

Strojno učenje za boljše javne storitve: zgodnja identifikacija iztokov iz omrežja pitne vode

Sašo Karakatič¹, Špela Pečnik¹, Grega Vrbancič¹, Rok Kukovec, Matej Levstek², Aleš Erker², Vili Podgorelec¹

¹ Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko, Maribor, Slovenija
saso.karakatic@um.si, spela.pecnik@um.si, grega.vrbancic@um.si, rok.kukovec@um.si, vili.podogrelec@um.si

² Mariborski vodovod, javno podjetje, d.o.o., Maribor, Slovenija
matej.levstek@mb-vodovod.si, ales.erker@mb-vodovod.si

Sinopsis Pitna voda je redek in dragocen vir, zato je ustrezno upravljanje tega vira bistveno za družbeni in gospodarski razvoj vsake države ali regije, saj je čista pitna voda uporabljena v vseh ključnih sektorjih gospodarstva, kot so kmetijstvo, industrija, energetika in promet. Pametne naprave so spremenile skoraj vse vidike našega življenja – in tudi pri upravljanju vodovodnega omrežja ni nič drugače. Moderni vodovodni sistemi so že opremljeni z napravami za sprotno beleženje pretokov skozi posamezne odseke vodovodnega omrežja. Prav tako so pogosti tudi sistemi, ki informacije o pretokih v obliki nadzornih sistemov prikazujejo v realnem času in s tem omogočajo spremljanje celotnega omrežja v vsakem trenutku. Še vedno pa za odkrivanje nepravilnosti v vodovodnem omrežju skrbijo bodisi človeški kontrolorji, ali pa primitivni sistemi na osnovi pravil. Pri človeških kontrolorjih je problem, da sta njihova čas in pozornost omejena. Prav tako pa je kontrolorja večšega identifikacije iztokov potrebno dodobra izobraziti. Obstoječi avtomatski sistemi pa so preveč primitivni, da bi zaznali kaj več kot nenadne ogromne iztoke – manjši iztoki, ki na dolgi rok naredijo veliko škode, pa ostanejo neodkriti. V tem članku je predstavljen sistem za zgodnje odkrivanje iztokov, ki temelji na tehnikah napredne analize podatkov in strojnem učenju. Sistem, ki bo opisan temelji na različnih tehnikah, kot so gručenje podobnih odsekov omrežnega sistema ter napoved pričakovane porabe vode določenega področja glede na letni čas in porabo v ostalih podobnih odsekih. Prvi rezultati kažejo, da je sistem zelo koristno orodje, ki kontrolorjem pomaga pri identifikaciji iztokov, ter te samodejno išče tudi v času, ko človeški kontrolorji niso na delovnem mestu.

Ključne besede:

Strojno učenje

Iztoki pitne vode

1 Uvod

Pitna voda je naravni vir in dobrina, katere pomen in potreba v svetu stalno narašča, zato učinkovito upravljanje celotnega vodovodnega sistema, od vodnih virov do končnih odjemalcev, postaja ključnega pomena. Izgubljena voda pomeni nepotrebno obremenjevanje vodnih virov in nepotrebno porabo električne energije, potrebne za črpanje in distribucijo vode do odjemalcev. Vodne izgube je potrebno zniževati in nadzorovati, saj pomenijo prihranek na vodnih virih z racionalizacijo stroškov in investicij ter zmanjšanje posegov v okolje in prostor v trajnostnem razvoju sistema. Predstavljajo izziv v vseh vodovodnih sistemih, v industrijskih območjih pa povzročajo tudi izpad proizvodnih procesov. V svetu se zaradi zastarele vodovodne infrastrukture izgubi v povprečju od 25 % do 50 % vode. V državah Evropske unije se ti deleži gibljejo od 2 % pa tudi do 60 %, v nekaterih evropskih državah pa tudi do 85 %. Učinkovito zmanjševanje vodnih izgub je tako eden glavnih izzivov vseh upravljalcev vodovodnih sistemov.

Mednarodna združenja za vodo in njihove smernice priporočajo zamenjavo najmanj dveh procentov najbolj dotrajane vodovodne infrastrukture na leto. Glede na trend v zadnjih letih vedno manjših vlaganj v vodovodno infrastrukturo in posledično manjšega števila rekonstrukcij cevovodov, tega cilja nismo dosegali, zato vodne izgube predstavljajo vedno večji izziv. Stremimo k temu, da stanje vodovodnega sistema, tudi z uporabo in preizkušanjem novih tehnologij in pristopov, optimiziramo.

Dodatni izziv predstavlja pomanjkanje zadostnih količin primernih kadrovskih virov, ki bi se ukvarjali z aktivnim pregledom iztekanj na terenu in z njihovim lociranjem. Rezultati novih tehnologij bi tako raziskovalce na terenu lahko dodatno obremenili, po drugi strani pa pripomogli pri učinkovitejših oz. natančneje usmerjenih pregledih.

1.1 Predlagana rešitev s strojnimi učenjem

Identifikacije iztokov iz vodovodnega omrežja navadno poteka, kar ročno preko pregleda vzorcev porabe pitne vode posameznih merilnih področij. Je pa tak ročen pregled problematičen, saj se večji iztoki ugotovijo prepozno, manjši pa so mnogokrat spregledani. Hkrati pa se mnogokrat najde lažen iztok vode, ki nepotrebno preobremeni ekipe na terenu. Zato je učinkovita identifikacija spretnost, ki jo operater mora razviti (največkrat dolgotrajen postopek ob obilici pomoči izkušenejših operaterjev in ekip na terenu), da se mu ne izmuznejo iztoki in, da ne sproži preveč lažnih alarmov.

Tako delo, identifikacija iztokov iz vodovodnega omrežja, je tako idealno situirano za avtomatizacijo, saj je delo monotono, ter zahteva obilo spretnosti in izkušenj. Do določene mere so postopki avtomatizacije identifikacije iztokov že vpeljani v raznorazne sisteme nadzora vodovodnega omrežja. Ampak največkrat je avtomatska identifikacija implementirana le v obliki preprostih pravil, ki so zapisana v stilu: »če je poraba ta teden dvakrat večja od porabe prejšnjega tedna, sproži alarm«. Taka naivna pravila sprožijo preveč lažnih alarmov in izpustijo preobilico pravih iztokov. Posledično so ob poplavi takih nepravilnih signalov naivni avtomatski sistemi operaterjem le v breme in ne v pomoč. [1]

Prav z namenom nadgradnje sistemov avtomatske identifikacije iztokov iz vodovodnega omrežja, je potrebno najti alternativne implementacije, ki ne delujejo po naivnih pristopih. Algoritmi strojnega učenja so ena izmed potencialnih rešitev, ki lahko pomagajo pri tem je »inteligentnejši« identifikaciji iztokov. Namreč, ne temeljijo na naivnih pravilih, ki bi jih vnaprej definirali ljudje (strokovnjaki ali programerji) ampak omogočajo sprotno prilagajanje pravil, hkrati pa imajo sposobnost delovanja po vzorcih iz vodovodnega omrežja, ki jih z naivnimi pravili ni možno opisati. [1]

V tem prispevku je opisan sistem, ki prav z uporabo algoritmov strojnega učenja uspe učinkovito identificirati iztoke iz vodovodnega omrežja pitne vode, ter s tem pomaga operaterjem sprejemati boljše odločitve za sprožitve alarmov ter posledično bolj učinkovito porabo časa ekip na terenu.

2 Iztoki pitne vode

Več kot 1.672 km razvejanega vodovodnega omrežja nas zavezuje k odgovornemu vzdrževanju celotnega sistema. Nepredvideni dogodki (prelomi cevi, puščanja...), ki povzročajo iztok oz. vodne izgube se večinoma pojavljajo zaradi dotrajane vodovodne infrastrukture. V analizah izgub vode sledimo smernicam mednarodnih združenj IWA (International Water Association), AWWA (American Water Works Association) in WHO (World Health Organization), ki obravnavajo obvladovanje vodnih izgub z upoštevanjem metod v analizi in kontroli sistema. V ta namen se tudi na območju RS teži k spremljanju vodne bilance po metodologiji IWA, kar se zahteva tudi po Uredbi o oskrbi s pitno vodo (Ur.l. 88/2013). Iz vodne bilance je razvidno, da je vtok v sistem enak vsoti delov prodane in neprodane vode, ki je vsota deležev:

- Neobračunane avtorizirane porabe (razlika med odčitki na števcih in prodano vodo).
- Navideznih izgub, ki so delno:
 - Neavtorizirana poraba (priključki na črno ali javna raba) in
 - posledica nenatančnih meritev (slaba merilna mesta oz. vodomeri, neustrezna kvaliteta, slaba proizvodnja, vzdrževanje ali dimenzioniranje) in
- Dejanskih izgub vode, ki so odraz stanja vodovodnega sistema oz. omrežja z vodooskrbnimi objekti in se pojavljajo na:
 - Vodih surove vode in sistemih za obdelavo vode,
 - transportnih in razdelilnih vodih in
 - priključkih do merilnega mesta.

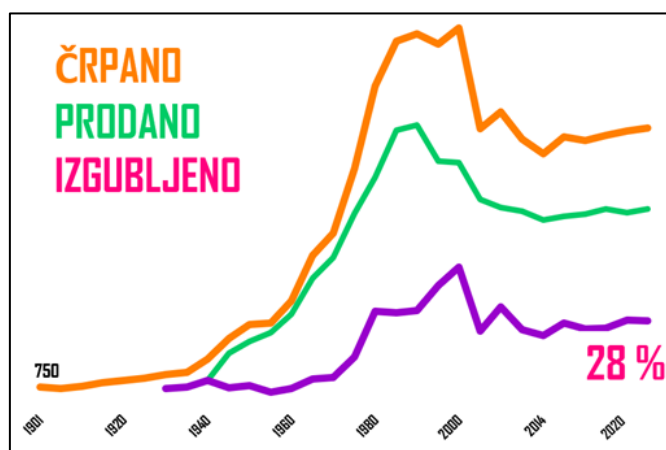
Metodologija IWA priporoča uporabo različnih načinov za analizo vodnih izgub (izgube na odjemalca, izgube na km cevovoda, UARL, CARL, ILI...). Zaradi lažjega prikaza in primerjave z drugimi sistemi se po izvedbi vodne bilance določijo obstoječe letne dejanske vodne izgube (CARL – Current Annual Real Losses). Za zmanjševanje dejanskih vodnih izgub je pomembno ugotavljati tudi neizogibne letne izgube (UARL – Unavoidable Annual Real Losses), ki predstavljajo minimalne podzemne izgube, katere je skoraj nemogoče odkriti z uporabo običajnih metod in naprav. Za ugotavljanje uspešnosti reševanja dejanskih izgub v vodooskrbnem sistemu se uporabi kazalnik iztekanja ILI (Infrastructure Leakage Index), ki predstavlja razmerje med obstoječimi dejanskimi letnimi izgubami in neizogibnimi letnimi izgubami in prikazuje, kako dobro je sistem vzdrževan. Skladno s tem je sistem Mariborskega vodovoda smiselno razdeljen po blokih:

- Blok I: Mestna občina Maribor in občine Miklavž, Hoče - Slivnica in Duplek,
- Blok II: Občini Ruše in Selnica ob Dravi,
- Blok III: Občine Pesnica, Kungota, Šentilj, Lenart, Sveta Ana v Slov. goricah, Sv. Jurij v Slov. goricah, Sv. Trojica v Slov. goricah, Benedikt, Cerkevjenjak in Gornja Radgona.

Tabela 1: Kazalnik ILI in delež izgubljene vode v letu 2021.

Opis	Črpano	V omrežje [m3]	Prodano [m3]	Vodne izgube [m3]	Dej.izgube CARL [m3]	UARL [m3]	Vodne izgube [%]	ILI
BLOK I	10.638.182	10.546.383	7.440.228	3.106.156	3.012.971	602.785	29,5	5,0
BLOK II	650.623	644.804	489.870	154.934	150.286	70.385	24	2,1
BLOK III	2.551.280	2.538.187	1.950.856	587.331	569.712	423.272	23,1	1,3
SKUPNO	13.840.085	13.729.374	9.880.953	3.848.421	3.732.969	1.116.036	28	3,3

Kazalnik ILI nam daje zadovoljiv podatek, da je sistem kot celota dobro vzdrževan. Kaže, da so vodne izgube v sistemu kot celoti, dobro upravljane, vendar obstaja potencial za izboljšave. Stanje sistema na področju Blokov II in III je zadovoljivo, nadaljnje zmanjševanje izgub na tem delu je lahko ekonomsko neupravičeno, posebno pozornost pa je potrebno nameniti področju Bloka I. Višji kazalnik ILI pripisujemo predvsem dejstvu, da je to urbano območje večje poselitve z večjo gostoto ter več gospodarskimi/industrijskimi odjemalci, kar posledično pomeni večji odjem vode na km cevovoda. Stremimo k temu, da sistem čim bolj vzdržujemo s sistemsko analizo in vlaganji tako v infrastrukturo, kot v razvoj ter s tem znižujemo vodne izgube ter porabo električne energije.



Slika 1: Količina črpane, prodane in izgubljene vode po letih.

Vir: lasten.

»Dejanske izgube« so rezultat slabega stanja in okvar na omrežju in kažejo na potrebo po investicijah v omrežje. Izgube so posledice okvar na ceveh, hišnih priključkih in armaturah. Nujna je opredelitev strukture izgub z določitvijo dejanskih vodnih izgub. Te je potrebno analizirati v bilanci vode znotraj obravnavanih območij (občine, bloki in zaključena merilna območja oz. DMA) na osnovi celovitega informacijskega sistema, centralne baze podatkov, programske opreme in strokovnih služb.

»Navidezne izgube« obvladujemo v sklopu sistematične zamenjave vodomeroev na merilnih mestih pri odjemalcih in na merilnih točkah na omrežju s sodobnejšimi in natančnejšimi. Zamenjava poteka za območje celotnega sistema, sočasno z uvedbo daljinskega odčitavanja vodomeroev, vendar je dinamika in realizacija vezana in odvisna od razpoložljivih sredstev. Od leta 2012 je potekala zamenjava vodomeroev vedno intenzivneje. Danes imamo na sistemu že več kot 99 % vodomeroev na daljinsko odčitavanje.

Zniževanje izgub na spojnih vodih (priključnem omrežju) je vezano na program zamenjave vodomeroev. Program zamenjav vodomeroev poteka v ciklusu petih let. Nadzorstvo nad izgubami vode za obvladovanje izgub, ob zamenjavi vodomera predvideva sočasen pregled priključka z ročnim detektorjem izgub. Slednje zahteva letni pregled 8.000 priključkov.

Zniževanje vodnih izgub je mogoče le ob izpolnjevanju trajnostne in razvojne naravnosti izvajanja dejavnosti oskrbe s pitno vodo. Program zniževanja vodnih izgub temelji na sodobni informacijski tehnologiji, sodobnem sistemu nadzora in vodenja proizvodnih procesov in distribucije pitne vode in preizkušanju novih pristopov.

2.1 Dosedanji način iskanja iztokov

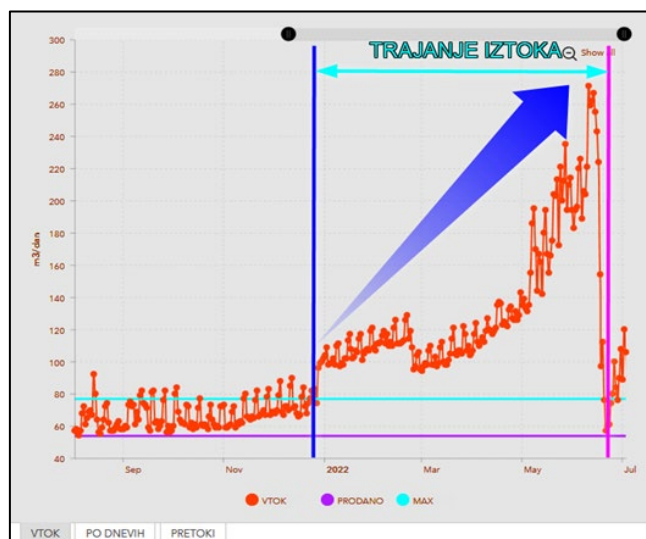
Mariborski vodovod ima za namene stalnega nadzora vodovodnega sistema vzpostavljeno centralno bazo podatkov, v kateri so zbrane vse meritve v sistemu, tako iz sistema daljinskega upravljanja (SDU), kot tudi podatki kontrolnih merilnikov z GSM/GPRS prenosom. Za namene iskanja nepredvidenih dogodkov v omrežju, ki povzročajo vodne izgube, se obdelujejo podatki o količini vode, trenutnih pretokih in tlakih. S celotnega vodovodnega sistema tako v bazi zberemo in obdelamo več kot 36.000.000 posameznih meritev na leto.

id	tag_name	description	type	unit	prikl_sfr	active	area	ID_GIS	Y	X	object_name_GIS	Aquis
1	MARIBOR_HP_HOCE-AM2 F_CV	Tlak tlačnega voda (F_CV)	2	bar	47412	1	HOCE-SLIVNICA	6465220	549415	150647	KV HP HOCE-RAZVANJE	0
2	MARIBOR_CP_VRBANSKI_PLATO-AM100 F_CV	V19 - tlak (F_CV)	2	bar	47418	1	MARIBOR	4585497	548507	158672	VRB PLATO VOD 19	0
3	MARIBOR_CP_VRBANSKI_PLATO-AM98 F_CV	V19 - nivo vode (F_CV)	2	cm	47418	1	MARIBOR	4585497	548507	158672	VRB PLATO VOD 19	0
4	MARIBOR_CP_VRBANSKI_PLATO-V9_S-AM4 F_CV	V9 - Pretok (F_CV)	2	l/s	56575	1	MARIBOR	4569597	548437	158612	VRB PLATO VOD 9	0
5	MARIBOR_CP_VRBANSKI_PLATO-V9_S-AM5 F_CV	V9 - tlak (F_CV)	2	bar	56575	1	MARIBOR	4569597	548437	158612	VRB PLATO VOD 9	0
6	MARIBOR_CP_VRBANSKI_PLATO-AM47 F_CV	V9 - nivo vode (F_CV)	2	cm	56575	1	MARIBOR	4569597	548437	158612	VRB PLATO VOD 9	0
7	MARIBOR_CP_VRBANSKI_PLATO-AM99 F_CV	V19 - Pretok (F_CV)	2	l/s	47418	1	MARIBOR	4585497	548507	158672	VRB PLATO VOD 19	0
8	MARIBOR_PP_RIBINSKO_SELO-AM5 F_CV	Pretok vode (F_CV)	2	l/s	30293	1	MARIBOR	6170297	550040	158380	PP RIBINSKO SELO	0
9	MARIBOR_PP_RIBINSKO_SELO-AM4 F_CV	Tlak tlačni cevovod (F_CV)	2	bar	30293	1	MARIBOR	6170297	550040	158380	PP RIBINSKO SELO	0
10	MARIBOR_HP_HOCE-AM3 F_CV	Pretok vode ho Hoče (F_CV)	2	l/s	47412	1	HOCE-SLIVNICA	6465220	549415	150647	KV HP HOCE-RAZVANJE	0
11	MARIBOR_CP_VRBANSKI_PLATO-AM59 F_CV	V10 - nivo vode (F_CV)	2	cm	47425	1	MARIBOR	965340	548401	158517	VRB PLATO VOD 10	0
12	MARIBOR_CP_VRBANSKI_PLATO-AM60 F_CV	V10 - Pretok (F_CV)	2	l/s	47425	1	MARIBOR	965340	548401	158517	VRB PLATO VOD 10	0
13	MARIBOR_CP_VRBANSKI_PLATO-AM61 F_CV	V10 - tlak (F_CV)	2	bar	47425	1	MARIBOR	965340	548401	158517	VRB PLATO VOD 10	0
14	MARIBOR_CP_VRBANSKI_PLATO-AM40 F_CV	V11 - nivo vode (F_CV)	2	cm	47419	1	MARIBOR	4587097	548479	158710	VRB PLATO VOD 11	0
15	MARIBOR_CP_VRBANSKI_PLATO-V11-AM3 F_CV	V11 - Pretok (F_CV)	2	l/s	47419	1	MARIBOR	4587097	548479	158710	VRB PLATO VOD 11	0
16	MARIBOR_CP_VRBANSKI_PLATO-V11-AM4 F_CV	V11 - tlak (F_CV)	2	bar	47419	1	MARIBOR	4587097	548479	158710	VRB PLATO VOD 11	0
17	MARIBOR_CP_VRBANSKI_PLATO-AM55 F_CV	V12 - nivo vode (F_CV)	2	cm	47426	1	MARIBOR	965339	548420	158563	VRB PLATO VOD 12	0
18	MARIBOR_CP_VRBANSKI_PLATO-AM56 F_CV	V12 - Pretok (F_CV)	2	l/s	47426	1	MARIBOR	965339	548420	158563	VRB PLATO VOD 12	0
19	MARIBOR_CP_VRBANSKI_PLATO-AM57 F_CV	V12 - tlak (F_CV)	2	bar	47426	1	MARIBOR	965339	548420	158563	VRB PLATO VOD 12	0
20	MARIBOR_CP_VRBANSKI_PLATO-AM63 F_CV	V14 - nivo vode (F_CV)	2	cm	47424	1	MARIBOR	965338	548382	158470	VRB PLATO VOD 14	0
21	MARIBOR_CP_VRBANSKI_PLATO-AM64 F_CV	V14 - Pretok (F_CV)	2	l/s	47424	1	MARIBOR	965338	548382	158470	VRB PLATO VOD 14	0
22	MARIBOR_CP_VRBANSKI_PLATO-AM65 F_CV	V14 - tlak (F_CV)	2	bar	47424	1	MARIBOR	965338	548382	158470	VRB PLATO VOD 14	0
23	MARIBOR_CP_VRBANSKI_PLATO-AM71 F_CV	V15 - nivo vode (F_CV)	2	cm	47423	1	MARIBOR	965337	548426	158478	VRB PLATO VOD 15	0

Slika 2: Centralna baza podatkov.

Vir: lasten.

Celotno vodovodno omrežje je dolgo več kot 1.672 km, zato ga je kot celoto nemogoče učinkovito obvladovati. Iz tega razloga omrežje delimo na »zaključena merilna območja« (District Metering Area - DMA), v katerih na dnevni ravni merimo količine vtokov, iztokov in trenutne pretoke, jih preračunamo in primerjamo s povprečno količino prodane vode. Trenutno imamo vzpostavljenih 136 takšnih merilnih območij. Na ta način dobimo seznam območij z največjimi razlikami med dobavljeno in povprečno prodano vodo (m³), na podlagi katerega se prioritarno določajo območja pregledov omrežja. Preračuni so narejeni na dnevni ravni. Podatke vode v DMA tudi grafično vizualiziramo in dnevno spremljamo trende. V primeru sprememb trenda, najprej preverimo morebitna redna vzdrževalna dela in nato po potrebi ukrepamo, saj nam povišani trendi po vsej verjetnosti nakazujejo nepredviden iztok vode, ki se pojavi zaradi lomov cevi, puščanj na spojnih elementih ali priključkih končnih odjemalcev.

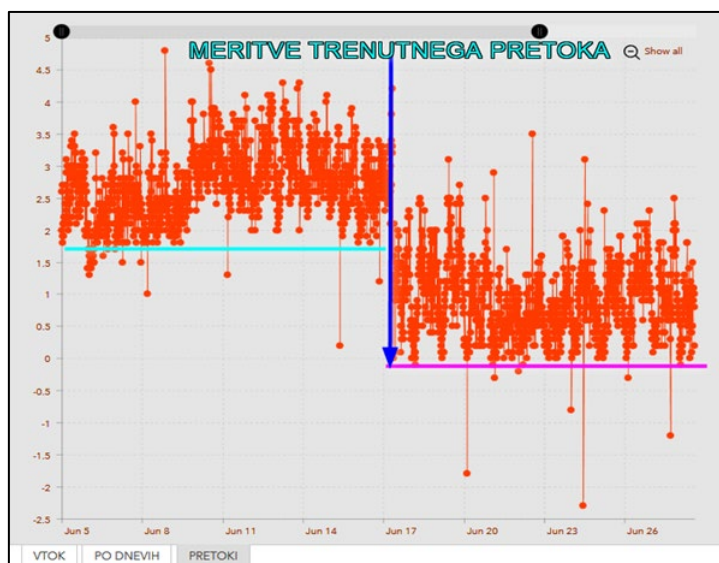


Slika 3: Primer začetka puščanja in sanacije.

Vir: lasten.

Z vedno boljšimi podatki kontrolnih vodomero, smo aplikacijo razvili tudi s preračuni po trenutnih pretokih (l/s), na 15 minutni ravni. Vsi vodomeri nam meritev trenutnih pretokov ne omogočajo, zato je zaključenih merilnih območij, ki jih spremljamo po trenutnem pretoku, nekoliko manj, kot tistih, ki jih spremljamo po dobavljeni količini vode, izraženi v dnevnem volumnu (m³). Pri iskanju vodnih izgub so med najpomembnejšimi kazalci

meritev minimalnega nočnega pretoka (Night Line Measurement – NLM) med 2:00 in 4:00 uro, ko je odjem vode najmanjši. Na ta način precej natančneje ocenimo količino izgubljene vode v določenem zaključnem območju.



Slika 4: Meritev trenutnega pretoka – puščanje in normalizacija po odpravi defekta.

Vir: lasten.

Območjem s povišano količino dobavljene vode sledi raziskava na terenu oziroma »aktivna kontrola iztoka«. Za namen raziskav vodovodnega omrežja na terenu imamo dve ekipi raziskovalcev, ki sta opremljeni s sodobno opremo za natančno lociranje puščanj. Raziskovalci uporabljajo geofon, senzorje šuma in korelatorje, s katerimi je možno natančno locirati prelom cevi oz. puščanje.

Senzorji šuma se zelo učinkovito obnesejo pri iskanju okvar, saj se zaradi večjega števila (komplet senzorjev) lahko pokrije večji del omrežja, ponekod celotno zaključeno merilno območje. Na litoželeznih, jeklenih in pocinkanih ceveh senzorji delujejo zelo dobro tudi na večjih razdaljah, medtem ko na ceveh iz PVC niso tako zanesljivi, prav tako jih je potrebno montirati bližje.



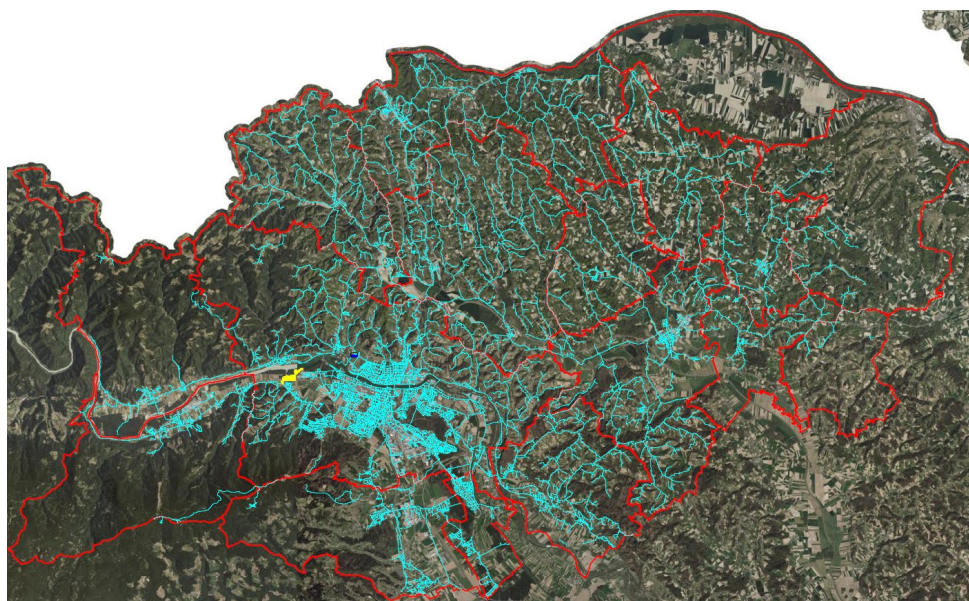
Slika 5: Loggerji šuma se z magnetom pripnejo na vodovodno cev/priključek/ventil.

Vir: lasten.

Senzorji šuma delujejo samostojno, njihova nadgradnja pa so senzorji šuma z možnostjo korelacije oz. korelatorji. Iskanje okvare je z njimi še učinkovitejše, saj omogočajo analizo шумov med posameznimi senzorji, izvedbo korelacije in s tem poenostavijo postopek lociranja puščanja. Po izvedbi korelacije tako v primeru defekta natančneje določimo razdaljo defekta od posameznega korelatorja.

3 Obstoječa infrastruktura

Vodovodni sistem Mariborskega vodovoda je največji enovit sistem za oskrbo z vodo v Sloveniji in letno načrpa približno 14.000.000 m³ vode. Obsega 1.672 km omrežja, 49 vodnjakov, 130 zbiralnikov vode, 134 prečrpalnih postaj, 7.770 hidrantov in 50.061 vodomero. Fizično povezuje in združuje skoraj 200.000 prebivalcev severovzhodne Slovenije. V celoti ali vsaj delno s pitno vodo oskrbuje odjemalce v 16 občinah in sicer iz Mestne občine Maribor, Občine Ruše, Selnica ob Dravi, Hoče-Slivnica, Miklavž na Dravskem polju, Duplek, Pesnica, Šentilj, Kungota, Lenart, Sveta Ana, Benedikt, Sveta Trojica v Slovenskih goricah, Sveti Jurij v Slovenskih goricah, Cerkevjak in Gornjo Radgono.



Slika 6: Vodovodni sistem je dolg 1.672 km in oskrbuje 16 občin na področju SV Slovenije.

Vir: lasten.

Celotna infrastruktura je nacionalnega pomena za dolgoročno ohranjanje zdravja in naravnih virov. Sistem, ki je izpostavljen tako obsežni infrastrukturi mora delovati tako, da ohranja kakovost vode, kot najpomembnejši vir življenja in stalno zmanjšuje stroške vzdrževanja in izgube vode zaradi dotrajanosti ali prelomov. Zraven vodovodne infrastrukture vsebuje »celovit informacijski sistem podjetja«, ki povezuje tako tehnični sistem za podporo upravljanja omrežja, kot tudi poslovni informacijski sistem, namenjen točni evidenci vseh poslovnih dogodkov na omrežju in sistemu, obračunu storitev in podporo vsem drugim procesom v podjetju.

Geografski informacijski sistem povezuje vodovodno omrežje in objekte ter drugo opremo v evidenci celotnega sistema, s tehničnim elektronskim arhivskim sistemom, opremljenim z digitaliziranimi načrti in skicami omrežja.

V procesih črpanja in distribucije vode se uporablja sodobna tehnologija in tehnika, ki se povezuje v sistem daljinskega nadzora in upravljanja vodooskrbnih objektov in omrežja. Tehnologija omogoča spremljanje in optimiziranje pretokov, tlakov in posledično zniževanje vodnih izgub. Sistem je v procesu nenehnega posodabljanja, v smislu vključevanja vedno več objektov in kontrolnih točk omrežja v daljinsko upravljanje.

Meritve z vodovodnih objektov se zbirajo v centralni bazi podatkov, ki zajema podatke sistema daljinskega upravljanja (SDU), kot tudi podatke kontrolnih merilnikov z GSM/GPRS prenosom.

Zbrane podatke dnevno analiziramo in jih vizualiziramo v nadzorni aplikaciji »WaterLoss MBV«, ki povezuje »Centralno bazo meritev« in »geografski informacijski sistem«. Zajema prikaz vseh 136 zaključenih merilnih območij (DMA), s prikazom področja, karakteristik omrežja (število priključkov, število priključenih oseb, dolžino omrežja, povprečen tlak), grafični prikaz z dnevnimi vtoki in primerjavo med dobavljeno in povprečno prodano vodo, na podlagi katerih se določajo prioriteta območja pregledov omrežja. Aplikacija omogoča tudi analizo zaključenih merilnih območij po minimalnih nočnih pretokih območjih z nameščenimi vodomeri, ki to omogočajo.

Odločitve sprejemamo tudi na podlagi trendov dobavljene količine vode – nenadna povišana količina dobavljene vode je razlog za alarm in posledično raziskavo omrežja na terenu oz. lociranje defekta (prelom cevi, puščanje).



Slika 7: Aplikacija za nadzor zaključenih merilnih območij »WaterLoss MBV«.

Vir: lasten.

4 Strojno učenje za iskanje iztokov

Iskanje iztokov iz vodovodnega omrežja bi po laičnem pregledu lahko izvajali ročno in postopoma ob vizualnem pregledu porabe pitne vode. Vsak nenaden dvig porabe pitne vode bi verjetno nakazoval pojav izpusta vode. Pa je res temu tako? Kaj pa, če je razlog za nenaden dvig vode v resnici zunanji pojav, ki odraža dejansko nenadno povečanje porabe pitne vode? Nekaj takih pojavov se redno in periodično dogaja tekom celotnega leta. En tak primer je, da ob prvem skoku temperatur v koncu pomladi ali začetku poletja nekateri domači porabniki, ki predstavljajo posameznike, ki živijo v hišah, začnejo s polnjenjem domačih bazenov iz vodovodnega omrežja pitne vode. Drugi primer je, da ob dolgotrajnem pojavu suše zalivajo domači vrt kar z vodo iz vodovodnega omrežja.

Prav tako je zmotno mišljenje, da pri iztokih vode pride do nenadnega in velikega dviga porabe pitne vode. Nekateri iztoki se začnejo z neznatnim povečanjem porabe, ki ga ni možno zaznati niti ročno, niti z nobeno tehniko strojnega učenja. Se pa tako povečanje porabe akumulira na dolgi rok. Če je iztok stalen in v res majhnih količinah, je tak iztok skoraj nemogoče zaznati, saj lahko predstavlja le spremembo pri porabniku (DMA-ju): mogoče rojstvo novega otroka, novo domačo žival ali novo napravo, ki porabi nekaj vode. Če pa se majhni iztoki postopoma povečujejo (kot je mnogokrat primer v realnosti), pa je tak konstanten dvig možno zaznati.

Seveda taki dvigi porabe niso pokazatelj iztoka iz vodovodnega omrežja. Nepremišljena reakcija interventne ekipe na terenu bi take ekipe le preobremenila. Prav zato je pomembno, da se za identifikacijo izlivov uporabijo naprednejše tehnike, kot pa le gol in laičen pregled porabe pitne vode.

Kot že namigujejo prejšnji odstavki, je identifikacija iztokov močno povezana s kvalitetno napovedjo porabe pitne vode. Če vemo napovedati kakšna bi morala biti poraba v določenem obdobju, lahko pregledamo tudi razliko med predvideno porabo in dejansko porabo. Vzorci razlike med tema vrednostma skozi določeno časovno obdobje pa povedo, če je razlika skladna z iztoki, ali pa gre za kakšen drugi pojav (mogoče se je v prazno hišo končno naselila družina).

4.1 Napovedovanje porabe

Edini podatek, ki ga z zadovoljivo mero zanesljivosti lahko zagotovimo v skorajda celotnem omrežju Mariborskega vodovoda je dejanska poraba porabnikov na določenem DMA-ju v časovnem rezu. V trenutnem sistemu se podatki o porabi v DMA-ju beležijo v petnajst minutnih intervalih. Analiza premajhnih intervalov lahko s seboj prinese preveč šuma: bodisi zaradi napak v merjenju, ali zaradi nenavadnih vzorcev porabe, ki niso povedni za napovedovanje. Posledično je analiza na tako majhnih intervalih neprimerna in je potrebna agregacija porabe v večje časovne intervale. V našem primeru so se enourni intervali izkazali za primerne – niso predolgi, saj so vzorci porabe čez dan razvidni, in niso prekratki, saj ni preveč šuma. [3]

Pri napovedovanju porabe pitne vode imamo izziv, ki mu pravimo univariatna analiza časovnih vrst. Namreč, podatki o pretekli porabi tvorijo podatke v strukturi časovne vrste. Ker pa imamo opravke le z eno vrednostjo, tj. dejansko porabo v časovnem rezu, imamo opravke z univariatno analizo.

Pri napovedovanju gibanja časovnih vrst se uporabljajo številne, dobro poznane tehnike, ki so bile implementirane v tem eksperimentu in katerih opisi sledijo v nadaljevanju.

4.1.1 Drseče povprečje

Metoda **drsečega povprečja** (angl. *moving average* ali *rolling mean*) je klasična metoda analize univariatnih časovnih vrst, kjer se za vsak časovni korak izračuna trend, ki temelji na preteklih podatkih. Ta metoda se mnogokrat uporablja tudi pri glajenju gibanja časovnih vrst in prav to glajenje laično obravnavamo kot trend gibanja časovne vrste. [4]

Izračun drsečega povprečja se naredi s pomočjo konvolucijskega filtra, ki za določeno točko v času pregleda k prejšnjih časovnih korakov in za vrednosti v teh k korakih izračuna povprečje. Formalna definicija izračuna vrednosti drsečega povprečja za najnovejši časovni korak je sledeča.

$$DP_n = \frac{v_{n-k+1} + v_{n-k+2} + \dots + v_n}{k} = \frac{1}{k} \sum_{i=n-k+1}^n p_i$$

Kjer v_n predstavlja vrednost v najnovejšem časovnem koraku, v_{n-k+1} predstavlja vrednost za k korakov nazaj, vmes pa so vse vrednosti med tema korakoma. Vrednost k predstavlja velikost okna konvolucijskega filtra. Večje ko je okno filtra, bolj gladka bo krivulja, ampak hkrati se lahko manjši vzorci gibanja spregledajo. Manjše ko je okno filtra, manj gladka je krivulja, ampak bolj je izračun dovzeten za šume, ki predstavljajo lažne vzorce.

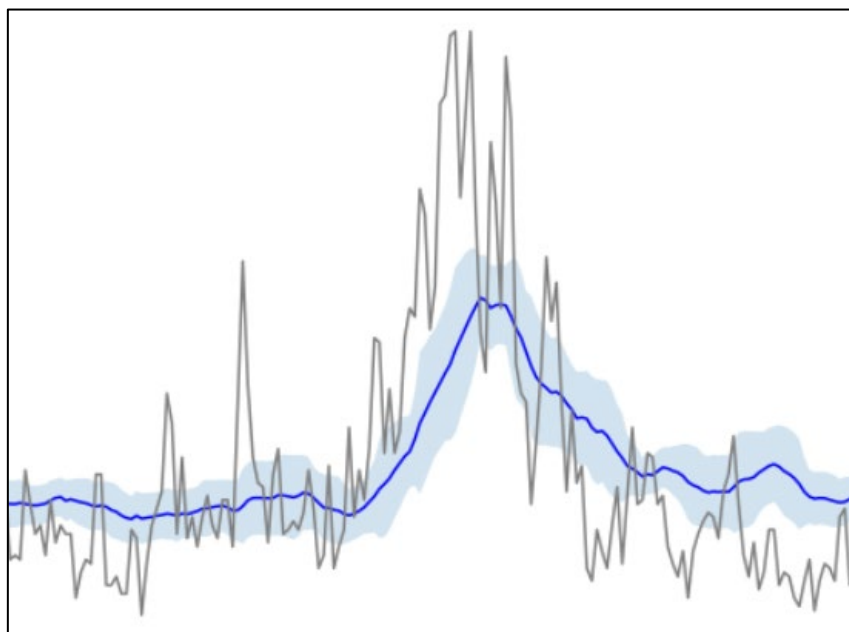
Iz preprostega drsečega povprečja lahko izpeljemo številne različice. Ena izmed pogostih je vsekakor **uteženo drseče povprečje**, kjer novejši podatki prispevajo večji delež k izračunu povprečja, kot pa starejši. To uteževanje je lahko linearno ali celo eksponentno (v tem primeru govorimo o eksponentnem uteženem drsečem povprečju).

Poleg drsečega povprečja lahko izračunamo tudi druge drseče opisno statistične vrednosti. Zelo prav nam pride tudi **drseč standardni odklon** (ali druga mera variabilnosti), ki nam kaže kako velika volatilitnost je bila v oknu konvolucijskega filtra. Slika 14 prikazuje drseče povprečje in drseč standardni odklon na podatkih porabe pitne vode. [4]

4.1.2 Dekompozicija časovnih vrst

Ko pa imamo za en DMA izračunan splošni vzorec porabe (bodisi z drsečim povprečjem ali kakšno drugo metodo), pa lahko nadaljujemo z postopkom, ki mu pravimo dekompozicija časovnih vrst. Od dejanske porabe pitne vode lahko odštejemo trend in ostane nam poraba, ki se ne sklada s splošnim vzorcem. Pravimo, da smo časovno vrsto razdelili na dva sestavna dela oz. na dve kompoziciji. Prav zato postopku pravimo dekompozicija.

Ni nujno, da se ustavimo tukaj, lahko z dekompozicijo nadaljujemo: ostanek ponovno s konvolucijo zgladimo, da dobimo vzorec porabe, ki predstavlja razliko od splošnega vzorca. Ta korak smo naredili le zaradi glajenja, ne za iskanje novega vzorca. Tokrat pa razlika med zglajenim ostankom in dejansko porabo predstavlja izredne nenavadnosti oz. celo nenadne iztoke pitne vode. [5]



Slika 8: Drseče povprečje in drseč standardni odklon.

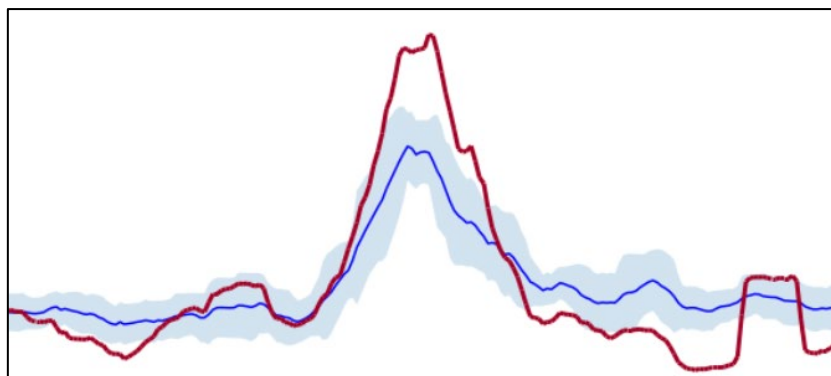
Vir: lasten.

V primeru, da imamo podatke preko več časovnih period (leto, mesec, teden...), lahko prvo komponento splošnega vzorca izračunamo celo le z enim DMA mestom. Zaradi spreminjajoče se narave porabe vode posameznih DMA-jev in nenehno spreminjajočega se vodovodnega omrežja, analiza podatkov čez daljše obdobje (čez obdobje dveh let) ni smiselno, saj pridemo do pojava premika podatkov (angl *data shift*) (na katerega vplivajo bodisi premiki kovariatov ali premiki koncepta), ko začnejo že naučeni modeli identifikacije iztoka delovati slabše.

4.1.3 Splošni vzorci porabe pitne vode

Taka drseča povprečja in standardne odklone lahko za več časovnih vrst tudi agregiramo – najlažje kar s povprečenjem preko dimenzije časa. Ta trik smo uporabili tudi pri implementaciji našega sistema, da smo dobili splošen vzorec porabe celotnega območja Mariborskega vodovoda.

Prav primerjava splošnega vzorca porabe pitne vode in vzorca porabe pitne vode enega DMA-ja pa predstavlja pomembno informacijo o odstopanju porabe posameznega DMA-ja napram pričakovani porabi. Prav razlika med vzorcem splošne porabe in porabo enega DMA-ja pa predstavlja pomemben indikator, ki ga lahko interpretiramo kot odstopanje, ki zna nakazovati na nevarnost iztoka pitne vode na posameznem DMA-ju. Razlika med splošno porabo in porabo posameznega DMA-ja je na spodnji sliki 15 prikazana v obliki obarvane črte vzorca porabe.



Slika 15: Splošen trend porabe pitne vode celotnega območja Mariborskega vodovoda (modro), ter trend porabe pitne vode enega DMA-ja (črno).

Vir: lasten.

Bolj ko je linija temna, večje odstopanje porabe pitne vode za izbran DMA opazamo in s tem je tudi večja nevarnost za potencialni iztok. Velja tudi obratno, bolj ko je svetla črta, bolj je poraba pitne vode izbranega DMA-ja skladna s porabo celotnega območja Mariborskega vodovoda in manjša je nevarnost za iztok na izbranem DMA-ju.

Seveda, neprimerno je interpretirati vsak manjšo obarvanje črte trenda posameznega DMA-ja. Čeprav vse uporabljene metode agregacije oz. izračuna trenda do določene mere odstranijo šume, so ti še vedno lahko prisotni. Namreč, prevelika robustnost na šume (odstranjevanje vseh nihanj) bi lahko izbrisala dejanske vzorce. Prav zato pa se interpretirajo le odtenki črt na nekoliko daljše časovno obdobje, ne le nekaj dnevna potencialna nevarnost za iztok. Po prvih preizkusih je interpretacija stalne nevarnosti v zadnjih petih dneh že dober pokazatelj nevarnosti za iztok, kar je vsekakor krajše časovno obdobje, kot pa mora preteči od iztoka po klasičnem ročnem sistemu iskanja iztokov.

4.2 Gručenje podobnih porabnikov

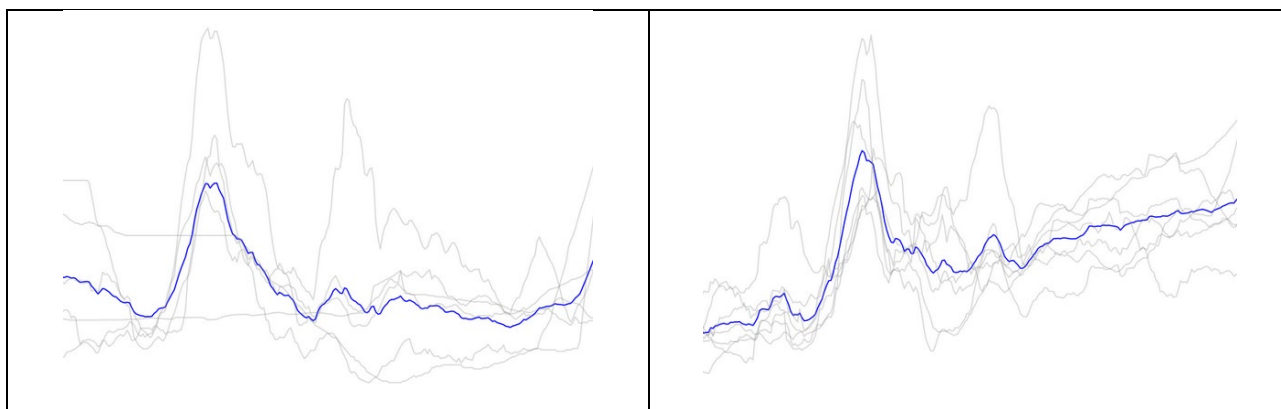
Vsekakor pa primerjava porabe vsakega DMA-ja s porabo na celotnem območju Mariborskega Vodovoda ni najbolj pravilna. Namreč, vsa območja si niso enakovredna po porabi. Nekatera območja (npr. center) so preplavljena z mešanimi porabniki – tako domačimi porabniki v hišah, kot stanovanjih, med katere se vmešajo tudi storitvene dejavnosti. To je vsekakor v nasprotju z industrijskimi območji (npr. Melje) ali območji, kjer prevladujejo porabniki, ki stanujejo le v hišah. Ker Mariborski Vodovod pokriva nekoliko širše območje, kot le mestno občino, se nekateri porabniki ukvarjajo tudi s kmetijskimi dejavnostmi. Vsak tak tip porabnika pitne vode ima nekoliko sebi specifične vzorce porabe. V industrijskih conah v času kolektivnih dopustov in v splošnem izven delovnega časa beležijo manjšo porabo pitne vode. Porabniki, ki živijo v družinskih hišah in imajo nekaj zelenice si ob prvem pojavu vročih dni največkrat polnijo domače bazene kar s pitno vodo. Prav tako pa območja s pretežno domačimi porabniki doživljajo bodisi dvige čez vikend (če družine preživijo te doma), čez popoldan (po službi), ter spuste čez praznike ali celo vikende (če te preživijo drugje kot doma) in ob nedoločenem času tudi čez poletje (ko gredo na dopust). Vsi ti dvigi in spusti porabi pitne vode so del vsakdanjega življenja in se ne smejo zamenjevati za nenavadnosti ali celo za pokazatelj iztoka pitne vode. [6]

Prav zaradi tega pa je tudi nesmiselno primerjati porabo posameznega DMA-ja s prav celotnim območjem Mariborskega vodovoda. Da res ugotovimo, če so nenadni dvigi ali spusti porabe nekaj pričakovanega ali pa gre za nekaj nenavadnega, bi bilo prav, da posamezne DMA-je primerjamo z območji, ki so v splošni porabi temu DMA-ju podobni – pri tem pa ignoriramo tiste, ki so v splošni porabi pitne vode popolnoma drugačni.

V namen iskanja podobnih pa se lahko uporabi tehnika strojnega učenja, ki se imenuje gručenje. Z gručenjem se iz podatkov najde vnaprej definirano število gruč DMA-jev, za katere velja sledeče:

- Tisti DMA-ji, ki so v isti gruči, imajo podobne profile porabe pitne vode, ter

- tisti DMA-ji, ki so v drugih gručah, se v profilu porabe pitne vode razlikujejo.



Slika 16: Dve različni gruči DMA-jev glede na trende porabe pitne vode.

Vir: lasten.

Gručenje podobnih časovnih vrst lahko poteka na več načinov. Za namen našega sistema sta bila implementirana dva. Prvi pristop je **k-mean gručenje za časovne vrste**, kjer se med posameznimi časovnimi koraki izračuna absolutna razdalja med porabami. Seštevek teh razdalj prikazuje razliko med porabama dveh DMA-jev. Cilj je, da se v isto gručo podajo taki DMA-ji, a je ta seštevek razdalj med njimi manjši napram seštevkom razdalja do DMA-jev v drugih gručah. Slika 16 prikazuje primer rezultata takega gručenja.

Drugi pristop, ki smo ga implementirali za namen gručenja, pa je gručenje s pomočjo **evolucijskega algoritma** s pomočjo NiaPy knjižnice [2]. Gručenje merilnih območij izvedemo na način, da v prvem koraku posamezna merilna območja naključno razporedimo v n gruč. V naslednjem koraku izračunamo povprečen standardni odklon porabe vode po dnevih za merilna območja znotraj posamezne gruče ter vrednosti standardnih odklonov združimo z uporabo pristopa utežene vsote. Utežene vsote vseh gruč seštejemo in to vrednost v nadaljevanju uporabimo kot cenitveno funkcijo (angl. *fitness function*) pri uporabi algoritma jDE [7], temelječega na diferencialni evoluciji s katerim skušamo poiskati takšno razporeditev merilnih območij v n gruč, da bo vrednost cenitvene funkcije čim manjša. Algoritem izvede 5000 iteracij, najboljšo najdeno razdelitev merilnih območij, to je rešitev, ki je dosegla najnižjo vrednost cenitvene funkcije pa uporabimo kot končno rešitev.

5 Zaključek

Identifikacija iztokov pitne vode je vedno bolj pereč problem ob vse pogostejših pojavih hudih sušnih obdobjih, predvsem v poletnih mesecih. Zanesljivost identifikacije pri ročnem pregledu podatkov ali čakanju na namige o iztoku iz okolja, prinaša preveč nepredvidljivosti in nepovratne škode na okolje, družbo in finance lokalnih skupnosti. Prav zato je predlagan sistem za samodejno identifikacijo s pomočjo algoritmov strojnega učenja zelo primeren za uporabo ob obstoječih sistemih nadzora vodovodnih sistemov.

S prispevkom smo prikazali kako lahko prav zastavljen sistem učinkovito pomaga kontrolorjem pri identifikaciji iztokov in s tem na daljši rok ohrani vire pitne vode kar se da učinkovito porabljene. Vsekakor pa tak sistem ne nadomešča človeške odločitve za posredovanje na terenu – je le pomoč, ki v obliki »človek nad strojem« (angl. *human over the loop*) pomaga pri odločitvah. Za še učinkovitejše in dobro integrirane sisteme pa so seveda potrebne še nadaljnje raziskave – tako algoritmov strojnega učenja za zanesljivejšo in hitrejšo identifikacijo iztokov, kakor za izkoriščanje še več podatkov, ki bi lahko tako identifikacijo naredilo še bolj zanesljivo.

Literatura

- [1] ADEDEJI Kazeem B., YSKANDAR Hamam, BOLANLE Tolulope Abe, in ADNAN M. Abu-Mahfouz »Towards achieving a reliable leakage detection and localization algorithm for application in water piping networks: An overview.« v IEEE Access, peta številka 2017, str. 20272-20285.
- [2] VRBANČIČ Grega, BREZOČNIK Lucija, MLAKAR Uroš, FISTER Dušan, in FISTER Iztok »NiaPy: Python microframework for building nature-inspired algorithms«, b Journal of Open Source Software številke 3.23, 2018, str. 613.
- [3] LIM Bryan in STEFAN Zohren »Time-series forecasting with deep learning: a survey.« v Philosophical Transactions of the Royal Society, številke 2194 v 2021.
- [4] HANSUN Seng »A new approach of moving average method in time series analysis.«, v zborniku konference 2013 conference on new media studies (CoNMedia), 2013, str. 1-4.
- [5] DAGUM Estela Bee »Time series modeling and decomposition.« v Statistica, številka 70 v 2010, str. 433-457.
- [6] AGHABOZORGI Saeed, ALI Seyed Shirshorshidi, in TEH Ying Wah »Time-series clustering—a decade review.« v Information systems, številka 53 v 2015, str. 16-38.
- [7] BREST Janez, Greiner Sašo, BOSKOVIC Borko, MERNIK Marjan in ZUMER Viljem »Self-adapting control parameters in differential evolution: A comparative study on numerical benchmark problems«, v IEEE Transactions on evolutionary computation št. 6, leta 2006, str. 646-657.