

# UPORABA METOD STROJNEGA UČENJA ZA PREUČEVANJE ODNOSOV MED ZNAČILNOSTMI BRANIK IN OKOLJEM

## ON THE USE OF MACHINE LEARNING METHODS TO STUDY THE RELATIONSHIPS BETWEEN TREE-RING CHARACTERISTICS AND THE ENVIRONMENT

Jernej JEVŠENAK<sup>1</sup>, Sašo DŽEROSKI<sup>2,3</sup>, Tom LEVANIČ<sup>4</sup>

(1) Gozdarski inštitut Slovenije, Večna pot 2, SI-1000 Ljubljana; jernej.jevsenak@gozdis.si

(2) Inštitut »Jožef Stefan«, Jamova cesta 39, SI-1000 Ljubljana; saso.dzeroski@ijs.si

(3) Mednarodna podiplomska šola Jožefa Stefana, Jamova cesta 39, SI-1000 Ljubljana; saso.dzeroski@ijs.si

(4) Gozdarski inštitut Slovenije, Večna pot 2, SI-1000 Ljubljana; tom.levanic@gozdis.si

### IZVLEČEK

Različne študije so pokazale, da lahko z nelinearnimi metodami bolje opišemo (modeliramo) odnos med branikami in okoljem. V naši študiji smo primerjali (multiplo) linearno regresijo (MLR) in štiri nelinearne metode strojnega učenja: modelna drevesa (MT), ansambel *bagging* modelnih dreves (BMT), umetne nevronske mreže (ANN) in metodo naključnih gozdov (RF). Za primerjavo teh metod modeliranja smo uporabili štiri množice podatkov. Natančnost naučenih modelov smo ocenili z metodo 10-kratnega prečnega preverjanja (ang. *10-fold cross-validation*) na naši množici in preverjanjem na dodatni testni množici. Na vseh množicah smo dobili boljše statistične kazalce za nelinearne metode s področja strojnega učenja, s katerimi lahko pojasnimo večji delež variance oz. dobimo manjšo napako. Nobena metoda se ni pokazala kot najboljša v vseh primerih, zato je smiselno predhodno primerjati več različnih metod in nato uporabiti najprimernejšo, npr. za rekonstrukcijo klime.

**Ključne besede:** strojno učenje, primerjava metod, dendroklimatologija, umetne nevronske mreže, modelna drevesa, ansambel modelnih dreves, naključni gozdovi, linearna regresija

### ABSTRACT

Many studies have shown that by using nonlinear methods, the relationship between tree-ring parameters and the environment can be described (modelled) better and in more detail. In our study, (multiple) linear regression (MLR) with four nonlinear machine learning methods are compared: artificial neural networks (ANN), model trees (MT), bagging of model trees (BMT) and random forests of regression trees (RF). To compare the different regression methods, four datasets were used. The performance of the learned models was estimated by using 10-fold cross-validation and an additional hold-out test. For all datasets, better results were obtained by the nonlinear machine learning regression methods, which can explain more variance and yield lower error. However, none of the considered methods outperformed all other methods for all datasets. Therefore, we suggest testing several different methods before selecting the best one, e.g. for climate reconstruction.

**Key words:** machine learning, method comparison, dendroclimatology, artificial neural networks, model trees, regression trees, ensembles of model trees, random forest, linear regression

GDK 52:111.8(045)=163.6  
DOI 10.20315/ASetL.114.2

Prispelo / Received: 4. 12. 2017  
Sprejeto / Accepted: 8. 1. 2018

## 1 UVOD

### 1 INTRODUCTION

Drevesa zmernih geografskih širin vsako leto proizvedejo eno širino branike, to je letni prirastek, v kateri so shranjene informacije o okolju v času njihovega nastanka. Z analizo izvrtkov tako dobimo vpogled v preteklost dreves, sestojev in posredno tudi okolja. Razumevanje rasti dreves je pomembno za optimalno sprejemanje odločitev pri načrtovanju in gojenju gozdov, za natančnejše napovedovanje rasti v prihodnosti in natančnejše rekonstrukcijo klime v preteklosti.

Z analizo branik in primerjavo njihovih značilnosti z okoljskimi podatki (npr. temperature) dobimo vpogled v odnos med rastnimi parametri in okoljem. Ta odnos se navadno opiše z linearnim regresijskim modelom. Različne študije kažejo, da je odnos med klimo in parametri branik bolj ali manj nelinearen (Evans in sod., 2006; Helama in sod., 2009). Linearnost je mogoča na ožjem spektru odzivnega intervala, na širšem spektru pa se spreminja, saj drevesa drugače uspevajo v optimalnih in suboptimalnih razmerah.

Z informacijsko revolucijo računalniki postajajo zmogljivejši in v zadnjih desetletjih se vse bolj uveljavljajo t.i. metode strojnega učenja (ang. *Machine Learning*). Strojno učenje je področje umetne inteligence, ki se osredotoča na načrtovanje algoritmov za prepoznavanje kompleksnih vzorcev in sprejemanje inteligentnih odločitev (Pohorec, 2013). Z metodami strojnega učenja lahko rešujemo različne vrste problemov:

- klasifikacijske probleme (npr. zdravo / odmrlo drevo; aktiven / neaktiven lastnik),
- razvrščanje (npr. razvrščanje dreves v skupine glede na prirastek),
- regresijske probleme (npr. odnos med parametri branik in okoljem).

V našem prispevku želimo 1) predstaviti izbrane nelinearne metode strojnega učenja za reševanje regresijskih problemov in 2) predstaviti in primerjati rezultate linearnih (okrajšava MLR) in nelinearnih modelov na štirih različnih naborih (množicah) podatkov.

## 2 METODE

### 2 METHODS

#### 2.1 Nelinearne metode strojnega učenja

##### 2.1 Nonlinear machine learning methods

###### 2.1.1 Drevesne metode

###### 2.1.1 Tree-based methods

Odločitvena drevesa (Quinlan, 1986) so hierarhične podatkovne strukture. Sestavljena so iz notranjih vozlišč, vej in listov. List vsebuje napovedi vrednosti razreda oz. ciljne spremenljivke. Če ima ciljna (odvisna) spremenljivka zvezne vrednosti, govorimo o regresijskem drevesu. Kadar so lahko v listih regresijskih dreves linearne enačbe, govorimo o modelnih drevesih (Quinlan, 1992; an. Model Trees, okrajšava MT; Slika 1b). Ta drevesa so analogna odsekoma linearnim funkcijam. Odločitveno drevo ponazarja graf ali model odločitev in možnih posledic. Interpretacija odločitvenih dreves je enostavna in lahko razumljiva (primer na sliki 4). Drevesa opisujejo odnos med odvisno in eno ali več neodvisnimi spremenljivkami. Odločitvena drevesa ne zahtevajo nobenih predpostavk o porazdelitvah spremenljivk.

Pri graditvi odločitvenega drevesa je osnovna naloga algoritma oceniti pomembnost posameznega atributa oz. neodvisne spremenljivke za podani učni problem. Najpogosteje se uporabljajo naslednje mere: informacijski prispevek (ang. *information gain*), razmerje informacijskega prispevka (ang. *gain-ratio*), Gini-indeks, Relief in razlika variance. Enostavna odločitvena drevesa se pogosto izkažejo za boljša, zato je treba kompleksna drevesa obrezati (ang. *pruning*).

Obrezovanje je ključno, da se izognemo problemu pretiranega prilaganja (ang. *overfitting*). O pretiranem prilaganju govorimo, kadar je model preveč podroben in kompleksen ter posledično zelo dobro deluje na učnih podatkih, vendar na neodvisnih podatkih daje izredno slabe rezultate.

Uporaba ansamblov pomeni, da končni model ponazarja kombinacijo (povprečje) več individualnih modelov. Natančnost ansambla za klasifikacijske in regresijske probleme je večja kot natančnost posameznih članov. Za uspešno delovanje ansamblov morajo biti posamezni členi raznoliki, torej se morajo učiti na različnih podmnožicah učnih podatkov oziroma uporabljati različne učne algoritme (Pohorec, 2013). V našem prispevku predstavljamo naključne gozdove (ang. *Random Forests*, okrajšava RF) in *bagging* (Breiman, 1996).

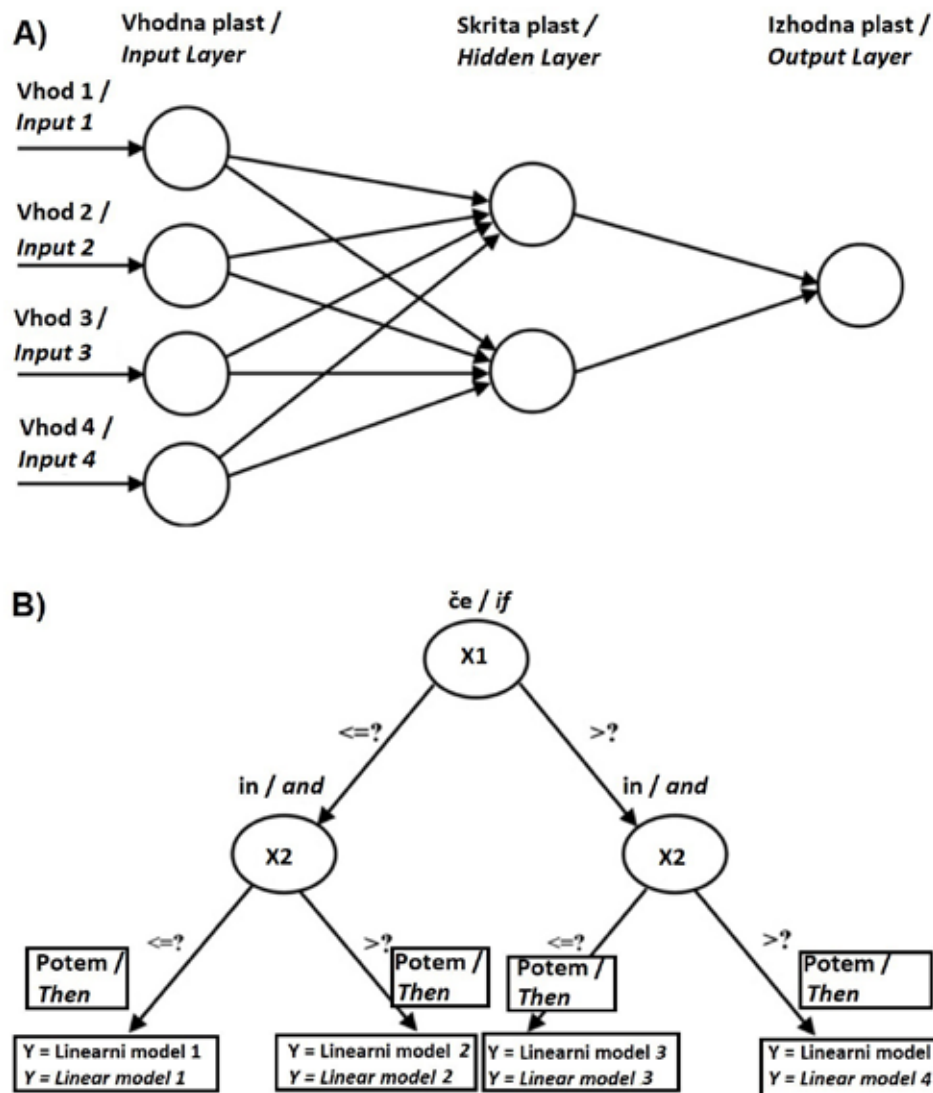
Metoda *bagging* izkorišča nestabilnost učnih algoritmov. Najprej se z uporabo vzorčenja z zamenjavo oz. postopkom *bootstrap* tvori  $n$  učnih podmnožic. Vsako podmnožico uporabimo za graditev ločenega modela. Ob uporabi ansambla vsak od  $n$  članov vrne svojo napoved. *Bagging* vrne kot končni rezultat povprečje vseh napovedi. V naši raziskavi smo metodo *bagging* uporabili skupaj z metodo modelnih dreves (BMT) in gradili gozdove tovrstnih dreves.

Naključne gozdove sestavljajo odločitvena drevesa, naučena z delnim naključnim postopkom. Z uporabo postopka *bootstrap* iz osnovne učne množice naredimo izbrano število učnih množic. Nad vsako množico nato zgradimo odločitveno drevo, pri čemer je atribut, na osnovi katerega razcepimo vozlišče v drevesu, vsakokrat izbran izmed določenega števila naključno izbranih atributov. Napovedi vseh zgrajenih dreves združimo z večinskim glasovanjem.

###### 2.1.2 Umetne nevronske mreže

###### 2.1.2 Artificial neural networks

Umetne nevronske mreže (Bishop, 1995; an. Artificial Neural Networks; okrajšava ANN; slika 1a) so skupina metod strojnega učenja, ki posnemajo delovanje človeških možganov. ANN imajo možnost razpoznavanja nelinearnih funkcij med vhodnimi in izhodnimi spremenljivkami in so zato zelo primerne za modeliranje nelinearnih odnosov. Sestavljene so iz vhodne, izhodne in ene ali več skritih plasti. Vhodna plast sestavljajo neodvisne spremenljivke, izhodna plast pa odvisna spremenljivka. Skrito plast sestavlja poljubno število nevronov, ki določajo kompleksnost ANN in so povezani z vhodnimi in izhodnimi pastmi. Te povezave vsebujejo uteži, ki se v iteracijskem procesu prilagajajo tako, da vhodne spremenljivke ustrezajo iz-



**Slika 1:** Shematski prikaz A) umetnih nevronske mreže in B) modelnih dreves. Vir: Gharaei-Manesh in sod. (2016)

**Fig. 1:** Schematic representation of A) artificial neural network and B) model trees. Source: Gharaei-Manesh et al. (2016).

hodnemu signalu. V vsaki iteraciji se izračuna napaka med simuliranimi in opazovanimi vrednostmi. Uteži se spreminjajo, dokler ni funkcija napake minimizirana. V postopek učenja uteži smo vključili Bayesovo regularizacijo (Burden in Winkler, 2008), ki prepreči pretirano prilaganje nevronske mreže podatkom.

## 2.2 Strategija primerjav različnih metod

### 2.2 The model comparison strategy

Za ocenjevanje linearnih in nelinearnih modelov smo podatke najprej razdelili na učno (kalibracijsko) množico in validacijsko (testno) množico (ang. *holdout set*). Slednja je vsebovala 10 % najstarejših let. Na kalibracijski smo najprej naredili 10 ponovitev 10-kratnega prečnega preverjanja (ang. *10-fold cross-validation*; okrajšava CV). Pri tej metodi podatke razdelimo na 10 enakih delov (ang. *folds*). Nato modele sistematično kalibriramo na 9 *foldih* in testiramo na preostalem *foldu*.

Končne modele (zgrajene na celotni učni množici) smo nato uporabili še na testni množici.

Optimalne vrednosti parametrov za nelinearne modele smo določili s pomočjo R paketa *caret* (Kuhn in sod., 2017). Funkcija *train* iz R paketa *caret* sistematično testira različne nastavitve parametrov in predlaga tiste vrednosti, ki dajejo najboljše rezultate na neodvisnih podatkih.

Za prečno preverjanje in validacijsko množico smo izračunali naslednje testne statistike: korelacijski koeficient ( $r$ ), *reduction of error* (RE; Fritts, 1976; Lorenz, 1956), *coefficient of efficiency* (Fritts, 1976), koren povprečne kvadrirane napake (RMSE; Willmott, 1981) in koren relativne kvadrirane napake (RRSE; Witten in sod., 2011). Modele z večjimi vrednostmi  $r$ , RE in CE ter z manjšimi napakami (RMSE in RRSE) štejemo za boljše.

### 2.3 Podatki za primerjavo metod strojnega učenja

#### 2.3 Data for method comparison of machine learning methods

Za primerjavo linearnih in nelinearnih metod modeliranja smo uporabili štiri množice podatkov (preglednica 2). Pri vseh množicah odvisno spremenljivko predstavljajo povprečne temperature v rastni sezoni (pri vsaki je kombinacija nekoliko drugačna, glej preglednico 2). Odvisno spremenljivko smo izbrali na podlagi preliminarne korelacijske analize.

Pri treh množicah podatkov (Albanija, Alpe in Srem) smo uporabili eno neodvisno spremenljivko (standardizirana širina branike), preostala množica pa vsebuje tri neodvisne spremenljivke, to so povprečna velikost trahej, stabilni izotopi ogljika ( $\delta^{13}\text{C}$ ) in stabilni izotopi kisika ( $\delta^{18}\text{O}$ ). Za lažjo interpretacijo posameznih vrednosti smo neodvisni spremenljivki  $\delta^{13}\text{C}$  in  $\delta^{18}\text{O}$  normalizirali po enačbi (1), pri čemer  $\bar{x}$  ponazarja povprečje za  $x$  in  $sd(x)$  njegov standardni odklon.

$$x_i = \frac{(x_i - \bar{x})}{sd(x)} \quad \dots(1)$$

### 3 REZULTATI IN DISKUSIJA

#### 3 RESULTS AND DISCUSSION

Primerjava med linearno regresijo in nelinearnimi metodami strojnega učenja je pokazala, da lahko v vseh primerih napredni nelinearni algoritmi z večjo natančnostjo opišejo odnos med odvisno in neodvisni-

mi spremenljivkami. Razlike med metodami sicer niso velike, a so na vseh množicah podatkov v prid nelinearnim metodam strojnega učenja. Na podlagi naših rezultatov ne moremo izpostaviti samo ene posamezne metode, ki bi na vseh rastiščih dala najboljše rezultate (slika 2).

Na rastišču Albanija je najboljše testne vrednosti statističnih kazalcev dala metoda ANN, čeprav smo z metodo MLR dobili le nekoliko slabše vrednosti testnih statistik. Na podatkih za macesnovo rastišče (Alpe) se je pri oceni napake s postopkom CV najboljše obnesla metoda ANN, pri oceni na testni množici pa metoda BMT.

Pri rastišču Srem smo pri oceni s postopkom CV dobili najboljše rezultate z metodo BMT, pri oceni na testni množici pa za metodo RF. Za rastišče Sorško polje sta metodi RF in BMT dali najboljše rezultate pri oceni s postopkom CV, metoda BMT pa pri oceni na testni množici. V nobenem primeru nismo dobili najboljših ocen z metodo MLR. Čeprav smo pri rastišču Srem uporabili tri neodvisne spremenljivke, so vrednosti statističnih kazalcev v splošnem primerljivi z drugimi rastišči. Torej dodajanje spremenljivk ne pomeni nujno tudi boljšega napovedovanja.

#### 4 ZAKLJUČKI

#### 4 CONCLUSIONS

Z regresijskimi statističnimi modeli opisujemo odnose med odvisnimi in neodvisnimi spremenljivkami. V naši študiji smo pokazali, da lahko z nelinearnimi

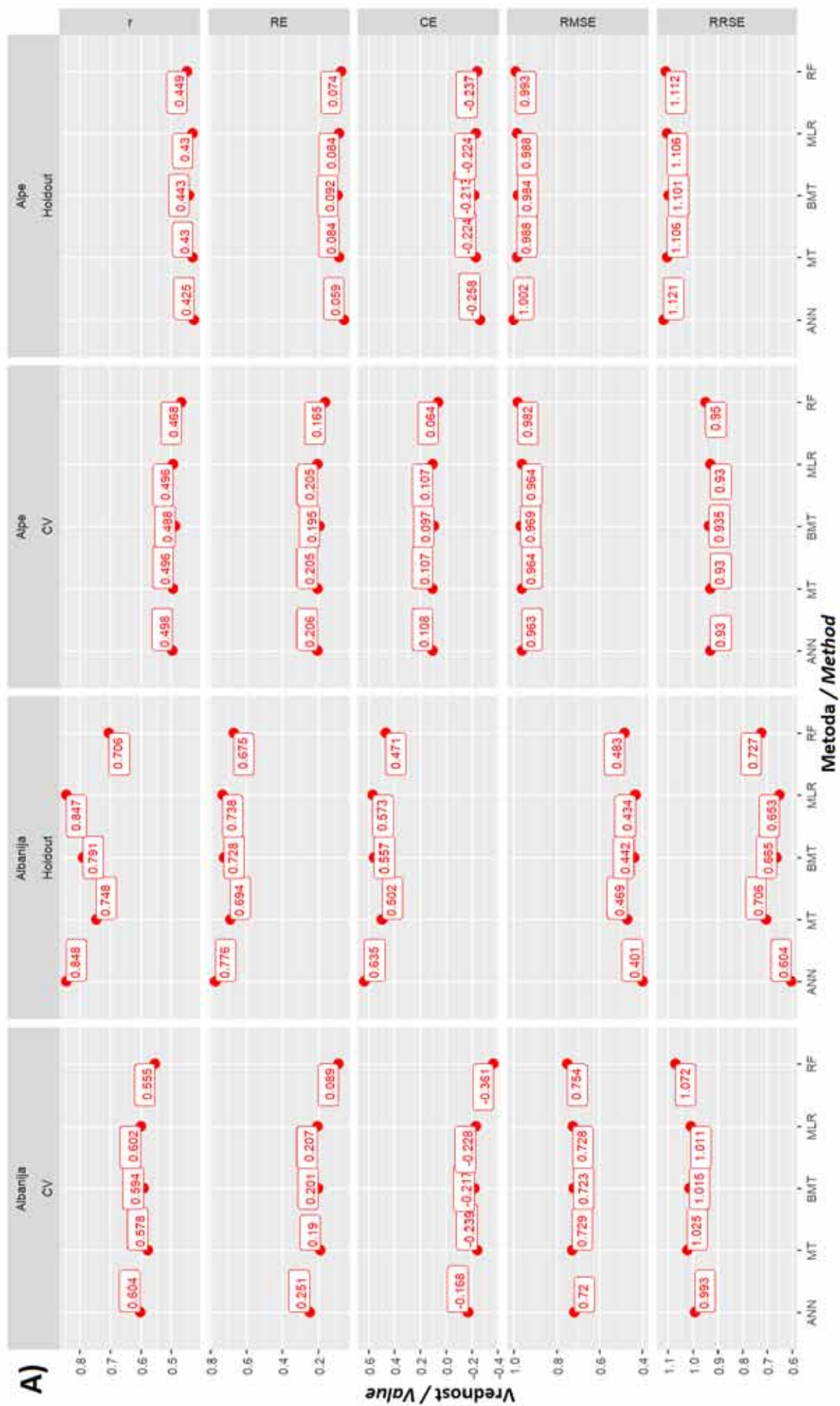
**Preglednica 1:** Optimizirane vrednosti parametrov nelinearnih metod strojnega učenja: umetne nevronske mreže (ANN), modelna drevesa (MT), bagging modelnih dreves (BMT) in naključni gozdovi (RF).

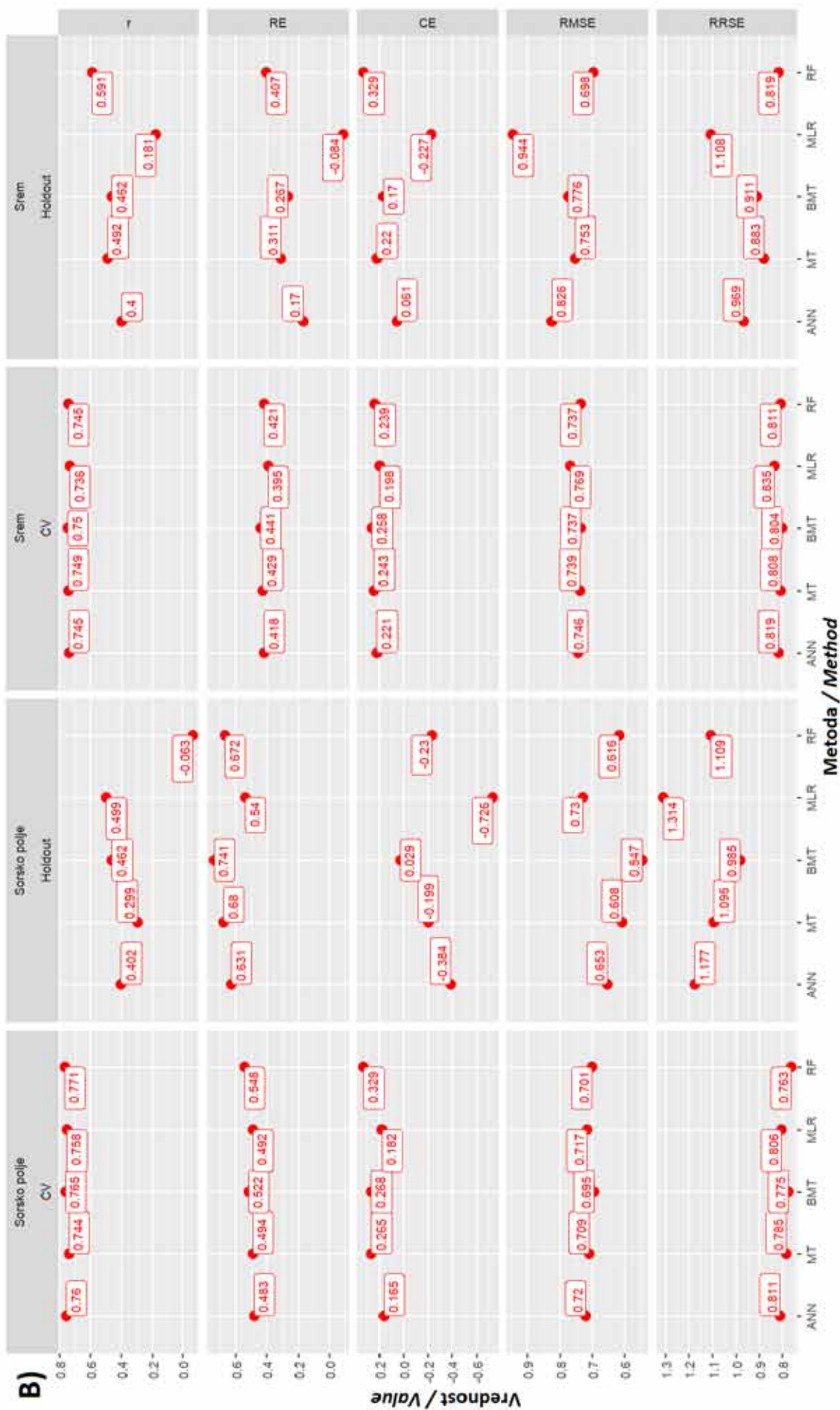
**Table 1:** Tuned values of parameters for nonlinear machine learning methods: artificial neural networks (ANN), model trees (MT), bagging of model trees (BMT) and random forests of regression trees (RF).

Metoda Method	Parameter – opis Parameter - description	Optimizirane vrednosti parametrov Tuned parameter values			
		Albanija Albania	Alpe Alps	Srem Srem	Sorško polje Sorsko plain
ANN	neurons – število nevronov	1	1	2	1
MT	M – minimalno število enot v listu	4	4	4	4
MT	N – uporabi neobrezana drevesa	TRUE	FALSE	TRUE	TRUE
MT	U – uporabi nezglajene napovedi	FALSE	FALSE	FALSE	FALSE
MT	R – sestavi regresijsko drevo (namesto modelnega)	FALSE	FALSE	FALSE	FALSE
BMT	P – velikost podmnožice, kot odstotek od velikosti kalibracijske množice	80	80	80	80
BMT	I – število iteracij	100	200	200	200
BMT	M – minimalno število enot na list	4	4	4	4
BMT	N – uporabi neobrezana drevesa	TRUE	FALSE	TRUE	TRUE
BMT	U – uporabi nezglajene napovedi	FALSE	FALSE	FALSE	FALSE
BMT	R – sestavi regresijsko drevo (namesto modelnega)	FALSE	FALSE	FALSE	FALSE
RF	P – velikost podmnožice, kot odstotek od velikosti kalibracijske množice	80	80	80	80
RF	I – število iteracij	200	200	200	200
RF	depth – največja globina drevesa	2	2	2	2

**Preglednica 2:** Pregled podatkov za primerjavo modelov**Table 2:** About the data used to compare models

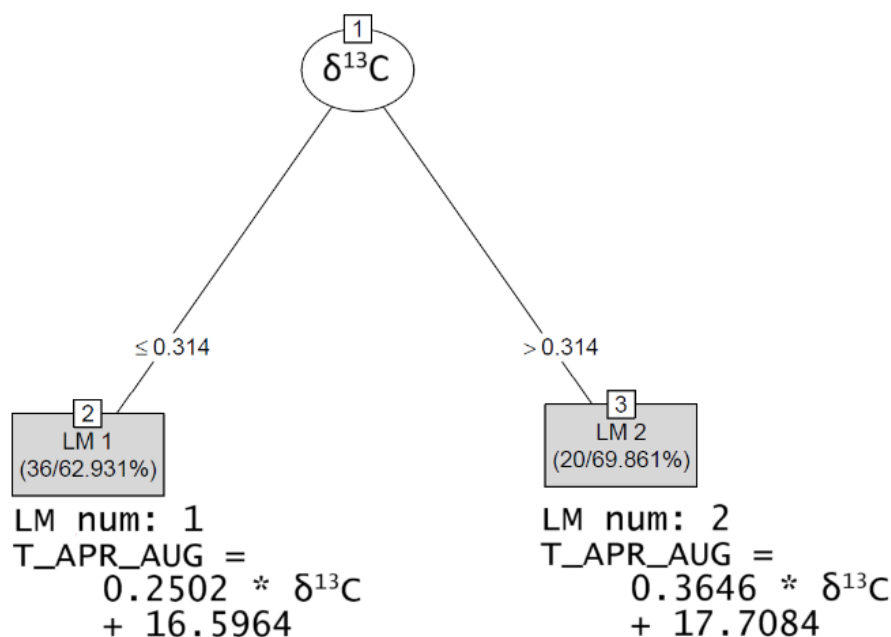
Množica Data set	Rastišče (drevesna vrsta) Site (Tree species)	Meteorološki podatki Meteorological data	Neodvisne spremenljivke Independent variables	Odvisna spremenljivka (povprečne temperature) Dependent variable (mean temperature)	Prečno preverjanje Cross-validation	Validacijska (holdout) množica Validation (holdout) set
1	Albanija / Albania ( <i>Pinus nigra</i> )	Albania, (KNMI Climate Explorer, 2017)	širina branike / tree-ring width	junij - julij June - July	1958 – 2009	1951 – 1957
2	Alpe / Alps ( <i>Larix decidua</i> )	Villacher Alpe (Auer in sod., 2007)	širina branike / tree-ring width	maj - julij May - July	1885 – 2009	1870 – 1884
3	Srem / Srem ( <i>Quercus robur</i> )	Ljubljana (ARSO, 2017)	širina branike / tree-ring width	april - julij / April - July	1950 – 2013	1942 – 1949
4	Sorško polje / Sorško plain ( <i>Quercus robur</i> )	Ljubljana (ARSO, 2017)	povprečna velikost traneje / mean vessel area stabilni izotopi ogljika / stable carbon isotope ratios stabilni izotopi kisika / stable oxygen isotope ratios	april - avgust / April - August	1967 - 2015	1960 - 1966





**Fig. 2:** Comparison of the results of different methods based on cross-validation (CV) and hold-out validation for A) the sites Albania and Alps and B) the sites Srem and Sorsko plain. The compared methods are: multiple linear regression (MLR), artificial neural networks (ANN), model trees (MT), bagging of model trees (BMT) and random forests of regression trees (RF).

**Slika 2:** Primerjava rezultatov prečnega preverjanja (ang. *cross-validation*, okrajšava CV) in dodatnega testa na validacijski množici (ang. *hold-out test*, okrajšava *Holdout*) za A) rastišči Albanija in Alpe ter B) rastišči Srem in Sorsko polje. Uporabljene metode so: multipla linearna regresija (MLR), umetne nevronske mreže (ANN), modelna drevesa (MT), *bagging* modelnih dreves (BMT) in naključni gozdovi regresijskih dreves (RF).



**Slika 3:** Primer modelnega drevesa za rastišče Sorško polje z dvema linearnima enačbama. Glede na vrednost neodvisne spremenljivke  $\delta^{13}\text{C}$  bosta za končno napoved uporabljeni LM1 ( $\delta^{13}\text{C} \leq 0.314$ ) ali LM2 ( $\delta^{13}\text{C} > 0.314$ ).

**Fig. 3:** An example of model tree for the site Sorsko plain with two linear equations. Based on value of the independent variable  $\delta^{13}\text{C}$ , LM1 ( $\delta^{13}\text{C} \leq 0.314$ ) or LM2 ( $\delta^{13}\text{C} > 0.314$ ) will be used for final prediction..

metodami strojnega učenja bolje opišemo odnos med parametri branik in temperaturami. Uporaba in razumevanje metod strojnega učenja postajata v znanosti vse pomembnejša za opisovanje okoljskih pojavov in napovedovanje dogodkov. V našem prispevku smo prikazali delovanje teh metod na primeru dendrokronologije, vendar pa je njihova potencialna uporabnost v gozdarstvu in sorodnih področjih precej večja.

Metode strojnega učenja so primerne za reševanje regresijskih in klasifikacijskih problemov, uporaba je možna na numeričnih in slikovnih podatkih. Ob nadaljnjem razvoju algoritmov lahko pričakujemo še natančnejše in boljše napovedne sposobnosti ter širšo uporabo metod strojnega učenja.

## 5 SUMMARY

Linear regression is the most often used method for studying the relationships between tree-ring parameters and the environment. Many studies have shown that by using nonlinear methods, this relationship could be described more precisely and accurately. In our study, (multiple) linear regression (MLR) is compared with four nonlinear machine learning methods: artificial neural networks (ANN), model trees (MT), bagging of model trees (BMT) and random forests of regression trees (RF). To compare the different regression methods, four dendroclimatological datasets were used. For all datasets, the dependent variable is

a combination of mean monthly temperatures. Three datasets (Albania, Alps and Srem) contain only one independent variable (standardized tree-ring width; TRW), while the remaining dataset includes three independent variables, i.e. mean vessel area (MVA), stable carbon ( $\delta^{13}\text{C}$ ) and oxygen ( $\delta^{18}\text{O}$ ) isotope ratios.

The model performance was estimated on independent data, i.e. data that was not used to train (calibrate) models. To estimate the model performance, 10-fold cross-validation on the training set and an additional hold-out test were used. The hold-out data consisted of 10 % of the oldest rings.

Our results showed that by using robust nonlinear machine learning methods, better predictions were obtained for all four datasets. The best results for the site Albania (*Pinus nigra*) were provided by ANN. For the Alpine site (*Larix decidua*), the best results were obtained for ANN (CV) and BMT (hold-out test). For the site Srem (*Quercus robur*), the best results were obtained for BMT (CV) and RF (hold-out test). For the multivariate dataset (site Sorsko plain, *Quercus robur*), the best results were calculated for BMT and RF (CV), and BMT (hold-out test), respectively. MLR did not provide the best results for any of the analysed sites. However, none of the considered methods outperformed all other methods for all datasets. Therefore, we suggest testing several methods before selecting the best one, e.g. for climate reconstruction.



**6 ZAHVALA****6 ACKNOWLEDGMENT**

Zahvaljujemo se Pahernikovi ustanovi, ki je v študijskih letih 2014/15, 2015/16 in delno 2016/17 podelila štipendijo za doktorski študij Jerneju Jevšenaku. Tom Levanič in Jernej Jevšenak sta bila finančno podprta s strani ARRS projektov J4-5519, J4-8216 in N4-0065 ter programske skupine P4-0107 »Gozdna biologija, ekologija in tehnologija«.

**7 VIRI****7 REFERENCES**

- ARSO. 2017. Arhiv - opazovani in merjeni meteorološki podatki po Sloveniji. <http://meteo.arso.gov.si/met/sl/archive/> (8. avgust, 2017)
- Auer I., Bohm R., Jurkovic A., Lipa W., Orlik A. in sod. 2007. HISTALP - historical instrumental climatological surface time series of the Greater Alpine Region. *International Journal of Climatology*, 27, 1: 17-46
- Bishop C. M. 1995. *Neural Networks for Pattern Recognition*. Oxford, University Press, Inc.: 482 str.
- Breiman L. 1996. Bagging predictors. *Machine Learning*, 24, 2: 123-140
- Burden F., Winkler D. 2008. Bayesian Regularization of Neural Networks. V: *Artificial Neural Networks: Methods and Applications*. Livingstone D. J. (ur.). (Artificial Neural Networks: Methods and Applications, 458). Totowa, NJ, Humana Press: 23-42
- Evans M. N., Reichert B. K., Kaplan A., Anchukaitis K. J., Vaganov E. A. in sod. 2006. A forward modeling approach to paleoclimatic interpretation of tree-ring data. *Journal of Geophysical Research-Biogeosciences*, 111, G03008
- Fritts H. C. 1976. *Tree Rings and Climate*. London, Academic Press: 567 str.
- Gharaei-Manesh S., Fathzadeh A., Taghizadeh-Mehrjardi R. 2016. Comparison of artificial neural network and decision tree models in estimating spatial distribution of snow depth in a semi-arid region of Iran. *Cold Regions Science and Technology*, 122, 26-35
- Helama S., Makarenko N. G., Karimova L. M., Kruglun O. A., Timonen M. in sod. 2009. Dendroclimatic transfer functions revisited: Little Ice Age and Medieval Warm Period summer temperatures reconstructed using artificial neural networks and linear algorithms. *Annales Geophysicae*, 27, 3: 1097-1111
- KNMI Climate Explorer. 2017. <https://climexp.knmi.nl> (22. avgust 2017)
- Kuhn M., Wing J., Weston S., Williams A., Keefer C. in sod. 2017. caret: Classification and Regression Training. R package version 6.0-76, <https://CRAN.R-project.org/package=caret>
- Lorenz E. N. 1956. *Empirical Orthogonal Functions and Statistical Weather Prediction*. Massachusetts, Massachusetts Institute of Technology, Department of Meteorology: 49 str.
- Pohorec S. 2013. Vrednotenje metod strojnega učenja z metodami obdelave naravnega jezika: disertacija. (Univerza v Mariboru). Maribor, samozaložba: 155 str.
- Quinlan J. R. 1986. Induction of decision trees. *Machine Learning*, 1, 1: 81-106
- Quinlan J. R. 1992. Learning with continuous classes. V: *Proceedings of the 5<sup>th</sup> Australian Joint Conference on Artificial Intelligence (AI '92)*. Hobart, World Scientific: 343-348
- Willmott C. J. 1981. On the validation of models. *Physical Geography*, 2, 2: 184-194
- Witten I. H., Frank E., Hall M. A. 2011. *Data Mining: Practical Machine Learning Tools and Techniques*. Burlington, Morgan Kaufmann Publishers: 629 str.

titles. The legislation is at the end of the references. The titles of references by one author are arranged in chronological order and with a letter after the year of issue if there are more references in one year.

Abbreviations of journals should follow the World List of Scientific Periodicals or a similar work.

**Examples:**

See the latest issue of *Acta Silvae et Ligni* journal.

**12. LATIN NAMES**

Latin names of genera, species and intraspecies taxons should be written in italics. In phytocoenological papers all taxa should be written in italics.

**13. FORMAT OF THE MANUSCRIPT (MS)**

The MS should be typewritten using 1,5 spacing on one side of A4 paper in DOC or RTF format. The original and two copies of the MS should be sent to the Editor, together with the MS on 3,5' diskette or CD. Illustrative materials should be enclosed separately. The figures should be attached as files in TIFF, JPEG or WMF format.

**14. TABLES, GRAPHS, PHOTOGRAPHS, AND SUPPLEMENTS.**

Their markings have to be the same as quoted in the text. They should have the serial numbers, headings and English text (Author's responsibility). The headings for tables should be above the tables and the headings for graphs and photographs should be below them. Colour illustrations are accepted only by previous agreement with the editor.

**Tables:**

Avoid using the space bar for adjusting the columns, use "tab" key or the tools for creating tables in Word. The dimensions should not exceed 12,5 cm in width and 15 cm in height; the font of letters and numbers must not be smaller than 8 points.

**Graphs:**

- Lines should be at least 0,15 mm thick, and thicker lines should be in the following increments: 0,15 mm - 0,30 - 0,45mm - 0,60 mm - 0,75 mm - 0,90 mm, the thinnest line possible is 0.25 points.
- Hatchings should be composed from different combinations of lines; if you use grey tones, they should be at the most in five shades (0% - white, 25%, 50%, 75% and 100% - black)
- Fonts - 10 points or more, but should not exceed 14 points. Use the fonts ARIAL or ARIAL CE
- All the text should be in Slovene and English for example: Premer / Diameter, Leto / Year
- The intensity of dot is very important, it should be at least 80 lines per inch
- Do not use MS Excel because the parameters of the graph (width of lines, dimensions of graph, hatchings...) cannot be changed. Professional programmes for drawing graphs are recommended: GRAPHER, SURFER, SIG-MA plot.
- Graphs should be submitted on 3,5' diskette in WMF or EPS format
- The background of the graph should be white, do not use hatchings.
- If only one kind of data is presented in histogram columns, these should be white with a black border - do not use hatchings.
- 3D graphs are not allowed, use 2D graphs instead
- The graphs should not contain colours, they should be monochrome
- The dimensions should not exceed 12,5 cm in width and 15 cm in height

**Photographs:**

- Quality black and white photographs or slides made with classical camera. If the picture is taken with a digital camera the resolution should be at least 1,3 million pixels
- The dimensions should not exceed 12,5 cm in width and 15 cm in height
- Scanned pictures should be scanned with a resolution of 240 dpi
- All the pictures should be submitted in original TIFF or JPEG file (not in BMP or other formats). A paper print-out of the article with photographs included should be submitted for correct placement of the photo into the text.
- Pen-and-ink drawings have to be made with good drawing equipment (ROTRING pen or similar) or you should use computer drawing programme (Corel DRAW, FreeHand...). Consider the minimum width of lines, which is 0,25 points which corresponds 0,15 mm. Poor quality photographs and drawings will not be accepted.