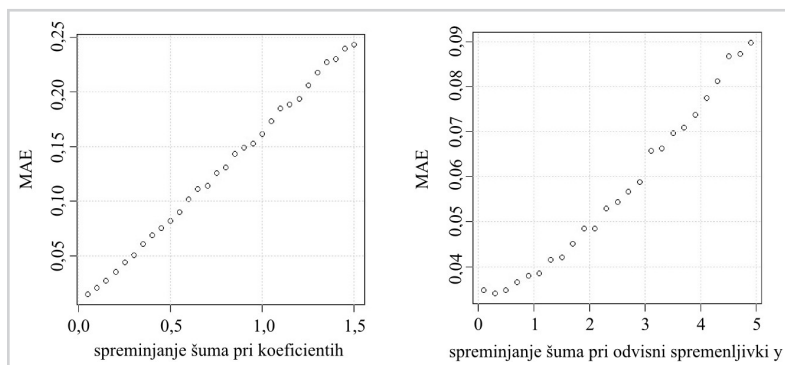
Na vnaprej pripravljenih umetnih podatkih, opisanih v razdelku 3.2, smo preizkusili metodo za iskanje povezav med modeli in lastnostmi objektov (atributi), na katere se ti modeli nanašajo. Pripravili smo 40 modelov. Na vsakega izmed njih se nanašajo 4 zvezni in 2 diskretna atributa ter 8 spremenljivk (vključno s prostim členom). Za vsak model je na voljo 100 primerov. Varianca normalne porazdelitve za generator šuma ( $\sigma^2$ ) pri koeficientih je enaka 0,2, varianca normalne porazdelitve za generator šuma ( $\sigma^2$ ) pri odvisni spremenljivki  $y$  pa je enaka 1,0.

Na sliki 4.16 in 4.17 je prikazanih več grafov z rezultati. Ker smo si matriko  $P$ , ki smo jo uporabili pri pripravi podatkov, shranili, smo jo lahko primerjali z matriko  $P'$ , ki smo jo poiskali s pomočjo naše metode, tako, da smo med njima računali povprečno

absolutno napako (MAE). Vsaka točka posameznega grafa predstavlja povprečno napako 50-ih ponovitev metode na podatkih. Iz grafov je razvidno povečevanje napake pri povečevanju šuma pri koeficientih ali šuma pri odvisni spremenljivki  $y$  (grafa na sliki 4.16) in zmanjševanje napake pri povečevanju števila modelov ali števila primerov (podatkov) na model (grafa na sliki 4.17).

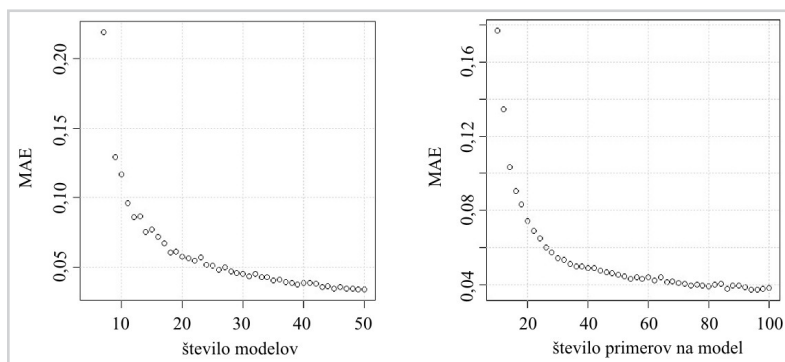
Slika 4.16

Preizkus metode za iskanje povezav med modeli in atributi na umetnih podatkih. Levo: spreminjanje šuma pri koeficientih (varianca normalne porazdelitve teče od 0,05 do 1,50). Desno: spreminjanje šuma pri odvisni spremenljivki  $y$  (varianca normalne porazdelitve teče od 0,1 do 5).



Slika 4.17

Preizkus metode za iskanje povezav med modeli in atributi na umetnih podatkih. Levo: spreminjanje števila modelov (od 5 do 50). Desno: spreminjanje števila primerov na model (od 5 do 100).



#### 4.3.4 Preizkus na problemu napovedovanja temperature cestišča

Metoda, predstavljena v razdelku 4.3.2, nam omogoča izračun koeficientov tudi v primeru, ko za določen model nimamo na voljo izmerjenih podatkov. Koeficiente lahko izračunamo s pomočjo vrednosti lastnosti objektov (atributov) in matrike  $P$ , saj

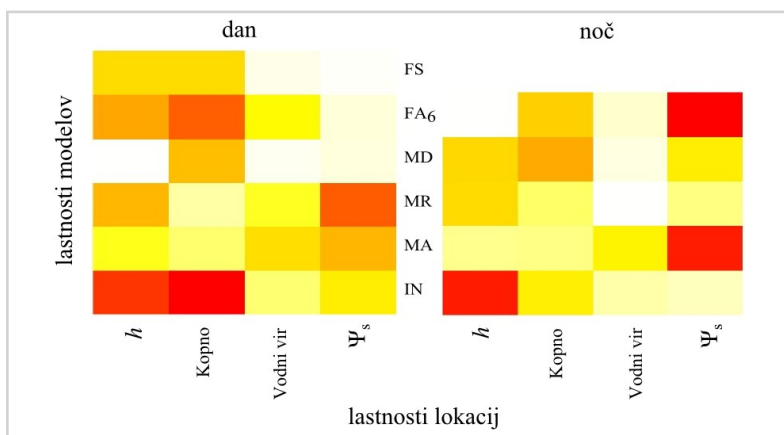
predpostavljamo, da  $P$  odraža podobnosti med modeli posameznih CVP in njihovimi lokacijami.

Pri problemu napovedovanja temperature cestišča nam attribute predstavljajo lastnosti lokacij CVP, modele nam predstavljajo CVP oziroma regresijski koeficienti, spremenljivke pa so meritve, iz katerih poizkušamo napovedati temperaturo cestišča. Za vsako CVP imamo na voljo meritve in attribute, ki so vezani na lokacijo teh postaj. Iz meritev za vsako CVP izračunamo regresijske koeficiente, s pomočjo katerih lahko napovemo temperaturo cestišča. Podobno kot v razdelku 4.1.3, smo na podlagi podatkov pripravili dnevne in nočne modele in tako dobili koeficiente dnevnega in nočnega modela za vsako CVP. Sledil je izračun RMS napake napovedi za vsako CVP za oba modela. Rezultati so predstavljeni v tabeli 4.10 (pod oznako M za dnevni in nočni model). ¶ Po drugi strani pa lahko koeficiente vsake postaje izračunamo tudi s pomočjo atributov, ki se nanašajo na CVP in prej izračunanih koeficientov drugih postaj. Poiščemo matriko  $P$  in jo pomnožimo z atributi CVP, za katero na primer nimamo na voljo meritev. Predvidevamo, da so na ta način pridobljeni koeficienti podobni koeficientom, ki bi jih lahko izračunali, če bi meritve imeli na voljo. Matriko  $P$ , ki smo jo poiskali tako za dnevne kot za nočne modele, smo predstavili na sliki 4.18 s toplotno karto, podobno kot smo to storili na sliki 4.15 v razdelku 4.3.1. Primerjava slednje s sliko 4.18 nam razkrije več podobnosti; ko je korelacija med lastnostmi modelov in lastnostmi lokacij večja, je pogosto večja tudi ustrezna absolutna vrednost v matriki  $P$ . ¶ Tudi za napovedi temperature cestišča, ki smo jih izvedli na podlagi koeficientov, izračunanih iz atributov, smo preverili RMS napako za vsako postajo za oba modela. Rezultati so predstavljeni v tabeli 4.10 (pod oznako A za dnevni in nočni model).

Na problemu napovedovanja temperature cestišča nas je zanimalo, koliko izračun koeficientov iz atributov napako poveča, v primerjavi s koeficienti, ki so izračunani iz meritev. Iz rezultatov je razvidno, da je napaka pri uporabi koeficientov, izračunanih iz atributov, pogosto zgolj nekoliko večja od napake, kjer so bili uporabljeni koeficienti, izračunani iz meritev. To zlasti velja za nočne modele, medtem ko je tudi pri dnevnih modelih v večini primerov povečanje napake še sprejemljivo. Slabše napovedi se kažejo zlasti pri bolj neobičajnih CVP, kot je na primer Črni Kal. ¶ Predvidevamo lahko, da bi vključitev dodatnih atributov (ti bi bili v najboljšem primeru pridobljeni s terenskimi meritvami) rezultate še izboljšalo. Sodeč po rezultatih testiranja metode na umetnih podatkih (razdelek 4.3.3), bi izboljšanje predvidoma dosegli tudi s povečevanjem števila CVP.

Slika 4.18

Matrika podobnosti  $P$  za dnevne in za nočne modele. Temnejša polja predstavljajo večje, svetlejša pa manjše absolutne vrednosti matrike  $P$ . IN pomeni prosti člen regresije, ostale oznake so obrazložene v tabeli 3.1, oznake lastnosti pa v tabeli 3.4. Zaradi lažje primerjave s sliko 4.15 smo obe matriki pripravili brez prostega člena regresije pri lastnostih lokacij.



Na desni strani tabele 4.10 so navedeni še stolpci, ki opisujejo RMS napake napovedi, katere dobimo, če poizkusimo napovedati temperaturo cestišča zgolj z enim dnevnim ali enim nočnim modelom za vse CVP (tega pripravimo na vseh dnevnih oziroma nočnih meritvah vseh CVP), kot tudi z enim skupnim modelom za vse CVP (tega pripravimo na vseh meritvah vseh CVP). Za lažjo primerjavo so na koncu tabele navedene še povprečja RMS napak. ⚡ Opazimo lahko, da se skupni modeli izkažejo nekoliko slabše. Če njihove napake primerjamo z napakami, kjer smo napovedi pripravili na podlagi koeficientov, izračunanimi iz atributov, je napaka skupnih modelov večja. Izjema je dnevni model, kjer je napaka podobna.

Tabela 4.10

Leva stran: RMS napake napovedi temperature cestišča v primerih, ko so koeficienti dnevnih ali nočnih modelov izračunani bodisi iz atributov bodisi iz meritev. Oznaka A pomeni, da so koeficienti izračunani iz atributov, oznaka M pa, da so koeficienti izračunani iz meritev. Desna stran: RMS napake napovedi temperature cestišča z zgolj z enim dnevnim (oziroma enim nočnim) modelom za vse CVP ali enim skupnim (dnevno-nočnim) modelom za vse CVP.

	Dnevni modeli		Nočni modeli		En dnevni model	En nočni model	Skupni model
	A	M	A	M			
Kranj	2,88	2,57	1,51	1,24	2,80	1,68	4,45
Jeprca	3,04	2,84	1,60	1,40	2,96	1,46	4,45
Trojane	3,35	3,07	1,52	1,30	3,66	1,36	4,16
Rimske	4,09	2,60	1,96	1,55	3,06	2,21	3,91
Črnova	3,04	2,22	1,90	1,39	3,20	1,54	3,81
Črni vrh	5,89	1,25	1,28	1,07	8,41	3,11	5,93
Soteska	3,61	3,37	1,28	1,24	4,29	1,77	4,32
Dolsko	3,12	2,99	1,69	1,50	3,16	1,98	4,64
Mislinja	5,18	3,35	1,82	1,63	3,55	1,80	5,01
Črmošnjice	3,14	2,68	1,56	1,56	2,77	1,65	4,50
Kresnica	3,39	3,26	1,88	1,60	3,94	1,73	5,47
Črni Kal	5,21	2,95	3,23	1,16	7,15	3,74	7,60
Devina	3,36	3,28	2,44	1,89	4,15	2,02	6,03
Ravbarkomanda	4,11	3,15	1,46	1,32	3,40	2,25	5,48
Malence	3,52	3,15	1,52	1,44	3,23	1,57	4,67
Slatina	3,27	3,17	1,59	1,40	3,69	1,66	5,24
Verd	4,15	2,98	1,78	1,76	3,08	1,88	4,19
Sava	4,04	3,02	1,52	1,43	3,16	1,79	5,52
Peračica	4,89	3,14	3,27	1,63	4,29	3,43	5,57
Moste	5,88	3,79	1,99	1,65	4,30	1,65	5,93
<i>Povprečje</i>	3,96	2,94	1,84	1,46	3,91	2,01	5,04





## *Zaključek*

V doktorski disertaciji je predstavljena razvita metoda za gradnjo hierarhičnih regresijskih modelov, ki temelji na združevanju vnaprej podanih množic podatkov na podlagi znanja o združljivosti modelov in/ali podobnosti modelov. Metoda v vsakem koraku združi tista dva modela, kjer je takšno združevanje, glede na določeno primerjavo napak, najugodnejše. Z združevanjem lahko prenehamo bodisi ko ta napaka preseže vrednost vnaprej določene dopustne napake bodisi ko dosežemo želeno število modelov. Metoda je preizkušena tako na umetnih podatkih kot tudi na podatkih s področja cestne meteorologije (napovedovanje temperature cestišča). ¶ Rezultati na umetnih podatkih kažejo, da pri večjem številu primerov združevanje ni tako pomembno, ko je primerov dovolj, pa je celo nepotrebno. S tem smo pokazali, da ko imamo na voljo premalo podatkov za ločene modele, lahko z njihovim združevanjem dosežemo boljše rezultate. Z metodo smo uspešno napovedovali temperaturo cestišča za 6 ur vnaprej, napovedi pa smo primerjali z napovedmi fizikalnega modela METRo. Ugotovili smo, da predlagana metoda predstavlja ustrezno alternativo fizikalnim modelom za napovedovanje temperature in stanja cestišča, uporabna pa je lahko tudi na podobnih, bolj splošnih podatkih.

Prilagodili smo postopek za empirično dekompozicijo in analizo razmerja napake na pristranskost in varianco. Postopek smo uporabili za teoretični preizkus delovanja razvite metode za združevanje modelov. Analiza je pri združevanju modelov pokazala, da z naraščanjem napake narašča pristranskost, varianca pa pada. Padanje variance je povezano z večjo množico učnih podatkov, zaradi česar se linearni modeli ne morejo ustrezno prilagoditi podatkom in so, gledano kot celota, manj kompleksni.

Pri analizi povezav med modeli in lastnostmi objektov, na katere se ti modeli nanašajo, smo pripravili metodo, ki predstavlja nadgradnjo metode za združevanje modelov. Ideja nadgradnje je, da z regresijo na podlagi podatkov odvisnih in neodvisnih spremenljivk vsakega modela poiščemo njihove koeficiente, nato pa, prav tako z regresijo, poiščemo še matriko podobnosti na podlagi prej izračunanih vrednosti koeficientov in vrednosti atributov. Iz matrike podobnosti lahko razberemo, kako pomembne so katere povezave med atributi in koeficienti, poleg tega pa nam omogoča, da iz atributov izračunamo koeficiente modelov, ki jih ne poznamo, ali obratno. ¶ Metodo smo uporabili na umetnih podatkih, kjer smo poiskali matriko podobnosti in jo primerjali z matriko, ki je bila uporabljena pri generiranju podatkov. Rezultati so pokazali povečevanje napake pri povečevanju šuma pri koeficientih ali povečevanju šuma pri odvisni spremenljivki in zmanjševanje napake pri povečevanju števila modelov ali povečevanju

števila primerov (podatkov) na model. Izkaže se, da predlagana metoda na osnovi množice objektov in modelov, ki jih opisujejo, uspešno poišče povezave med njimi. Metodno smo na drugačen način uporabili tudi na problemu napovedovanja temperature cestišča, kjer nas je zanimalo, koliko izračun koeficientov iz atributov napako napovedi poveča, v primerjavi s koeficienti, ki so izračunani iz meritev. Iz rezultatov je razvidno, da je napaka pri uporabi koeficientov, izračunanih iz atributov, pogosto zgolj nekoliko večja od napake, kjer so bili uporabljeni koeficienti, izračunani iz meritev. To zlasti velja za nočne modele, medtem ko je tudi pri dnevnih modelih v večini primerov povečanje napake še sprejemljivo. Poleg tega smo te rezultate primerjali še z zgolj enim dnevnim ali z enim nočnim modelom za vse CVP (tega smo pripravili na vseh dnevnih oziroma nočnih meritvah vseh CVP), kot tudi z enim skupnim modelom za vse CVP (tega smo pripravili na vseh meritvah vseh CVP). Takšni skupni modeli so se izkazali nekoliko slabše. Če njihove napake primerjamo z napakami, kjer smo napovedi pripravili na podlagi koeficientov, izračunanih iz atributov, je napaka skupnih modelov večja. Izjema je dnevni model, kjer je napaka podobna. To pomeni, da lahko s pomočjo poznavanja lastnosti objektov pripravimo ustreznejše modele, kot če teh lastnosti objektov ne bi poznali.

### 5.1 Povzetek praktičnih prispevkov

Znanstveni prispevki disertacije so naštetih v razdelku 1.3. Tu povzemimo le še njeno praktično uporabnost.

- Pokazali smo (razdelek 4.1.3), da je mogoče sestavljati dovolj natančne modele za napovedovanje temperature cestišča tudi zgolj s statističnimi metodami, na podlagi izmerjenih podatkov in podatkov vremenskih napovedi. Prednost takšnega modeliranja pred fizikalnim je, da ne zahteva nastavljanja parametrov modela, slabost pa, da zahteva vnaprej zbrane podatke.
- To slabost lahko delno odpravimo z metodo gručenja, ki smo jo razvili v okviru disertacije. Metoda združuje množice podatkov, dokler ne pride do množic, ki dajejo najmanjšo napako modelov. Delovanje metode smo teoretično utemeljili z razcepom napake na varianco in pristranskost (razdelek 4.2), njeno delovanje in pravilnost utemeljitve pa preizkusili tudi empirično (razdelek 4.1.3).
- Takšen postopek modeliranja še vedno zahteva pridobivanje izmerjenih podat-

kov iz CVP, kar pa praktično ni izvedljivo za celotno cestno omrežje. V razdelku 3.1.2 smo zato definirali množico podatkov o lokacijah CVP, ki jih je mogoče zbrati razmeroma enostavno in pokazali, da ti dovolj dobro opisujejo lastnosti lokacij CVP v tem smislu, da imata lokaciji, ki sta si podobni glede na takšen opis, tudi podobne modele za napovedovanje temperature cestišča. V razdelku 4.3.4 smo pokazali, kako lahko tudi za lokacije, na katerih ni CVP, sestavimo napovedne modele na podlagi podatkov, zbranih za druge, podobne lokacije. Kakor kažejo poizkusi, imajo lahko tudi ti modeli zadovoljivo napovedno točnost.

## 5.2 Nadaljnje delo

Pri analizi povezav med modeli in lastnostmi objektov, na katere se ti modeli nanašajo, smo predpostavili, da so te lastnosti poznane vnaprej. Zanimivo bi bilo razviti metodo za iskanje novih lastnosti za opis objektov. Postopek bi bil lahko avtomatski ali ročen. Avtomatski postopek bi sestavljal nove lastnosti na podlagi pogostih zvez med posameznimi parametri in lastnostmi. Na podoben način je mogoče lastnosti sestavljati tudi ročno, pri čemer pa mora računalnik ekspertu pomagati s primerno vizualizacijo povezav med modeli in objekti.

Rezultate metode za gradnjo hierarhičnih regresijskih modelov na problemu napovedovanja temperature cestišča smo primerjali z rezultati fizikalnega modela METRo. Metodo bi bilo smiselno uporabiti tudi kot nadgradnjo fizikalnih modelov in tako izboljšati njihove rezultate. V tem primeru bi z metodo poizkušali napovedovati napako fizikalnega modela na podoben način, kot smo to storili pri napovedovanju temperature cestišča brez uporabe fizikalnega modela.

Pomemben izziv pri napovedovanju stanja cestišča je doseči zadovoljive napovedi tudi na tistih mestih, kjer meritve CVP niso na voljo<sup>1</sup>. V kolikor bi razpolagali z ustreznim številom dovolj raznolikih CVP, bi lahko s predlagano metodo za analizo povezav poiskali matriko podobnosti, ki bi zadovoljivo opisala povezave med modeli in njihovimi lastnostmi lokacij. S pomočjo teh povezav bi lahko pripravili modele na lokacijah, kjer so nam poznane zgolj njihove lastnosti. Če bi podatke o meritvah CVP izključili že iz priprave prvotnih modelov, bi se pri modelih na lokacijah brez CVP lahko izognili njihovi uporabi. Meteorološki modeli postajajo vedno bolj zanesljivi,

<sup>1</sup>Prim. razdelek 2.2.4 na strani 22.

njihove napovedi pa prostorsko in časovno podrobnejše. Tudi zaradi tega razloga bi z uporabo zgolj meteoroloških podatkov v modelih za napovedovanje stanja cestišča na lokacijah brez CVP lahko dosegali zadovoljive rezultate.



*Programska koda za pripravo  
umetnih podatkov*

A

```
1 import functools
2 import os
3 import random
4 import numpy as np
5
6
7 class Model:
8     def __init__(self, attributes, variates, coefs, intercept, P, noise):
9         self.attributes = attributes
10        self.variates = variates
11        self.coefs = coefs
12        self.intercept = intercept
13        self.P = P
14        self.noise = noise
15
16    def gen_x(self):
17        return [v() for v in self.variates]
18
19    def get_y(self, x):
20        return sum(xi*ci for xi, ci in zip(x, self.coefs)) + self.intercept + self.noise()
21
22    def gen_xy(self):
23        x = self.gen_x()
24        y = self.get_y(x)
25        return x, y
26
27    def gen_XY(self, X, Y):
28        X[:] = [self.gen_x() for _ in X]
29        Y[:] = [self.get_y(x) for x in X]
30        return X, Y
31
32
33 def build_models_and_data(all_attrs, P, variables, X, Y):
34     models = []
35     n_points = len(X) / len( all_attrs )
36     for i, attrs in enumerate(all_attrs):
37         # Koeficienti : P * attrs + šum
38         coefs = P.dot(attrs) + np.random.normal(0, 0.2, size=len( variables ))
39         intercept = np.random.normal(0, 0.2)
40         model = Model(attrs, variables, coefs, intercept, P, functools.partial( random.gauss,
41                                     0, 0.1))
42         model.gen_XY(X[i * n_points: (i + 1) * n_points], Y[i * n_points: (i + 1) * n_points])
43         models.append(model)
```



```

43     return models, X, Y
44
45 def build_attributes (n_models, n_cont_attrs, n_disc_attrs):
46     # Matriko A z vrednostmi atributov inicializiram na N(0,1)
47     n_attrs = n_cont_attrs + n_disc_attrs
48     attrs = np.random.normal(size=(n_models, n_attrs))
49     # Diskretne attribute postavimo za prvo polovico modelov na 0, za drugo na 1
50     attrs [: n_models//2, n_cont_attrs:] = 0
51     attrs [n_models//2:, n_cont_attrs:] = 1
52     # Nato jih permutiramo po stolpcih
53     for col in range(n_cont_attrs, n_attrs):
54         np.random.shuffle(attrs[:, col])
55     return attrs
56
57 def build_attrs_to_coefs (n_attrs, n_vars):
58     # Priprava matrike podobnosti P. Vsak koeficient je linearna kombinacija vrednosti
59     # določenega števila atributov. Koliko atributov kombiniramo v koeficient, žebamo
60     # naključno: N(1,1) * število atributov / 3. V kolikor je ta vrednost manjša ali enaka
61     # 0, je koeficient neodvisen od atributov.
62     P = np.zeros((n_vars, n_attrs))
63     for t in range(n_vars):
64         for k in range(int (random.gauss(1, 1) * n_attrs // 3)):
65             P[t, random.randint(0, n_attrs-1)] = random.gauss(0, 1)
66     return P
67
68 def lin_lin (n_models, n_cont_attrs, n_disc_attrs, n_vars, n_points_per_mod, X=None, Y=None):
69     P = build_attrs_to_coefs (n_cont_attrs + n_disc_attrs, n_vars)
70     attrs = build_attributes (n_models, n_cont_attrs, n_disc_attrs)
71     variables = [functools.partial (random.gauss, 0, 1) for _ in range(n_vars)]
72     if X is None:
73         X = np.empty((n_models * n_points_per_mod, n_vars))
74     if Y is None:
75         Y = np.empty(n_models * n_points_per_mod)
76     return build_models_and_data(attrs, P, variables, X, Y)
77
78 def multiple_groups(n_groups, n_models_per_group, n_cont_attrs, n_disc_attrs, n_vars,
79                     n_points_per_mod):
80     tot_points_group = n_models_per_group * n_points_per_mod
81     tot_points = n_groups * tot_points_group
82     X = np.empty((tot_points, n_vars))
83     Y = np.empty(tot_points)
84     models = []
85     for i in range(0, tot_points, tot_points_group):
86         new_models, _, _ = lin_lin (

```

```
83         n_models_per_group,
84         n_cont_attrs,
85         n_disc_attrs,
86         n_vars,
87         n_points_per_mod,
88         X[i : i + tot_points_group],
89         Y[i : i + tot_points_group])
90     models.append(new_models[0])
91     return models, X, Y
92
93
94 def save_models(fname, models):
95     open(fname, "wt").write("\n".join(", ".join(map(str, m.attributes)) for m in models))
96
97 def save_data(fname, X, Y):
98     XY = np.hstack((X, Y.reshape((len(Y), 1))))
99     open(fname, "wt").write("\n".join(", ".join(map(str, xy)) for xy in XY))
100
101 def save_P(fname, models):
102     np.savetxt(fname, models[0].P)
```

---

## SLIKE

2.1	Linearni regresijski modeli z eno neodvisno spremenljivko . . . . .	11
2.2	Cestnovremenska postaja Moste . . . . .	13
2.3	Cestni merilnik Vaisala DRS <sub>511</sub> . . . . .	15
2.4	Cestni merilnik Boschung BOSO II . . . . .	15
2.5	Cestni merilnik Luftt IR <sub>21</sub> . . . . .	16
2.6	Cestna merilnika Vaisala DST <sub>1111</sub> in DSC <sub>1111</sub> . . . . .	17
2.7	Termalni posnetek cestnega odseka v bližini CVP Ravbarkomanda . . . . .	18
2.8	Digitalna kamera, opremljena z lečo ribjega očesa in primer črnobelega fotografije neba . . . . .	21
2.9	Meritve temperature zraka in cestišča vzdolž cestnega omrežja . . . . .	24
2.10	Prikaz nekaterih energijskih tokov na površini cestišča . . . . .	27
2.11	Primer triurne prostorske napovedi temperature cestišča na Danskem . . . . .	31
2.12	Shema slovenskega cestnovremenskega informacijskega sistema . . . . .	32
2.13	DARS cestnovremenski informacijski sistem . . . . .	33
2.14	Fazni diagram vode in <i>NaCl</i> v odvisnosti od koncentracije <i>NaCl</i> in temperature cestišča . . . . .	36
3.1	CVP Mislinja . . . . .	39
3.2	Temperatura cestišča na CVP Mislinja . . . . .	40
4.1	Shematski prikaz prvega koraka postopka združevanja . . . . .	51
4.2	Preizkus metode za gradnjo hierarhičnih modelov na umetnih podatkih . . . . .	52
4.3	Toplotna karta, dendrogram s prikazom postopka združevanja in graf povprečne RMS napake na CVP Jeprca . . . . .	54

4.4	Toplotna karta, dendrogram s prikazom postopka združevanja in graf povprečne RMS napake na CVP Mislinja . . . . .	55
4.5	Toplotna karta, dendrogram s prikazom postopka združevanja in graf povprečne RMS napake na CVP Črmošnjice . . . . .	56
4.6	Korelacija med napako metode in napako temperature zraka . . . . .	59
4.7	Dekompozicija napake metode za gradnjo hierarhičnih regresijskih modelov na podatkih CVP Jeprca . . . . .	65
4.8	Dekompozicija napake metode za gradnjo hierarhičnih regresijskih modelov na podatkih CVP Črmošnjice . . . . .	66
4.9	Gručenje lokacij glede na njihove lastnosti . . . . .	68
4.10	Gručenje dnevnih modelov glede na njihove regresijske koeficiente . . . . .	69
4.11	Gručenje nočnih modelov glede na njihove regresijske koeficiente . . . . .	69
4.12	Graf razdalj med nočnimi modeli in lastnostjo pripadajočih lokacij $\Psi_s$ (vidnost neba) . . . . .	74
4.13	Graf razdalj med nočnimi modeli in lastnostmi $\Psi_s(S)$ , $\Psi_s(J)$ , $\Psi_s(V)$ in $\Psi_s(Z)$ pripadajočih lokacij (posamezne vidnosti neba) . . . . .	75
4.14	Graf razdalj med dnevnimi modeli in lastnostmi pripadajočih lokacij $\Psi_s(S)$ , $\Psi_s(J)$ , $\Psi_s(V)$ in $\Psi_s(Z)$ (posamezne vidnosti neba) . . . . .	76
4.15	Opazovanje povezav med lastnostmi dnevnih in nočnih modelov in lastnostmi lokacij s pomočjo barvne karte korelacij . . . . .	77
4.16	Preizkus metode za iskanje povezav med modeli in atributi na umetnih podatkih (spreminjanje šuma) . . . . .	80
4.17	Preizkus metode za iskanje povezav med modeli in atributi na umetnih podatkih (spreminjanje modelov ali primerov) . . . . .	80
4.18	Matrika podobnosti $P$ za dnevne in za nočne modele . . . . .	82

## TABELE

2.1	Tipični dejavniki, ki vplivajo na temperaturo cestišča . . . . .	19
3.1	Okrajšave vhodnih spremenljivk in njihov pomen . . . . .	41
3.2	Regresijski koeficienti dnevnih modelov, ki se nanašajo na posamezne lokacije ( $M_d$ ) . . . . .	42
3.3	Regresijski koeficienti nočnih modelov, ki se nanašajo na posamezne lokacije ( $M_n$ ) . . . . .	43
3.4	Podatki o lastnostih lokacij CVP ( $L$ ), na katere se nanašajo modeli temperature cestišča . . . . .	44
4.1	Izbrane vhodne spremenljivke z vrednostmi koeficientov končnih modelov . . . . .	57
4.2	Napake 6–urnih napovedi temperature cestišča . . . . .	62
4.3	Matriki evklidskih razdalj za dnevne modele . . . . .	71
4.4	Matriki evklidskih razdalj za nočne modele . . . . .	72
4.5	Matrika evklidskih razdalj za lokacije . . . . .	73
4.6	Korelacije med lastnostmi dnevnih modelov in lastnostmi lokacij . . . . .	75
4.7	Korelacije med lastnostmi nočnih modelov in lastnostmi lokacij, na katere se ti modeli nanašajo . . . . .	75
4.8	RMS napaka modelov glede na združevanje lokacij po lastnosti <i>Kopno</i> in znotraj tega po lastnosti $\Psi_s$ . . . . .	78
4.9	RMS napaka modelov glede na združevanje lokacij po lastnosti <i>Kopno</i> in znotraj tega po lastnosti <i>Vodni vir</i> . . . . .	79
4.10	RMS napake pri uporabi metode za analizo povezav na problemu napovedovanja temperature cestišča . . . . .	83



## LITERATURA

- [1] H. Akaike. A new look at the statistical model identification. *IEEE Transactions of Automatic Control*, 19(6):716–723, 1974.
- [2] A. C. Atkinson. *Plots, Transformations, and Regression*. Oxford University Press, Oxford (Velika Britanija), 1985.
- [3] D. M. Bates in D. G. Watts. *Nonlinear Regression Analysis and Its Application*. Wiley, New York (ZDA), 1988.
- [4] A. Beden in M. Ivačić. Integrated road weather information systems in Slovenia. *13<sup>th</sup> International Road Weather Conference (SIRWEC 2010)*, Quebec (Kanada), februar 2010.
- [5] M. Benko. Road weather forecasts and MDSS in Slovakia. *16<sup>th</sup> International Road Weather Conference (SIRWEC 2012)*, Helsinki (Finska), maj 2012.
- [6] V. J. Berrocal, A. E. Raftery, T. Gneiting in R. C. Steed. Probabilistic weather forecasting for winter road maintenance. *Journal of the American Statistical Association*, 105(490):522–537, 2010.
- [7] M. J. Best. A model to predict surface temperatures. *Boundary Layer Meteorology*, 88(2):279–306, 1998.
- [8] B. Bica, A. Kann in I. Meirold-Mautner. Enhanced road weather warnings and improved communication strategies within central europe as part of the INCA-CE project. *16<sup>th</sup> International Road Weather Conference (SIRWEC 2012)*, Helsinki (Finska), maj 2012.
- [9] K. Blennow. Sky view factors from high-resolution scanned fish-eye lens photographic negatives. *Journal of Atmospheric and Oceanic Technology*, 12(6):1357–1362, 1995.
- [10] J. Bogren. Screening effects on road surface temperature and road slipperiness. *Theoretical and Applied Climatology*, 43(1–2):91–99, 1991.
- [11] J. Bogren in T. Gustavsson. Nocturnal air and road surface temperature variations in complex terrain. *International Journal of Climatology*, 11(4):443–455, 1991.
- [12] J. Bogren, T. Gustavsson, M. Karlsson in U. Postgård. The impact of screening on road surface temperature. *Meteorological Applications*, 7(2):97–104, 2000.
- [13] J. Bogren, T. Gustavsson in S. Lindqvist. A description of a local climatological model used to predict temperature variations along stretches of road. *Meteorological Magazine*, 121(1440):157–164, 1992.
- [14] R. R. Bouckaert. Practical bias variance decomposition. *Lecture Notes in Computer Science*, 5360:247–257, 2008.
- [15] L. Bouilloud, E. Martin, F. Habets, A. Boone, P. Le Moigne, M. Marchetti J. Livet, A. Foidart, L. Franchistéguy, S. Morel, J. Noilhan in P. Pettré. Road surface condition forecasting in France. *Journal of Applied Meteorology and Climatology*, 48(12):2513–2527, 2009.
- [16] A. V. Bradley, J. E. Thornes in L. Chapman. Variation and prediction of urban canyon geometry from sky-view factor transects. *Atmospheric Science Letters*, 2:155–165, 2001.
- [17] A. V. Bradley, J. E. Thornes, L. Chapman, D. Unwin in M. Roy. Modelling spatial and temporal road thermal climatology in rural and urban areas using a GIS. *Climate Research*, 22(1):41–55, 2002.
- [18] L. Breiman, J. Friedman, R. Olshen in C. Stone. *Classification and Regression Trees*. Wadsworth and Brooks, California (ZDA), 1984.
- [19] D. Brunt. *Physical and dynamical meteorology*, poglavje VI, str. 124–146. Cambridge University Press, Cambridge (Velika Britanija), druga izdaja, 1944.
- [20] L. Chapman in J. E. Thornes. A blueprint for 21<sup>st</sup> century road ice prediction. *11<sup>th</sup> International Road Weather Conference (SIRWEC 2002)*, str. 421–436, Sapporo (Japonska), januar 2002.

- [21] L. Chapman in J. E. Thornes. Real time sky-view factor calculation and approximation. *Journal of Atmospheric and Oceanic Technology*, 21(5):730–741, 2004.
- [22] L. Chapman in J. E. Thornes. A geomatics-based road surface temperature prediction model. *Science of the Total Environment*, 360(1–3):68–80, 2006.
- [23] L. Chapman in J. E. Thornes. Small-scale road surface temperature and condition variations across a road profile. *14<sup>th</sup> International Road Weather Conference (SIRWEC 2008)*, Praga (Česka Republika), maj 2008.
- [24] L. Chapman in J. E. Thornes. What spatial resolution do we need for a route-based weather decision support system? *Theoretical and Applied Climatology*, 104(3–4):551–559, 2011.
- [25] L. Chapman, J. E. Thornes in A. V. Bradley. Modelling of road surface temperature from a geographical parameter database. Part 1: Statistical. *Meteorological Applications*, 8(4):409–419, 2001.
- [26] L. Chapman, J. E. Thornes in A. V. Bradley. Modelling of road surface temperature from a geographical parameter database. Part 2: Numerical. *Meteorological Applications*, 8(4):421–436, 2001.
- [27] L. Chapman, J. E. Thornes in A. V. Bradley. Rapid determination of canyon geometry parameters for use in surface radiation budgets. *Theoretical and Applied Climatology*, 69(1–2):81–89, 2001.
- [28] L. Chapman, J. E. Thornes, J.-P. Muller in S. McMuldroch. Potential applications of thermal fish-eye imagery in urban environments. *IEEE Geoscience and Remote Sensing Letters*, 4(1):56–59, 2007.
- [29] L. P. Crevier in Y. Delage. METRo: A new model for road-condition forecasting in Canada. *Journal of Applied Meteorology*, 40(11):2026–2037, 2001.
- [30] T. Cypra in W. Seidl. Maintenance decision support system (MDSS) ASFINAG/Austria — experience of a comprehensive winter maintenance management system. *16<sup>th</sup> International Road Weather Conference (SIRWEC 2012)*, Helsinki (Finska), maj 2012.
- [31] P. Domingos. A unified bias-variance decomposition. *17<sup>th</sup> International Conference on Machine Learning (ICML 2000)*, San Francisco (ZDA), 2000.
- [32] Državni zbor Republike Slovenije. Zakon o cestah (ZCes-1). *Uradni list Republike Slovenije*, 109:16850–16876, 30.12.2010.
- [33] I. Eliasson. Urban nocturnal temperatures, street geometry and landuse. *Atmospheric Environment*, 30(3):379–392, 1996.
- [34] M. Eriksson in J. Norrman. Analysis of station locations in a road weather information system. *Meteorological Applications*, 8(4):437–448, 2001.
- [35] European Committee for Standardization. Winter maintenance equipment — Road weather information systems — Part 3. Preliminary standard prEN 15518-3:2009, European Committee for Standardization, Bruselj (Belgija), 2009.
- [36] D. A. Freedman. *Statistical Models, Theory and Practice*. Cambridge University Press, Cambridge (Velika Britanija), 2009.
- [37] L. Fridstrøm, J. Ifver, S. Ingebrigtsen, R. Kulmala in L. K. Thomsen. Measuring the contribution of randomness, exposure, weather, and daylight to the variation in road accident counts. *Accident Analysis & Prevention*, 27(1):1–20, 1995.
- [38] R. Fry, L. Slade, G. Taylor in I. Davy. A GIS based approach to predicting road surface temperatures. *Lecture Notes in Computer Science*, 4857:16–29, 2007.
- [39] A. Gelman in J. Hill. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, Cambridge (Velika Britanija), druga izdaja, 2007.
- [40] H. Goldstein. *Multilevel Statistical Models*. Edward Arnold, London (Velika Britanija), druga izdaja, 1995.
- [41] V. Golenić, S. Mandžuka in G. Puž. The use of advanced road weather information system in republic of Croatia. *15<sup>th</sup> International Road Weather Conference (SIRWEC 2010)*, Quebec (Kanada), februar 2010.
- [42] I. Grabec, K. Kalcher in F. Švegl. Modeling and forecasting of traffic flow. *Nonlinear Phenomena in Complex Systems*, 13(1):53–63, 2010.
- [43] C. Gregorc in R. Rupar. Sistem VIPOS za zbiranje in obdelavo podatkov o cestni infrastrukturi. *6. slovenski kongres o cestah in prometu*, str. 152–157, Portorož (Slovenija), oktober 2002.
- [44] C. S. B. Grimmond, S. K. Potter, H. N. Zutter in C. Souch. Rapid methods to estimate sky-view factors applied to urban areas. *International Journal of Climatology*, 21(7):903–913, 2001.
- [45] K. Gurney. *An Introduction to Neural Networks*. Routledge, London (Velika Britanija), 1997.
- [46] T. Gustavsson. Variation in road surface temperature due to topography and wind. *Theoretical and Applied Climatology*, 41(4):227–236, 1990.
- [47] T. Gustavsson. Evaluation of a local climatological model — test carried out in the country of Halland, Sweden. *Meteorological Magazine*, 122(1456):257–267, 1993.



- [48] T. Gustavsson. Thermal mapping — a technique for road climatological studies. *Meteorological Applications*, 6(4):385–394, 1999.
- [49] T. Gustavsson in J. Bogren. Thermal mapping by use of infrared technique. *4<sup>th</sup> International Conference on Weather and Road Safety*, Firenze (Italija), november 1988.
- [50] T. Haiden, A. Kann, C. Wittmann, G. Pistotnik, B. Bica in C. Gruber. The integrated nowcasting through comprehensive analysis (INCA) system and its validation over the eastern Alpine region. *Weather and Forecasting*, 26(2):166–183, 2011.
- [51] D. S. Hammond, L. Chapman in J. E. Thornes. Verification of route-based road weather forecasts. *Theoretical and Applied Climatology*, 100(3–4):371–384, 2010.
- [52] P. A. Hansen in S. Brodersen. The Danish road weather information system. *14<sup>th</sup> International Road Weather Conference (SIRWEC 2008)*, Praga (Češka Republika), maj 2008.
- [53] M. Hippi, I. Juga in P. Nurmi. A statistical forecast model for road surface friction. *13<sup>th</sup> International Road Weather Conference (SIRWEC 2010)*, Quebec (Kanada), februar 2010.
- [54] M. Hoffmann, P. Nutz in R. Blab. New findings in winter maintenance and their implementation in Austria. *16<sup>th</sup> International Road Weather Conference (SIRWEC 2012)*, Helsinki (Finska), maj 2012.
- [55] Campbell Scientific Inc. IRS21 Luft intelligent road surface sensor. Priročnik, Campbell Scientific Inc., Utah (ZDA), 2009.
- [56] M. Ivačić, A. Beden, A. Šajin Slak, R. Kršmanec, S. Čarman in M. Korošec. Road weather information systems (RWIS) in Slovenia. *1<sup>st</sup> International Conference on Road and Rail Infrastructure (CETRA 2010)*, str. 979–986, Opatija (Hrvaška), maj 2010.
- [57] G. James in T. Hastie. Generalizations of the bias/variance decomposition for prediction error. Tehnično poročilo, Stanford University, California (ZDA), 1997.
- [58] J. Kazys in P. Kytra. Remote sensors tests on Lithuanian roads. *14<sup>th</sup> International Road Weather Conference (SIRWEC 2008)*, Praga (Češka Republika), maj 2008.
- [59] R. Kohavi in D. H. Wolpert. Bias plus variance decomposition for zero-one loss functions. *Machine Learning: Proceedings of the Thirteenth International Conference (ICML 1996)*, San Francisco (ZDA), 1996.
- [60] J. Kondo in N. Okusa. A simple numerical prediction model of nocturnal cooling in a basin with various topographic parameters. *Journal of Applied Meteorology and Climatology*, 29(7):604–619, 1990.
- [61] K. Korotenko. An automated system for prediction of icing on the road. *Lecture Notes in Computer Science*, 2331:1193–1200, 2002.
- [62] T. Kramberger in J. Žerovnik. A contribution to environmentally friendly winter road maintenance: optimizing road de-icing. *Transportation Research Part D*, 13(5):340–346, 2008.
- [63] R. Kršmanec, A. Šajin Slak, S. Čarman in M. Ivačić. Next steps in forecasting road surface temperature and developing MDSS. *19<sup>th</sup> International Symposium on Electronics in Transport (ISEP 2011)*, Ljubljana (Slovenija), marec 2011.
- [64] R. Kršmanec, A. Šajin Slak, S. Čarman in M. Korošec. METRo model testing at Slovenian road weather stations and suggestions for further improvements. *16<sup>th</sup> International Road Weather Conference (SIRWEC 2012)*, Helsinki (Finska), maj 2012.
- [65] R. Kršmanec, A. Šajin Slak in J. Demšar. Statistical approach for forecasting road surface temperature. *Meteorological Applications*. doi: 10.1002/met.1305.
- [66] S. K. Linden in S. D. Drobot. The evolution of METRo in a roadway DSS. *15<sup>th</sup> International Road Weather Conference (SIRWEC 2010)*, Quebec (Kanada), februar 2010.
- [67] P. C. Mahalanobis. On the generalised distance in statistics. *National Institute of Sciences of India*, 1(2):49–55, 1936.
- [68] A. Mahura, C. Petersen, B. Sass, P. Holm in T. Pedersen. Road stretch weather forecasting: Thermal mapping data applicability. Znanstveno poročilo, Danish Meteorological Institute in Danish Road Directorate, Copenhagen (Danska), 2007.
- [69] T. McClellan, P. Boone in M. A. Coleman. Maintenance decision support system (MDSS). Končno poročilo, Indiana Department of Transportation, Indiana (ZDA), 2009. <http://www.in.gov/indot/files/MDSSReportWinter08-09.pdf>.
- [70] P. McCullagh in J. Nelder. *Generalized Linear Models*. Chapman and Hall/CRC, Boca Rato (Florida), druga izdaja, 1989.
- [71] J. Norrman, M. Eriksson in S. Lindqvist. Relationships between road slipperiness, traffic accident risk and winter road maintenance activity. *Climate Research*, 15(3):185–193, 2000.
- [72] P. Nurmi, M. Hippi in I. Juga. Evaluation of FMI's new forecast model of road surface friction. *15<sup>th</sup> International Road Weather Conference (SIRWEC 2010)*, Quebec, Kanada, februar 2010.

- [73] T. R. Oke. *Boundary Layer Climates*. Routledge, London in New York, 1992.
- [74] E. Pasero in W. Moniaci. Learning and data driver methods for short term meteo forecast. *Lecture Notes in Computer Science*, 3931:105–112, 2006.
- [75] T. S. Pedersen, C. Petersen, K. Sattler, A. Mahura in B. Sass. Physiographic data for road stretch forecasting. *15<sup>th</sup> International Road Weather Conference (SIRWEC 2010)*, Quebec (Kanada), februar 2010.
- [76] K. R. Petty in W. P. Mahoney. The U.S. federal highway administration winter road maintenance decision support system (MDSS): recent enhancements & refinements. *14<sup>th</sup> International Road Weather Conference (SIRWEC 2008)*, Praga (Česka Republika), maj 2008.
- [77] U. Postgård. *Road Climate Variations Related to Weather and Topography*. Doktorska disertacija, University of Gothenburg, Gothenburg (Švedska), 2000.
- [78] D. M. Ramakrishna in T. Viraraghavan. Environmental impact of chemical deicers — a review. *Water, Air, and Soil Pollution*, 166(1–4):49–63, 2005.
- [79] P. J. Rayer. The meteorological office forecast road surface temperature model. *The Meteorological Magazine*, 116(1379):180–191, 1989.
- [80] B. H. Sass. A numerical model for prediction of road temperature and ice. *Journal of Applied Meteorology*, 31(12):1499–1506, 1992.
- [81] J. Shao. The prediction of road surface-state and simulations of the shading effect. *Boundary-Layer Meteorology*, 73(4):411–419, 1995.
- [82] J. Shao. Improving nowcasts of road surface temperature by a backpropagation neural network. *Weather and Forecasting*, 13(1):164–171, 1998.
- [83] J. Shao in P. J. Lister. An automated nowcasting model of road surface temperature and state for winter road maintenance. *Journal of Applied Meteorology*, 35(8):1352–1361, 1996.
- [84] J. Shao, P. J. Lister, G. D. Hart in H. B. Pearson. Thermal mapping — reliability and repeatability. *8<sup>th</sup> International Road Weather Conference (SIRWEC 1996)*, Birmingham (Velika Britanija), april 1996.
- [85] J. Shao, J. C. Swanson, R. Patterson, P. J. Lister in A. N. McDonald. Variation of winter road surface temperature due to topography and application of thermal mapping. *Meteorological Applications*, 4(2):131–137, 1997.
- [86] A. Sherif in Y. Hassan. Modelling pavement temperature for winter maintenance operations. *Canadian Journal of Civil Engineering*, 31(2):369–378, 2004.
- [87] A. Šajin Slak, S. Čarman in R. Kršmanc. Suitability of the new paradigm for winter observation of road condition. *15<sup>th</sup> International Road Weather Conference (SIRWEC 2010)*, Quebec (Kanada), februar 2010.
- [88] A. Šajin Slak, S. Čarman, R. Kršmanc, M. Ivačić, R. Černivec in L. Herga. Vidiki razvoja cestnovremenskega informacijskega sistema. *10. slovenski kongres o cestah in prometu*, str. 478–483, Portorož (Slovenija), oktober 2010.
- [89] A. Šajin Slak in R. Kršmanc. Vpliv posipnih materialov na okolje. *Strokovno srečanje naročnikov in izvajalcev zimskega vzdrževanja cest*, Bled (Slovenija), oktober 2009.
- [90] A. Šajin Slak, R. Kršmanc in S. Čarman. Improved weather information for road sector in Slovenia (INCA-CE project). *16<sup>th</sup> International Road Weather Conference (SIRWEC 2012)*, Helsinki (Finska), maj 2012.
- [91] D. G. Steyn, J. D. Hay, I. D. Watson in G. T. Johnson. The determination of sky-view factors in urban environments using video imagery. *Journal of Atmospheric and Oceanic Technology*, 3(4):759–764, 1986.
- [92] J. G. Sugrue, J. E. Thornes in R. D. Osborne. Thermal mapping of road surface temperatures. *Physics in Technology*, 14:212–213, 1983.
- [93] R. C. Tabony. The variation of surface temperature with altitude. *Meteorological Magazine*, 114(1351):37–48, 1985.
- [94] N. Takahashi, R. A. Tokunaga, M. Asano in N. Ishikawa. Developing a method to predict road surface temperatures —applying heat balance model considering traffic volume—. *13<sup>th</sup> International Road Weather Conference (SIRWEC 2006)*, Turin (Italija), marec 2006.
- [95] N. Takahashi, R. A. Tokunaga, M. Asano in N. Ishikawa. Development and operation of the winter maintenance support system. *14<sup>th</sup> International Road Weather Conference (SIRWEC 2008)*, Praga (Česka Republika), maj 2008.
- [96] B. W. Thompson. Small-scale katabatics and cold hollows. *Weather*, 41(5):146–153, 1986.
- [97] J. E. Thornes. *The prediction of ice on motorways*. Doktorska disertacija, University College London, London (Velika Britanija), 1984.
- [98] J. E. Thornes. Thermal mapping and road-weather information systems for highway engineers. A. H. Perry in L. J. Symons, editors, *Highway Meteorology*, str. 39–67. Taylor and Francis, London (Velika Britanija), 1991.

- [99] J. E. Thornes, G. Cavan in L. Chapman. XRWIS: the use of geomatics to predict winter road surface temperatures in Poland. *Meteorological Applications*, 12(1):83–90, 2005.
- [100] J. Tilley in L. Johannek. Evaluation of Vaisala spectro pavement sensor. Preliminarno poročilo projekta Aurora 2006-04, University of North Dakota, Severna Dakota (ZDA), 2008.
- [101] Vaisala Oyj. Vaisala remote road surface state sensor DSC1111. Priročnik, Vaisala Oyj, Helsinki (Finska), 2009.
- [102] Vaisala Oyj. Vaisala remote road surface temperature sensor DST1111. Priročnik, Vaisala Oyj, Helsinki (Finska), 2009.
- [103] Vaisala Oyj. Vaisala road surface and depth sensor DRS511. Priročnik, Vaisala Oyj, Helsinki (Finska), 2009.
- [104] J. H. Ward. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301):236–244, 1963.
- [105] S. P. White, J. E. Thornes in L. Chapman. *A Guide to Road Weather Information Systems*. University of Birmingham, Birmingham (Velika Britanija), 2006. [http://www.sirwec.org/documents/rwis\\_web\\_guide.pdf](http://www.sirwec.org/documents/rwis_web_guide.pdf).
- [106] F. Wilcoxon. Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6):80–83, 1945.
- [107] L. Wilkinson in M. Friendly. The history of the cluster heat map. *The American Statistician*, 63(2):179–184, 2009.
- [108] Z. Ye, C. K. Strong, X. Shi in S. M. Conger. Analysis of maintenance decision support system (MDSS), benefits & costs. Končno poročilo, Western Transportation Institute, Idaho (ZDA), 2009. [http://www.westerntransportationinstitute.org/documents/reports/AW1408\\_Final\\_Report.pdf](http://www.westerntransportationinstitute.org/documents/reports/AW1408_Final_Report.pdf).