

# Uporaba metod strojnega učenja za modeliranje deleža prestreženih padavin na eksperimentalnem porečju reke Dragonje

Luka Štravs, Mojca Šraj, Mitja Brilly<sup>1</sup>

## Povzetek

V okviru raziskave so bile na podlagi izvedenih meritev in analize rezultatov le-teh z metodami strojnega učenja ugotovljene značilnosti vpliva zaraščanja opuščeni kmetijskih površin na vodno bilanco porečja Dragonje. Gozdovi na na novo zaraščeni površinah prestrežejo skoraj vso vodo padavinskih dogodkov z manj kot 2,5 mm padavin, kar pomeni, da v takih primerih ne pride do površinskega in podpovršinskega odtoka v Dragonjo ali napajanja zalog podzemne vode v porečju. Za dogodke z več kot 2,5 mm padavin pa generirani modeli kažejo, da novi gozd prestreže od 42,8 % padle vode pri dogodkih s količino padavin od 2,5 do 7,5 mm in 23,2 % padle vode pri dogodkih s količino padavin večjo od 7,5 mm, ko pa imata na vrednost deleža prestreženih padavin močan vpliv tudi povprečna hitrost vetra in intenziteta padavin. Lahko torej ugotovimo, da metode strojnega učenja za generiranje modelov v obliki odločitvenih dreves predstavljajo zelo uporaben in obetajoč pripomoček pri iskanju novih znanj v podatkovnih bazah.

## Uvod

Hidrološka znanost se ukvarja s kroženjem vode, njeno pojavnostjo ter distribucijo, gibanjem in njenimi fizikalno-kemijskimi karakteristikami (Chow, 1964). Glavna raziskava v okviru hidrologije je na področju kroženja vode med atmosfero, zemeljskim površjem in vodnimi sistemi v obliki površinskih ali podzemnih voda (Brilly and Šraj, 2000). Gozdna hidrologija se pretežno ukvarja s kroženjem vode na območjih gozdov; proučuje spreminjanje pojavnosti in poti vode iz atmosfere preko gozdnega ekosistema v tla, podzemno in površinsko vodo ter njen povratek nazaj v atmosfero.

Padavine so glavni vir vode v gozdnem hidrološkem krogu (slika 1). Večinoma predstavljata padavine dež ali sneg, v obmorskih in goratih območjih pa tudi megla. Velik del padavin, ki padejo nad gozdom, prestrežejo drevesne krošnje, manjši del pa jih pade skozi odprtine med krošnjami in listi naravnost na tla. Znatna količina padavin, ki jih prestrežejo krošnje, izhlapi nazaj v ozračje med ali takoj po nalivu ( $E_i$ ). Del prestreženih padavin pade s krošenj oz. listov na tla kot kapljanje, potem ko je kapaciteta krošnje že zapolnjena. Manjši del prestreženih padavin pa se steka z listov na veje in z vej po deblu do tal ( $S_f$ ). Ker padavin, ki padejo skozi odprtine med krošnjami in listi, ter padavin, ki kasneje prikaplajo do tal, ne moremo ločeno meriti, jih obravnavamo skupaj ( $T_f$ ). Vsota padavin, ki padejo skozi odprtine med krošnjami in listi naravnost na tla, in padavin, ki kasneje prikaplajo do tal ( $T_f$ ), ter odtoka po deblu ( $S_f$ ) je navadno bistveno manjša od količine padlih padavin. Bilanca je torej naslednja:

$$P = T_f + S_f + E_i \quad \rightarrow \quad E_i = P - (T_f + S_f)$$

Če intenziteta vseh padavin, ki dosežejo tla ( $T_f + S_f$ ), preseže infiltracijsko kapaciteto zemljine, višek odteče kot površinski odtok. Zaradi zelo velike infiltracijske kapacitete

---

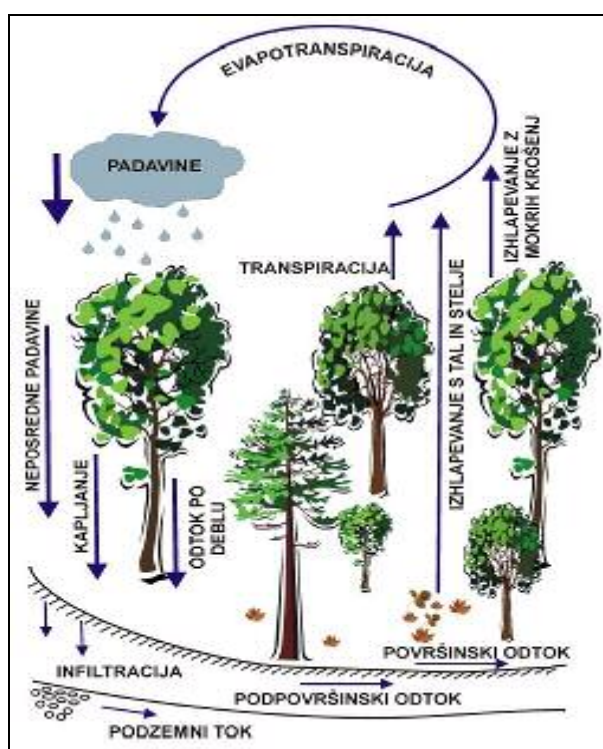
<sup>1</sup> Fakulteta za gradbeništvo in geodezijo, Jamova 2, SI-1000 Ljubljana.

organske zemljine v večini gozdov je ta vrsta odtoka redka. Infiltrirana voda v tleh predstavlja zaloge podtalnice, ki bogati tudi vodotoke.

Na količino oziroma delež prestreženih padavin (Mikoš et al., 2002) na gozdnih območjih vplivata predvsem vegetacija in klimatski faktorji:

- kapaciteta krošnje, ki je odvisna od vrste, velikosti, oblike in starosti dreves v gozdu, površine listja in njihove orientacije;
- gostota vegetacije;
- vrsta, intenziteta, trajanje in frekvenca padavin;
- klimatske razmere (temperatura, veter, vlažnost)
- in drugi.

Na podlagi raziskav je Ovington (1954) ugotovil, da delež prestreženih padavin varira od 6 do 93 % v odvisnosti od razmer oziroma pogojev na posameznih območjih.



Slika 1 – Komponente hidrološkega kroga v gozdu (Šraj, 2003a).

V sodelovanju z Vrije Universiteit iz Amsterdama so bile na porečju Dragonje izvedene bazične raziskave hidroloških procesov. Na območju porečja reke Dragonje je bilo v zadnjih desetletjih opaženo intenzivno zaraščanje opuščenih kmetijskih površin. V tem obdobju so se znižale vrednosti nizkih in visokih pretokov, ni pa bilo opaznih sprememb v padavinskem in temperaturnem režimu. Glavni namen te raziskave je bil ugotoviti vpliv zaraščanja opuščenih zemljišč z gozdom in drugo zarastjo na vodno bilanco na porečju Dragonje in ugotoviti vpliv posameznih klimatskih faktorjev na procese kroženja vode.

Za analizo rezultatov meritev v okviru bazičnih hidroloških raziskav smo uporabili metodi strojnega učenja M5 in J4.8, s pomočjo katerih smo generirali modele v obliki odločitvenih dreves. Metode strojnega učenja so bile že večkrat uspešno uporabljene na področju hidrologije (Štravs et al., 2004; Solomatine & Dulal, 2003).

### Karakteristike porečja

Porečje reke Dragonje s prispevnim območjem 90,5 km<sup>2</sup> se nahaja na jugozahodnem delu Slovenije oziroma na severnem delu istrskega polotoka. Povprečna letna količina padavin je na obalnem delu porečja 900 mm, v bolj celinskem delu pa letno pade 1200 mm padavin. Za razliko od pravega mediteranskega padavinskega režima ni izrazite sezone razporeditve padavin s sušnimi poletnimi meseci in mokrimi zimskimi, temveč so vsi meseci približno enakomerno namočeni (Šraj, 2003b). Povprečna letna temperatura na obalnem območju porečja je 14 °C, v notranjosti pa 10 °C.

Ekonomske razmere v šestdesetih, še bolj pa v sedemdesetih letih prejšnjega stoletja so povzročile demografske spremembe, kar se je odrazilo v spremembi rabe pašnikov in travnikov, kasneje pa tudi obdelovalnih površin. Hkrati z opuščanjem ekstenzivne kmetijske rabe in depopulacijskimi procesi so se izvajala tudi protierozijska vegetacijsko stabilizacijska dela. Vse to je pospešilo zaraščanje območja in v zadnjih treh desetletjih se je zaraščenost porečja Dragonje povečala s povprečno 25 % na več kot 60 % (Globevnik, 2001).



Slika 2 – Porečje Dragonje (Šraj, 2003b).

Slovenska obala je dobro znana tudi kot vododeficitarno območje, na katerem se še posebej v času turistične sezone čuti pomanjkanje razpoložljivih kapacitet vodnih virov za oskrbo s pitno vodo.

### Metode

#### Meritve

Za izvedbo meritev posameznih komponent gozdnega hidrološkega kroga sta bili izbrani dve gozdni raziskovalni ploskvi, ki sta medsebojno oddaljeni 400 m in obe na približno 200 m nadmorske višine, v 30- do 35-letnem gozdu nad sotočjem Dragonje in Rokave. Prva raziskovalna ploskev s površino 1420 m<sup>2</sup> se nahaja na severno orientiranem pobočju v porečju Rokave, druga raziskovalna ploskev s površino 615 m<sup>2</sup> pa na južno orientiranem pobočju v porečju Dragonje.

Na obeh raziskovalnih ploskvah so bile za vsak padavinski dogodek merjene padavine nad krošnjami dreves, količina padavin, ki padejo direktno ali pa prikapljuje skozi krošnje dreves, in količina odtoka po deblu. Padavine nad krošnjami so bile merjene z ombrografom in totalizatorji z ročnim praznjenjem (slika 3). Količina padavin, ki padejo direktno ali pa prikapljuje skozi krošnje dreves, je bila merjena s kovinskimi žlebovi, opremljenimi z avtomatskimi zapisovalci podatkov in premičnimi totalizatorji, katerih lokacija se je po vsakem praznjenju spremenila. Odtok po deblu je bil merjen na gabru ter hrastu na severni ploskvi in jesenu ter hrastu na južni ploskvi; te vrste dreves predstavljajo najbolj pogoste vrste dreves na posameznih raziskovalnih ploskvah.



Slika 3 – Meritve količine padavin nad krošnjami (levo) in meritev količine padavin, ki padejo skozi krošnje dreves (desno).

Vse spremenljivke so bile merjene v 10-minutnih intervalih. Temperatura zraka, relativna vlažnost ter smer in jakost vetra so bili merjeni na bližnjih meteoroloških postajah Kubed in Boršt. Več o meritvah posameznih procesov je možno prebrati v Šraj (2003b).

### Modeliranje

Glavna ideja pri generiranju modelov v obliki dreves (angl. tree-like models) je razvoj enostavnih in strukturno transparentnih (za razliko od npr. nevronske mreže) modelov, ki so enostavni tako z vidika uporabe kot tudi razlage. Za razvoj modelov, ki predstavljajo empirične povezave med merjenimi spremenljivkami in deležem prestreženih padavin, smo v naši raziskavi uporabili metodi strojnega učenja M5 in J4.8, kot sta implementirani v programskem orodju WEKA, ki je bilo razvito na Univerzi Waikato na Novi Zelandiji (Witten & Frank, 2000).

S tem, ko metodi strojnega učenja ponudimo dovolj vhodnih (atributov) in izhodnih podatkov, ki opisujejo modelirani proces, se ta iz posameznih primerov nauči vzorcev obnašanja v okviru modeliranega procesa. Pri tem se celotni podatkovni prostor vhodnih spremenljivk deli na podprostore, katerih meje so v procesu razvoja modela optimizirane, znotraj posameznih podprostorov celotnega podatkovnega n-dimenzijskega prostora pa veljajo določeni vzorci, ki se jih lahko aproksimira z npr. nominalno ali konstantno numerično vrednostjo in linearno funkcijo.

Odločitvena drevesa se generirajo na podlagi iterativnega deljenja podatkov v podatkovne podprostore celotnega podatkovnega prostora, pri čemer je cilj, da je razdalja

med posameznimi podatkovnimi podprostorimi čim večja (Breiman et al., 1984; Quinlan 1986, 1992; Kompare, 1995; Mitchell, 1997; Witten & Frank, 2000; Solomatine & Dulal, 2003). Glavne komponente odločitvenih dreves so odločitvena vozlišča, veje in listi. Odločitveni proces se začne pri glavnem oziroma zgornjem odločitvenem vozlišču (na sliki 5 je to atribut 'količina padavin', ki loči drevo v dve veji glede na to, ali je vrednost atributa 'količina padavin' večja ali manjša od 2,4), v katerem je specificiran razdelitveni kriterij. Glede na odgovor na ta razdelitveni kriterij se odločitveno drevo razdeli v veje. Vsaka veja lahko vodi do podrejenih vej ali pa zaključnih delov odločitvenih dreves, ki se imenujejo listi in predstavljajo vrednost, ki jo zavzame model.

Rezultati modeliranja so torej modeli v obliki odločitvenih dreves, ki predstavljajo zaporedje pravil, ki pripeljejo do vrednosti razreda, numerične vrednosti ali linearne funkcije uporabljenih atributnih spremenljivk. Glede na to lahko razdelimo vrste odločitvenih dreves (angl. decision trees) v:

- klasifikacijska drevesa (angl. classification trees) z nominalnimi vrednostmi razreda kot listi modela,
- regresijska drevesa (angl. regression trees) s konstantnimi numeričnimi vrednostmi kot listi modela in
- modelna drevesa (angl. model trees) z linearnimi kombinacijami oziroma funkcijami uporabljenih atributov kot listi modela.

```

dQ <= 2.78 :
|   Qstart <= 23 :
|   |   dQ <= 0.783 : k = -0.0259
|   |   dQ > 0.783 : k = -0.0372
|   Qstart > 23 : k = -0.0510
dQ > 2.78 :
|   dQ <= 6.84 : k = -0.0617
|   dQ > 6.84 : k = -0.0938

```

Slika 4 – Primer regresijskega drevesa.

Primer regresijskega drevesa, katerega atributa pri razvoju modela sta bila  $dQ$  in  $Q_{START}$ , modelirana spremenljivka pa  $k$ , je na sliki 4 in se tolmači oziroma uporablja na naslednji način:

- če je  $dQ$  nižji od 0,783,  $Q_{START}$  pa nižji od 23, potem se uporabi vrednost  $k = -0,0259$ ;
- če je  $dQ$  višji od 0,783 in nižji od 2,78,  $Q_{START}$  pa nižji od 23, potem se uporabi vrednost  $k = -0,0372$ ;
- če je  $dQ$  nižji od 2,78,  $Q_{START}$  pa višji od 23, potem se uporabi vrednost  $k = -0,0510$ ;
- če je  $dQ$  višji od 2,78 in nižji od 6,84, potem se uporabi vrednost  $k = -0,0617$ ;
- če je  $dQ$  višji od 6,84, potem se uporabi vrednost  $k = -0,0938$ .

#### Podatki

V obdobju enega leta je bilo zabeleženih 369 dogodkov; 173 na južni in 196 na severni raziskovalni ploskvi. Ločnica med dvema zaporednima dogodkoma je bilo sušno obdobje, v katerem so se krošnje dreves lahko posušile. Za vsak posamezen dogodek so bili tako zbrani naslednji atributni podatki:

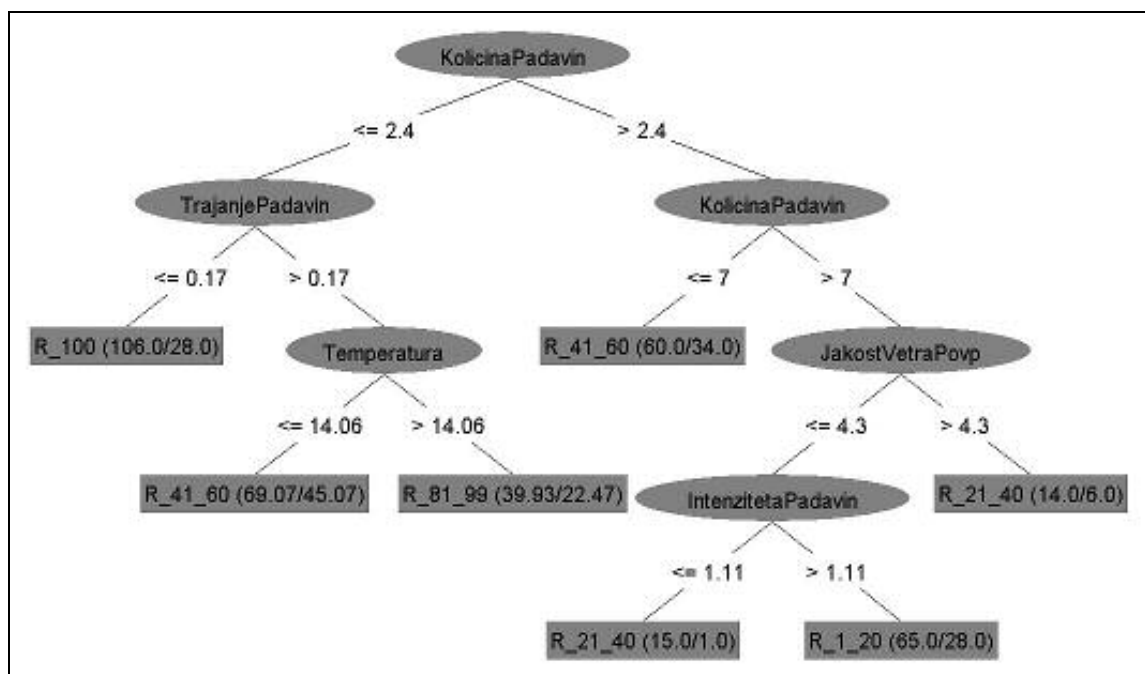
- orientacija raziskovalne ploskve (severna, južna),
- količina padavin (v mm),
- trajanje naliva (v urah),
- intenziteta naliva (v mm na uro),
- povprečna temperatura zraka (v °C),
- relativna vlažnost (v %) in
- povprečna hitrost vetra (v m/s).

Količina padavin za posamezne dogodke je bila v razponu od 0,2 do 100,2 mm, nalivi so trajali od 5 minut do skoraj 40 ur, intenziteta padavin pa je varirala od 0,15 do celo 44 mm na uro.

### Rezultati

S pomočjo metod strojnega učenja M5 in J4.8 smo razvili tri modele v obliki odločitvenih dreves.

V primeru #1 (slika 5) smo razvili klasifikacijsko drevo, katerega atributi so bili orientacija raziskovalne ploskve, količina, trajanje in intenziteta padavin, temperatura in relativna vlažnost zraka in povprečna hitrost vetra v času dogodka. Modelirana spremenljivka je bil odstotni delež prestreženih padavin (v odnosu do celotne količine padlih padavin nad krošnjami), ki smo ga za potrebe tega primera klasificirali v 7 razredov, in sicer: razred R\_0 (slika 5) predstavlja dogodke, ko je bil relativni delež prestreženih padavin 0 %, R\_1\_20 predstavlja dogodke z relativnim deležem prestreženih padavin od 1 do 20 % in tako naprej do razreda R\_100, ko je bilo v okviru posameznega dogodka prestreženo 100 % padavin.



Slika 5 – Razviti model v obliki klasifikacijskega drevesa za primer #1 (metoda strojnega učenja J4.8).

Iz modela (slika 5), ki pravilno klasificira 56 % znanih in 46 % neznanih primerov, lahko razberemo, da je v primeru dogodkov z manj kot 2,4 mm padlih padavin in

trajanjem, krajšim od 10 minut, prestrežena vsa padla voda (razred R\_100). V takih pogojih padavine ne dosežejo gozdnih tal in torej ne pride do procesov napajanja podzemnih rezervoarjev preko infiltracije in ne pride do procesov površinskega ali podpovršinskega odtoka. Na drugi strani pa je pri padavinskih dogodkih z enako količino padavin (do 2,4 mm) in trajanjem, daljšim od 10 minut, prestreženo med 41 in 60 % padavin, če je temperatura nižja od 14 °C, in prestreženo med 81 in 99 % padavin, če je temperatura višja od 14 °C. Razviti model torej za dogodke z manj kot 2,4 mm padavin kot pomembna faktorja, ki vplivata na delež prestreženih padavin, opredeli trajanje naliva in povprečno temperaturo v času dogodka.

Pri dogodkih z več kot 2,4 in manj kot 7,0 mm padavin je prestreženo med 41 in 60 % (razred R\_41\_60) padavin, torej približno polovica vseh padavin. Pri dogodkih z več kot 7 mm padavin se kot faktorja, ki pomembno vplivata na proces, pojavita tudi povprečna hitrost vetra in intenziteta padavin; prestreženih je do 40 % padavin v odvisnosti od vrednosti obeh spremenljivk pri posameznem dogodku. Zanimivo je, da struktura razvitega modela navkljub pričakovanjem ne kaže vpliva orientacije pobočja (severno/južno) na proces. To bi lahko bila tudi posledica tega, da so razlike v procesih na severnem in južnem pobočju predstavljene v obliki različnih vrednosti posameznih atributov dogodkov; torej v povprečni količini padavin, temperaturi zraka, relativni vlažnosti in vetrovni sliki.

V primerih #2 (slika 6) in #3 (slika 7) smo razvili dve regresijski drevesi, katerih atributi so bili orientacija raziskovalne ploskve, količina, trajanje in intenziteta padavin, temperatura in relativna vlažnost zraka in povprečna hitrost vetra v času dogodka. Modelirana spremenljivka je bil odstotni delež prestreženih padavin, ki za razliko od primera #1 ni bil klasificiran v razrede, temveč izražen v numerični obliki v intervalu med 0 in 100 %. Glavna razlika med obema generiranima modeloma je bila kompleksnost le-teh; rezultat v obliki modela v primeru #3 predstavlja obrezano (angl. 'pruned') oziroma manj kompleksno drevo. Manjša oziroma skoraj zanemarljiva je tudi razlika v obnašanju obeh modelov pri verifikaciji na neznanih primerih; srednja absolutna napaka razvitega modela v primeru #2 je 12,58 %, srednja absolutna napaka razvitega modela v primeru #3, ki predstavlja manj kompleksno obliko modela v primeru #2, pa je 14,80 % padavin.

```

KolicinaPadavin <= 2.5 :
|   Temperatura <= 14.2 :
|   |   TrajanjePadavin <= 1.67 : ProcentEi = 81.2
|   |   TrajanjePadavin > 1.67 : ProcentEi = 47.2
|   Temperatura > 14.2 :
|   |   KolicinaPadavin <= 0.5 : ProcentEi = 98
|   |   KolicinaPadavin > 0.5 : ProcentEi = 85.7
KolicinaPadavin > 2.5 :
|   KolicinaPadavin <= 7.5 : ProcentEi = 42.8
|   KolicinaPadavin > 7.5 :
|   |   JakostVetraPovp <= 3.45 :
|   |   |   IntenzitetaPadavin <= 1.13 : ProcentEi = 30.1
|   |   |   IntenzitetaPadavin > 1.13 : ProcentEi = 17
|   |   JakostVetraPovp > 3.45 :
|   |   |   TrajanjePadavin <= 6.5 : ProcentEi = 15.6
|   |   |   TrajanjePadavin > 6.5 : ProcentEi = 40

```

Slika 6 – Razviti model v obliki regresijskega drevesa za primer #2 (metoda strojnega učenja M5).

Iz obeh razvitih modelov, še posebej pa iz manj kompleksnega regresijskega drevesa v primeru #3 (slika 7), je razvidno, da je pri dogodkih z manj kot 2,5 mm padavin in temperaturo zraka nižjo od 14,2 °C prestreženih 81,2 % padavin, če je dogodek krajši od 1,67 ure, in da je prestreženo 47,2 % padavin, če je dogodek daljši od 1,67 ure. Če je temperatura v času dogodka z manj kot 2,5 mm količine padavin višja od 14,2 °C, pa je prestreženo kar 95,5 % padavin nad krošnjami dreves. Pri dogodkih z več kot 2,5 in manj kot 7,5 mm količine padavin drevesne krošnje prestrežejo 42,8 % padavin, pri dogodkih z več kot 7,5 mm količine padavin nad krošnjami dreves pa je v povprečju prestreženih 23,2 % padavin. Generirano regresijsko drevo primera #2 (slika 5) nam v primeru dogodkov z več kot 7,5 mm količine padavin razkriva še dodatne klimatske faktorje, ki vplivajo na proces; to sta povprečna hitrost vetra ter intenziteta in trajanje padavin. Pri dogodkih z nad 7,5 mm padavin je v primeru povprečne hitrosti vetra nižje od 3,45 m/s prestreženih približno 10 % manj padavin kot pri dogodkih s povprečno hitrostjo vetra višjo od 3,45 m/s.

```

KolicinaPadavin <= 2.5 :
|   Temperatura <= 14.2 :
|   |   TrajanjePadavin <= 1.67 : ProcentEi = 81.2
|   |   TrajanjePadavin > 1.67 : ProcentEi = 47.2
|   Temperatura > 14.2 : ProcentEi = 95.5
KolicinaPadavin > 2.5 :
|   KolicinaPadavin <= 7.5 : ProcentEi = 42.8
|   KolicinaPadavin > 7.5 : ProcentEi = 23.2

```

Slika 7 – Razviti model v obliki regresijskega drevesa za primer #3 (metoda strojnega učenja M5).

### Zaključki

V okviru raziskave so bile na podlagi izvedenih meritev in analize rezultatov le-teh z metodami strojnega učenja ugotovljene značilnosti vpliva zaraščanja opuščениh kmetijskih zemljišč na vodno bilanco porečja Dragonje. Gozdovi na novozaraščениh površinah prestrežejo skoraj vso vodo padavinskih dogodkov z manj kot 2,5 mm padavin, kar pomeni da v takih primerih ne pride do površinskega in podpovršinskega odtoka v Dragonjo ali napajanja zalog podzemne vode v porečju. Izjema pri dogodkih z manj kot 2,5 mm padavin so le dogodki daljši od 1,67 ure in s temperaturo nižjo od 14 °C, kar bi lahko predstavljalo tudi vpliv horizontalnih padavin oziroma megle; takrat je namreč pretreženih le približno polovica padavin. Za dogodke z več kot 2,5 mm padavin pa generirani modeli kažejo, da novi gozd prestreže od 42,8 % padle vode pri dogodkih s količino padavin od 2,5 do 7,5 mm in 23,2 % padle vode pri dogodkih s količino padavin večjo od 7,5 mm, ko pa imata na vrednost deleža prestreženih padavin močan vpliv tudi povprečna hitrost vetra in intenziteta padavin.

Klasifikacijska in regresijska drevesa, ki smo jih iz rezultatov meritev zgenerirali s pomočjo metod strojnega učenja J4.8 in M5, transparentno kažejo stopnje vplivov in interakcij posameznih faktorjev, ki pomembno vplivajo na delež prestreženih padavin na gozdnih območjih. Če rezultate z obeh raziskovalnih ploskev ekstrapoliramo na celotno porečje Dragonje, lahko ugotovimo, da je vpliv spremenjene dejanske rabe zemljišč v zadnjih desetletjih na vodno bilanco znotraj porečja precej visok in nikakor zanemarljiv. Z vidika oskrbe z vodo, ki (zaenkrat) predstavlja velik razvojni problem obalnega območja, lahko na spremembe v vodni bilanci zaradi zaraščanja opuščениh kmetijskih zemljišč



gledamo kot na izrazito nezaželen pojav, saj je kar približno tretjina vode na območjih, poraslih z gozdom, dobesedno izgubljena.

Metode strojnega učenja za generiranje modelov v obliki odločitvenih dreves predstavljajo zelo uporaben in obetajoč pripomoček za iskanje novih znanj v podatkovnih bazah. Pri tem pa je treba poudariti, da je sam razvoj modela le eden izmed korakov pri iskanju novih znanj, saj izbor, zbiranje in priprava podatkov, nadzor nad razvojem modela in njegova interpretacija, ki pa jih lahko opravi samo področni strokovnjak, predstavljajo prav tako pomembne stopnje razvoja novih znanj.

### **Viri in literatura**

- Breiman L, Friedman JH, Olshen RA, Stone CJ. (1984) Classification and regression trees. Wadsworth, Belmont.
- Brilly M, Sraj M. (2000) Osnove hidrologije (Principles of Hydrology). University Textbook, University of Ljubljana, Faculty of Civil and Geodetic Engineering.
- Chow VT. (1964) Handbook of applied hydrology. McGraw-Hill, New York.
- Dawson CW, Wilby RL. (2001) Hydrological modelling using artificial neural networks. Progress in Physical Geography 25(1): 80–108.
- Globevnik L. (2001) Celosten pristop k urejanju voda v povodjih. Doktorska disertacija, Univerza v Ljubljani.
- Kompare B. (1995) The use of artificial intelligence in ecological modelling. Ph. D. Thesis, Royal Danish School of Pharmacy, Copenhagen, Denmark.
- Mikoš M, Kranjc A, Matičič B, Müller J, Rakovec J, Roš M, Brilly M. (2002). Hidrološko izrazje. Acta hydrotechnica 20/32: 3–324.
- Mitchell T. (1997) Machine Learning. MIT Press and The McGraw-Hill Companies, Inc.
- Ovington JD. (1954) A comparison of rainfall in different woodlands. Forestry London 27, pp 41–53.
- Quinlan JR. (1986) Induction of Decision Trees. Machine Learning 1: 81–106.
- Quinlan JR. (1992) Learning with continuous classes. In: Proceedings of the Fifth Australian Joint Conference on Artificial Intelligence, pp 343–348.
- Solomatine DP, Dulal KN. (2003) Model trees as an alternative to neural networks in rainfall-runoff modelling. Hydrological Sciences Journal 48: 399–411.
- Šraj M. (2003) Estimating leaf area index of the deciduous forest in the Dragonja watershed – Part I: Methods and measuring. Acta Hydrotechnica 21/35: 105–127.
- Šraj M. (2003) Modeliranje in merjenje prestreženih padavin. Doktorska disertacija, Univerza v Ljubljani.
- Štravs L, Kobold M, Brilly M. (2004) Modeli kratkoročnih napovedi pretokov visokih voda na Savinji. Zbornik – Mišičev vodarski dan, Maribor.
- Witten I H, Frank E. (2000) Data mining: Practical machine learning tools and techniques with java implementations. Morgan Kaufmann Publishers, San Francisco, USA.