

# Avtomatsko določanje višine temperaturne inverzije iz podatkov vertikalnih sondaž

Jana Faganeli Pucer

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko, Večna pot 113, 1000 Ljubljana, Slovenija  
E-pošta: jana.faganeli@fri.uni-lj.si

**Povzetek.** Dihanje onesnaženega zraka z delci  $PM_{10}$  negativno vpliva na zdravje. Pozimi so izpusti delcev  $PM_{10}$  večji zaradi notranjega ogrevanja, vremenske razmere za disperzijo onesnaževal pa predvsem v kotlinah in dolinah osrednje Slovenije skrajno neugodne. V Sloveniji najvišje vrednosti  $PM_{10}$  izmerijo v dnevih s temperaturno inverzijo, zato so atributi, povezani z njo, zelo pomembni pri napovedovanju dnevnih ravni onesnaženosti. Agencija Republike Slovenije za okolje vsakodnevno izvaja vertikalne sondaže, kjer merijo spreminjanje temperature z nadmorsko višino. Iz podatkov vertikalnih sondaž meteorologi določijo prisotnost temperaturne inverzije in njeno višino.

V našem delu predstavimo avtomatsko metodo za obdelavo podatkov temperatur vertikalnih sondaž, ki temelji na diskretni Morsovi teoriji (ADMT). Rezultat našega postopka so prevojnne točke, v katerih se spremeni smer spreminjanja temperature (naraščanje ali padanje) z nadmorsko višino. Iz tako obdelanega signala izpeljemo štiri attribute, ki kažejo značilnosti temperaturne inverzije. Atributi vključujejo višino in globino inverzije, temperaturno razliko inverzije in temperaturo pri tleh. Vpliv atributov na ravni  $PM_{10}$  določimo s pomočjo bayesovske linearne regresije. Vsi izpeljani atributi izkazujejo signifikanten vpliv na dnevne ravni  $PM_{10}$ . V nadaljnjem delu nameravamo vključiti nove attribute v model za napovedovanje  $PM_{10}$ .

**Ključne besede:** temperaturna inverzija, vertikalna sondaža,  $PM_{10}$ , diskretna Morsova teorija, bayesovska linearna regresija

## Automatic processing of meteorological sounding data to detect the temperature inversion layer

Temperature inversions are common in basins and valleys of central Slovenia increasing  $PM_{10}$  levels. Meteorologists and air quality experts usually assess the temperature inversions from the meteorological sounding data. The paper presents a method based on the discrete Morse theory (ADMT) for automatic processing of meteorological sounding data. The ADMT algorithm is applied to show temperature variations with altitude as a function of critical points taking place at changed direction of the temperature (temperature can increase or decrease with increasing altitude). Based on this data four different features are derived. The derived features comprise the depth, height, and the temperature difference of the temperature inversion. The ground temperature is also used as a feature. Using the Bayesian linear regression the significance of the derived features is assessed. The paper shows a significant impact of the new derived features describing the temperature inversion on the  $PM_{10}$  levels.

**Keywords:** temperature inversion, meteorological sounding,  $PM_{10}$ , discrete Morse theory, Bayesian linear regression

## 1 UVOD

Dihanje onesnaženega zraka negativno vpliva na zdravje. Najbolj je na udaru najranljivejši del populacije (otroci, starejši, kronični bolniki) [1]. Čeprav se v Evropi že veliko let borimo proti povišanim ravnem najbolj

škodljivih onesnaževal zraka, so povišane ravni delcev ( $PM_{10}$ ,  $PM_{2.5}$ ), dušikovih oksidov in ozona še vedno pogoste [2]. V Evropi dopustne ravni posameznih onesnaževal določa Evropska direktiva o kakovosti zunanjega zraka in čistejšem zraku za Evropo [3].

V Sloveniji so se zadnja leta ravni  $SO_2$  in  $NO_2$  občutno znižale, največja težava so povišane vrednosti delcev  $PM_{10}$  in  $PM_{2.5}$ . Se je pa zaradi oglaševanja ogljične nevtralnosti biomase in njene cenovne dostopnosti povečala njena poraba, kar negativno vpliva na koncentracije delcev. V Sloveniji na povišane ravni  $PM_{10}$  pozimi poleg povečanih izpustov zaradi notranjega ogrevanja vplivajo tudi vremenske razmere. Večina mest celinske Slovenije leži v neprepišnih kotlinah ali dolinah, kjer so temperaturne inverzije pozimi zelo pogoste. Ta vpliva na ozračje tako, da se zračna masa meša le do višine, na kateri začne temperatura padati z nadmorsko višino. Zelo nizka temperaturna inverzija pozimi, ko so izpusti visoki, lahko zelo vpliva na povišanje ravni delcev  $PM_{10}$  pri tleh.

Agencija Republike Slovenije za okolje (ARSO) je v Sloveniji pristojna za izvajanje meritev kakovosti zraka in ocenjevanje skladnosti s standardi EU [3]. ARSO je pristojen tudi za opravljanje vseh meteoroloških meritev. Temperaturne profile (spreminjanje temperature z višino) merijo z meteorološkimi sondami (vertikalna sondaža) vsak dan zgodaj zjutraj. Iz tako pridobljenega temperaturnega profila je najlažje vizualno oceniti, ali je

zjutraj prisotna temperaturna inverzija in na kateri višini je. Podatki vertikalnih sondaž so na voljo ravno pred jutranjim objavljanjem napovedi onesnaženosti zraka s  $PM_{10}$ . Stokovnjaki podatke vizualizirajo v obliki grafa temperature v odvisnosti od višine in temperaturno inverzijo ocenijo vizualno. S tem podatkom si pomagajo, ko napovedujejo dnevno povprečno vrednost  $PM_{10}$ .

Temperaturna inverzija je znan atribut pri avtomatskem napovedovanju ravni PM. V [4] smo, podobno kot v [5], temperaturno inverzijo ocenili kot razliko temperature v niže in više ležečem kraju (v Ljubljani in na Krvavcu). Temperaturno inverzijo smo ocenili tudi s pomočjo napovedi meteorološkega modela European Center for Medium-Range Weather Forecasting (ECMWF) [6]. V [7], [8] so za vpliv temperaturne inverzije na raven delcev uporabili napovedi meteorološkega modela. Pomanjkljivost ocene temperaturne inverzije z razliko temperatur v dveh točkah je, da ne moremo oceniti, do katere višine sega. Pomanjkljivost ocene z meteorološkim modelom pa je, da je sam model nepopoln. Resolucija meteorološkega modela je ponavadi premajhna, da bi lahko natančno napovedali temperaturno inverzijo v ozkih dolinah. Oba načina nista primerna za ocenjevanje nizkih temperaturnih inverzij, ki pa najbolj vplivajo na povišane ravni onesnaževal pri tleh.

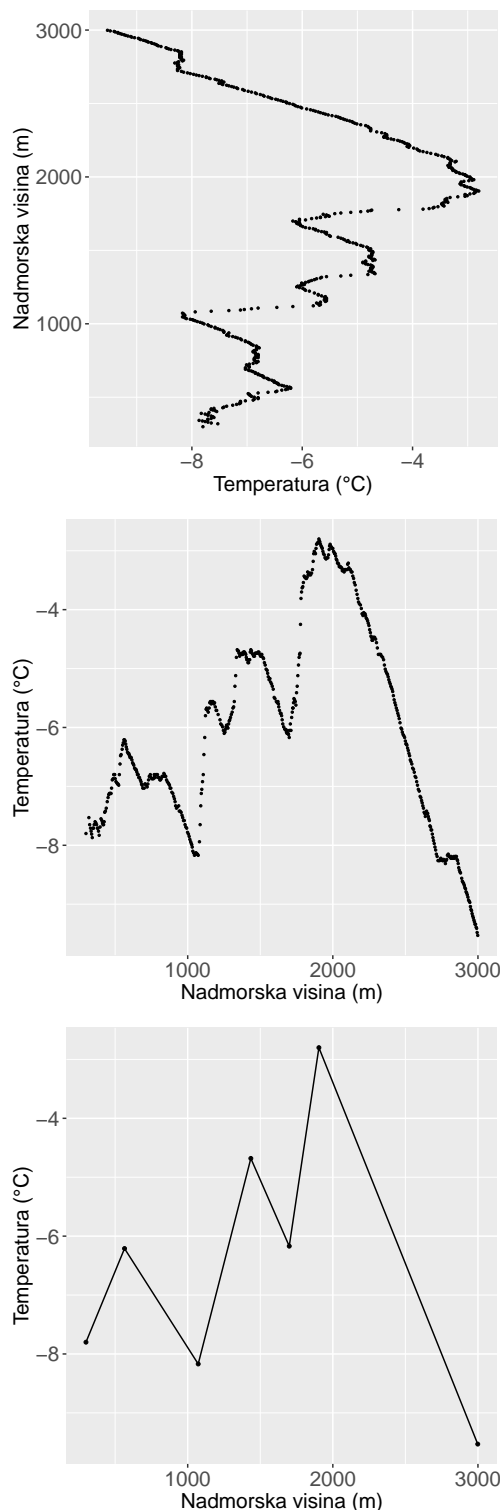
Delo obravnava avtomatsko analizo podatkov vertikalnih sondaž, ki prikazujejo spreminjanje temperature z nadmorsko višino z algoritmom, ki temelji na diskretni Morsovi teoriji (ADMT). Sledi avtomatsko tvorjenje atributov, ki opisujejo temperaturno inverzijo in povezavo teh atributov z dnevnimi ravnimi delcev  $PM_{10}$ .

## 2 MATERIALI IN METODE

### 2.1 Podatki

Vertikalne sondaže izvaja ARSO vsak dan zgodaj zjutraj (večinoma med 4. in 5. uro) s pomočjo meteoroloških sond, ki so opremljene z merilniki zračnega tlaka, temperature, vlage in smeri in jakosti vetra. Sonde ponesejo meteorološki baloni iz strehe ARSO do višine okoli 18000 metrov in pošiljajo podatke v realnem času s pomočjo radijskih signalov (radijska telemetrija). V našem delu nas zanima spreminjanje temperature z nadmorsko višino, kot prikazuje slika 1. Na onesnaženost zraka najbolj vpliva dogajanje v nižjih plasteh ozračja, ker ležijo mesta celinske Slovenije večinoma na okoli 300 metrov nadmorske višine, za analizo izberemo zračno maso do višine 3000 metrov. Kot vhodni podatek vzamemo točke, ki prikazujejo povezavo višine in temperature od višine 299 m do 3000 m nadmorske višine.

Meritve onesnaženosti zraka izvaja ARSO na različnih lokacijah v Sloveniji, ena od postaj je pri sedežu ARSO v Ljubljani, za Bežigradom. Za analizo smo uporabili meritve delcev  $PM_{10}$  v obdobju od junij 2012 do marca 2019. Podatki so na voljo kot dnevne povprečne vrednosti v  $\mu\text{g}/\text{m}^3$ .



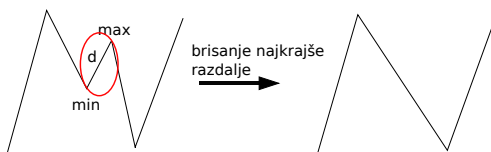
Slika 1: Slike prikazujejo spreminjanje temperature z nadmorsko višino 26. januarja 2019. Na zgornji sliki je standardni prikaz spreminjanja temperature z nadmorsko višino, srednja slika prikazuje rotacijo, ki jo naredimo, da lahko na podatke apliciramo algoritem ADMT. Spodnja slika prikazuje rezultat postopka ADMT. Vertikalna sondaža je zdaj opisana le z ekstremnimi točkami.

## 2.2 Diskretna Morsova teorija in njena enodimenzionalna adaptacija

Zvezna Morsova teorija [9] je postopek iz diferencialne topologije. Omogoča nam neposredno obravnavo gladkih mnogoterosti s pomočjo analize odvedljivih funkcij. Diskretna Morsova [10] teorija (ADMT) je kombinatorična prilagoditev Morsova teorije. Temelji na diskretnih gradientnih vektorjih, katerih kritični elementi opisujejo topologijo strukture. Je pripomoček, ki nam pomaga določiti ekvivalentnost med topološkimi prostori, ki nastanejo iz diskretnih matematičnih struktur. Opisuje, kako lahko matematično strukturo opišemo z njenimi kritičnimi točkami. Teorija je bila uporabljena v različnih aplikacijah: za kompresijo mrež [11], obdelavo slik [12], učenje kvalitativnih modelov [13] in obdelavo signalov EKG [14].

Za obdelavo podatkov vertikalnih sondaž uporabimo enodimenzionalno prilagoditev diskretne Morsova teorije [15], kot smo jo uporabili v [14] in smo jo poimenovali adaptacija diskretne Morsova teorije (ADMT). V našem primeru osnovni algoritem [15] nekoliko spremenimo. V nasprotju s [15] in tako kot v [14] nikoli ne izbrišemo začetne in končne točke (skrajnih nadmorskih višin). V [14] iščemo razliko amplitud lokalnih minimumov in maksimumov (y-os), pri vertikalnih sondažah pa nas zanimajo razlike med minimumi in maksimumu po x-osi (temperature), zato signal zarotiramo (glej srednjo sliko 1). Algoritem najprej poišče vse lokalne minimume in maksimume signala in izračuna razliko med njihovimi amplitudami. Glavna značilnost algoritma je brisanje, odstranjevanje parov minimum-maksimum, ki so med seboj oddaljeni (po y-osi) za manj kot vnaprej podan prag. Slika 2 prikazuje brisanje najkrajše razdalje v sekvenci minimumov in maksimumov, v tem primeru se izbrišeta tudi pripadajoča minimum in maksimum in razdalji levo od minimuma in desno od maksimuma (glej sliko 2). Ustvari se nova razdalja med preostalima nepovezanima ekstremoma.

V našem primeru razdalje imenujemo razlike v temperaturi med minimumi in maksimumi (y-os zarotiranega signala). Za prag smo izbrali vrednost  $1^\circ\text{C}$  oziroma brisanje končamo, ko nam ostaneta samo en minimum in maksimum (ena razdalja). Za prag izberemo  $1^\circ\text{C}$  in ker so razlike manjše od  $1^\circ\text{C}$  je razlika tako majhna, da še ne moremo govoriti o temperaturni inverziji.



Slika 2: Slika prikazuje brisanje najkrajše razdalje v sekvenci petih razdalij (6 ekstremov).

**Algoritem 1:** Enodimenzionalna adaptacija diskretne Morsove teorije (ADMT), ki smo jo uporabili za obdelavo vertikalnih sondaž.

**Data:**

- seznam točk, parov  $(x_i, y_i)$ , kjer je  $y_i = f(x_i)$ .
- vrednost praga.

**Result:** seznam kritičnih točk, ki ostanejo, ko končamo krajšanje.

Lmin=vsi lokalni minimumi Lmax=vsi lokalni maksimumi Če ima več sosednjih točk  $(x_j, \dots, x_k)$  isto vrednost ( $y_j = \dots = y_k$ ), določimo, da je samo zadnja, torej  $(x_k, y_k)$ , kritična točka.

dy=vse razdalje sosednjih Lmin in Lmax, razdalja je mišljena po y osi P=določen prag **while**  $\min(dy) < P$

```

do
  if  $\min(dy)$  je začetna ali končna razdalja then
    ohranimo robno točko;
    izbrišemo njej sosedno točko;
    povežemo preostali ekstrem s končno točko;
  else
    izbrišemo  $\min(dy)$ -par minimum in
    maksimum in povezave med njima;
    povežemo preostala ekstrema;
  end
end

```

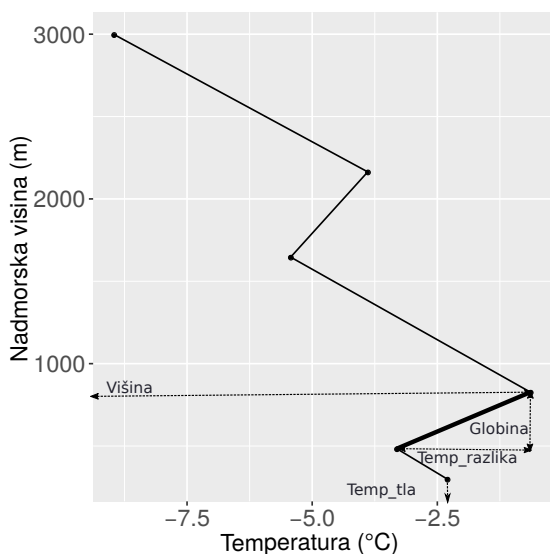
## 2.3 Izračun atributov

Po aplikaciji postopka ADMT (brisanjem) nam ostanejo le pari minimum-maksimum, ki so med seboj oddaljeni za več kot  $1^\circ\text{C}$  (glej spodnjo sliko 1). Attribute določimo tako, da čim bolj ponazarjajo značilnosti, ki jih strokovnjaki opazujejo na grafih vertikalnih sondaž, ko napovedujejo ravni delcev  $\text{PM}_{10}$ .

Strokovnjake zanima najnižja temperaturna inverzija (plast zraka pri tleh), zato iščemo prvo razdaljo (med minimumom in maksimumom), kjer temperatura narašča z nadmorsko višino. Na sliki 3 je taka razdalja narisana odebeltano. Atributi so prikazani na sliki 3:

- Ekstremno točko, ki ustreza maksimumu temperature na tej razdalji, določimo kot atribut in ga poimenujemo Višina.
- Temperaturno razliko, ki je enaka izbrani razdalji (glej sliko 3, odebeltano razdaljo), imenujemo Temp\_razlika.
- Razliko v višinah ekstremov, ki pripadata tej razdalji, imenujemo Globina.
- Ker vemo, da temperatura zraka pri tleh vpliva na raven delcev (bolj je hladno, več je izpustov notranjega ogrevanja), dodamo kot atribut še temperaturo, ki jo sonda izmeri pri tleh (Temp\_tla).

Če v obdelanih podatkih vertikalne sondaže ni nobene razdalje, kjer temperatura narašča, določimo, da je Temp\_razlika enaka 0. Atributoma Višina in Globina pa dodelimo vrednost 5000, ki je zunaj opazovanega območja. Atributoma ne želimo prirediti vrednosti 0, saj iz prakse vemo, da nižja ko je inverzija, manj vpliva na



Slika 3: Rezultati vertikalne sondaže z dne 8. januarja 2019 po obdelavi z ADMT. Označeni so izpeljani atributi.

ravni  $PM_{10}$ .

#### 2.4 Določanje vpliva atributov na koncentracije delcev $PM_{10}$

Vpliv izpeljanih atributov na ravni delcev  $PM_{10}$  določimo s pomočjo bayesovske linearne regresije [16]. Bayesovska linearne regresija nam poda iste odvisnosti med atributi in izhodno spremenljivko kot navadna linearne regresija. Razlika je, da nam bayesovska linearne regresija ne da točkovnih napovedi, temveč napovedi v obliki verjetnostnih porazdelitev. Podobno tudi parametri regresije niso več določeni kot optimalne točke (v smislu najmanjše kvadratične napake), ampak kot aposteriorne porazdelitve. Tak pristop nam omogoči določanje negotovosti parametrov. Za prilagajanje bayesovske linearne regresije smo uporabili R [17] paket arm [18].

Bayesovsko linearne regresijo prilagajamo tako, da vzamemo prej izpeljane attribute kot vhodne spremenljivke, dnevne povprečne ravni  $PM_{10}$  pa kot izhodno spremenljivko. Dneve, za katere ni na voljo katera od vhodnih spremenljivk ali izhodna spremenljivka, odstranimo. Za vsak atribut določimo vpliv tako, da opazujemo njegovo povprečno vrednost (dobimo enak rezultat kot za koeficiente navadne linearne regresije) in njihov 95-odstotni aposteriorni bayesovski interval zaupanja. Če interval ne vsebuje ničle, lahko sklepamo, da je vpliv atributa na izhodno spremenljivko znaten.

### 3 REZULTATI

Vsi izpeljani atributi izkazujejo znaten vpliv na izhodno spremenljivko (koncentracijo  $PM_{10}$ ), saj njihov bayesovski interval zaupanja ne vsebuje vrednosti 0. Vplivi atributov so pričakovani, temperatura tal negativno vpliva

| Atribut      | Srednja vrednost | CI025   | CI975   |
|--------------|------------------|---------|---------|
| Temp_tla     | -1.05            | -1.12   | -0.98   |
| Temp_razlika | 2.22             | 2.00    | 2.44    |
| Višina       | -0.0046          | -0.0055 | -0.0037 |
| Globina      | 0.0035           | 0.0027  | 0.0043  |

Tabela 1: Tabela prikazuje rezultate ocenjenega vpliva atributov na dnevno koncentracijo delcev  $PM_{10}$ . Parameter "Srednja vrednost" pomeni srednjo vrednost koeficienta v linearni regresiji, CI025 in CI975 pa meji 95 % bayesovskega aposteriornega intervala zaupanja.

na ravni  $PM_{10}$ , saj bolj ko je hladno, več je izpustov zaradi notranjega ogrevanja in višje so ravni  $PM_{10}$ . Temperaturna razlika med minimalno in maksimalno temperaturo v temperaturni inverziji izkazuje pozitiven vpliv, večja ko je temperaturna razlika, višje so ravni  $PM_{10}$ . Če je večja razlika med minimalno in maksimalno temperaturo v temperaturni inverziji, je manjša verjetnost, da se bo temperaturna inverzija tekom dneva "razbila". V tam primeru je zelo velika verjetnost, da bo temperaturna inverzija trajala in tako negativno vplivala na ravni delcev  $PM_{10}$ . Če je temperaturna razlika manjša, se temperaturna inverzija tekom dneva "razbije", mešalna plast zraka se dvigne in ravni  $PM_{10}$  padejo. Višina inverzije (Višina) izkazuje negativen vpliv, torej nižje ko je prevojna točka inverzije (maksimalna temperatura, preden začne temperatura padati z višino), manjši volumen ima plast, v kateri se zrak meša. Globina inverzije izkazuje pozitiven vpliv, večja ko je plast inverzije (po višini), višje so koncentracije. Globina in Višina sta odvisna atributa. Če je najnižja plast inverzije tudi prva opazovana razdalja, opisujeta atributa isto. Kljub temu atribut Globina izkazuje signifikanten vpliv v linearni regresiji, kar pomeni, da njegova vključitev prinese dodatno znanje v model.

### 4 SKLEP

V članku smo predstavili postopek za avtomatski izračun atributov iz podatkov vertikalnih sondaž (spreminjanje temperature z višino). Uporabili smo postopek iz diferencialne topologije, ki smo ga prej uspešno uporabili za obdelavo signalov EKG [14], imenovan ADMT. V tem primeru nismo uporabili implementacije s kopico, saj so bili tukaj signali veliko krajši, zato je bila direktna implementacija dovolj hitra. Atributi, ki smo jih izpeljali po predhodni obdelavi signala z ADMT, so izkazali signifikanten vpliv pri modeliranju dnevne ravni  $PM_{10}$ . Izkazovali so podoben vpliv, kot ga poznamo iz prakse.

V nadaljnjem delu bomo izpeljane attribute vključili v model za napovedovanje delcev  $PM_{10}$ . Prednost postopka je, da popolnoma avtomatizira tvorjenje atributov iz podatkov vertikalnih sondaž, tako da lahko takšne attribute vključimo v model, ki deluje v operativi. Pridobljeni atributi so uporabni tudi za izdelavo letnih statistik, opis značilnosti in trendov temperaturnih inverzij.

Signal, ki ga obdelamo z ADMT, je tudi kompaktnjša oblika zapisa signala (informacije zasedejo manj prostora).

## ZAHVALA

Zahvaljujem se Nevi Pristov iz Sektorja za meteorološko, hidrološko in oceanografsko modeliranje Agencije RS za okolje za podatke vertikalnih sondaž in Sektorju za kakovost zraka za ravni PM<sub>10</sub>.

**Jana Faganeli Pucer** je leta 2007 diplomirala na Fakulteti za elektrotehniko Univerze v Ljubljani. Leta 2014 je doktorirala na Fakulteti za računalništvo in informatiko Univerze v Ljubljani, kjer je bila mlada raziskovalka. Med letoma 2014 in 2017 je delala na Agenciji Republike Slovenije za okolje. Od leta 2017 je zaposlena kot asistentka na Fakulteti za računalništvo in informatiko Univerze v Ljubljani.

## LITERATURA

- [1] World Health Organization *et al.*, "Health aspects of air pollution with particulate matter, ozone and nitrogen dioxide: report on a WHO working group, Bonn, Germany 13-15 January 2003," 2003.
- [2] European Environment Agency, "Air quality in Europe-2018 report," 2017.
- [3] European Council, "Directive 2008/50/EC of the European Parliament and of the Council," *Decision of Council*, 2008.
- [4] J. Faganeli Pucer, G. Pirš, and E. Štrumbelj, "A bayesian approach to forecasting daily air-pollutant levels," *Knowledge and Information Systems*, vol. 57, no. 3, pp. 635–654, 2018.
- [5] E. Stadlober, S. Hörmann, and B. Pfeiler, "Quality and performance of a pm10 daily forecasting model," *Atmospheric Environment*, vol. 42, no. 6, pp. 1098–1109, 2008.
- [6] F. Molteni, R. Buizza, T. N. Palmer, and T. Petroliaigis, "The ECMWF ensemble prediction system: Methodology and validation," *Quarterly journal of the royal meteorological society*, vol. 122, no. 529, pp. 73–119, 1996.
- [7] M. A. Pohjola, M. Rantamäki, J. Kukkonen, A. Karppinen, and E. Berge, "Meteorological evaluation of a severe air pollution episode in helsinki on 27-29 december 1995," *Boreal Environment Research*, vol. 9, no. 1, pp. 75–87, 2004.
- [8] P. Perez and J. Reyes, "An integrated neural network model for pm10 forecasting," *Atmospheric Environment*, vol. 40, no. 16, pp. 2845–2851, 2006.
- [9] J. W. Milnor, *Morse theory*. No. 51, Princeton university press, 1963.
- [10] R. Forman, "A user's guide to discrete Morse theory," *Sém. Lothar. Combin.*, vol. 48, p. 35, 2002.
- [11] T. Lewiner, H. Lopes, and G. Tavares, "Applications of Forman's discrete Morse theory to topology visualization and mesh compression," *IEEE Transactions on Visualization and Computer Graphics*, vol. 10, no. 5, pp. 499–508, 2004.
- [12] V. Robins, P. J. Wood, and A. P. Sheppard, "Theory and algorithms for constructing discrete Morse complexes from grayscale digital images," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 8, pp. 1646–1658, 2011.
- [13] J. Žabkar, I. Bratko, G. Jerše, J. Prankl, and M. Schlemmer, "Learning qualitative models from image sequences," in *22nd International Workshop on Qualitative Reasoning*, pp. 146–149, 2008.
- [14] J. Faganeli Pucer and M. Kukar, "A topological approach to delineation and arrhythmic beats detection in unprocessed long-term ecg signals," *Computer methods and programs in biomedicine*, vol. 164, pp. 159–168, 2018.
- [15] J. Žabkar, *Učenje kvalitativnih odvisnosti*. PhD thesis, Univerza v Ljubljani, Fakulteta za računalništvo in informatiko, 2010.
- [16] P. D. Hoff, *A first course in Bayesian statistical methods*. Springer Science & Business Media, 2009.
- [17] H. Wickham, R. François, L. Henry, and K. Müller, *dplyr: A Grammar of Data Manipulation*, 2018. R package version 0.7.6.
- [18] A. Gelman, Y.-S. Su, M. Yajima, J. Hill, M. G. Pittau, J. Kerman, T. Zheng, and V. Dorie, "arm: Data analysis using regression and multilevel/hierarchical models," *R package, version 9.01*, 2009.