NAPOVEDOVANJE DREVESNIH VRST IZ GEOMETRIJE IN INTENZITETE OBLAKA AEROLASERSKIH TOČK VRHOV DREVESNIH KROŠENJ

PREDICTING TREE SPECIES BASED ON THE GEOMETRY AND INTENSITY OF AERIAL LASER SCANNING POINT CLOUD OF TREETOPS

Nina Kranjec, Mihaela Triglav Čekada, Milan Kobal

UDK: 528.715:633/635.055
Klasifikacija prispevka po COBISS.SI: 1.01
Prispelo: 20. 4. 2021
Sprejeto: 27. 5. 2021

DOI: 10.15292/geodetski-vestnik.2021.02.234-259
SCIENTIFIC ARTICLE
Received: 20. 4. 2021
Accepted: 27. 5. 2021

SI | EN

IZVLEČEK

Na osnovi laserskih oblakov točk 240 posameznih dreves, ki smo jih identificirali tudi na terenu, smo razvili odločitvena drevesa za ločevanje listavcev in iglavcev ter posameznih drevesnih vrst (rdeči bor, navadna bukev, gorski javor, veliki jesen, evropski macesen, navadna smreka). Kot pojasnjevalne spremenljivke smo uporabili volumen zgornjega dela drevesne krošnje (višine 3 m) in povprečno intenziteto laserskih odbojev. Uporabili smo štiri nize aerolaserskih podatkov: iz maja 2012, septembra 2012, marca 2013 in julija 2015. Ugotovili smo, da najzanesljivejše rezultate za napovedovanje izbranih drevesnih vrst daje kombinacija volumna in povprečne intenzitete prvih treh laserskih nizov (uspešnost modela 60 %). Nekoliko nižjo uspešnost modela dobimo, če uporabimo samo povprečno intenziteto prvih treh nizov (54 %). Najslabšo uspešnost modela daje intenziteta niza 4, ki predstavlja lasersko skeniranje Slovenije (LSS) (31 %) oziroma volumen (21 %). Uspešnejše je razločevanje listavcev in iglavcev, ki na testnih podatkih dosega uspešnost 75 % oziroma 95 % (kombinacija volumna in povprečne intenzitete združenih prvih treh laserskih nizov). Če uporabimo samo intenzitete LSS, listavce in iglavce lahko ločimo z uspešnostjo 81 %.

ABSTRACT

Based on the laser point clouds of 240 individual trees that were also identified in the field, we developed decision trees to distinguish deciduous and coniferous trees and individual tree species: Picea abies, Larix decidua, Pinus sylvestris, Fagus sylvatica, Acer pseudoplatanus, Fraxinus excelsior. *The volume of the upper part of the tree crown (height of 3 m) and the average intensity of the laser reflections were used as explanatory variables. There were four aerial laser datasets: May 2012, September 2012, March 2013 and July 2015. We found that the combination of the volume and the average intensity of the first three laser datasets was the most reliable for predicting the selected tree species (60% model performance). A slightly poorer model performance was obtained if only the average intensity of the first three datasets was used (54% model performance). The worst model performance was given by the intensities (31 % model performance) or the volumes (21 % model performance) of dataset 4, which represents the national laser scanning of Slovenia (LSS). The best performing was the deciduous and coniferous separation, which achieved 75% and 95% success based on the test data (combination of volume and average intensity of the first three laser datasets). Using only the LSS intensities, deciduous and coniferous trees could be separated with 81% success.*

KLJUČNE BESEDE

lidar, intenziteta, geometrija drevesa, drevesne vrste, strojno učenje, odločitveno drevo

KEY WORDS

lidar, intensity, the geometry of tree, tree species, machine learning, decision tree

1 INTRODUCTION

Aerial laser scanning allows detailed studies of the forest at the landscape level, as well as at the level of forest stands and individual trees. This is because the laser beam is reflected from the top of the tree, branches, trunk and finally from the ground. These data can be used to study not only the horizontal but also the vertical structure of the forest. This allows an estimation of the proportion of coniferous and deciduous trees in mixed forests, the tree density, the health status and the damage to the trees (Kobal et al., 2014; Mongus et al., 2018; Benčina and Kobal, 2019; Karna et al., 2020). The first national laser scanning of Slovenia, which was performed in 2014–2015, is therefore also a very useful source of data for a detailed study of Slovenian forests (Šturm et al., 2016).

Laser scanning data can be studied in vector format, i.e., as the point cloud, and in derivative raster format, where we analyse the relationships between the different raster digital elevation models (relief, surface, canopy) or other raster thematic derivatives (e.g., voxels), where we display the values of selected attributes from the laser point cloud (e.g., intensity).

Individual trees can be distinguished in the laser point cloud by using methods that examine the local geometry in the point cloud, by delineating trees based on treetop identification, detecting edges between tree crowns, finding shadows, and region growing (Kobal et al., 2014; Mongus and Žalik, 2015; Eysn et al., 2016; Burt et al., 2018; Kansanen et al., 2019). A more challenging and not yet fully explored area is the identification of individual tree species based on independent laser data. Procedures for tree-species classification based only on laser data can be divided into (Shi et al., 2018):

- methods that examine the geometric characteristics of a particular tree species (e.g., crown shape and density, leaf and branch distribution, canopy openness);
- methods that study the radiometric properties of laser reflections (e.g., intensity value of a single reflection).

The intensity of an individual laser-beam reflection depends on the shape of the leaves or needles, the orientation and size of the leaves or needles, the density of the leaves or needles in the tree crown and the distribution of tree branches in the canopy (Holmgren and Person, 2004; Korpela et al., 2010; 2013). It also depends on the type of laser scanner and the parameters of the flight mission used (Triglav Čekada, 2011).

In Slovenia, Triglav Čekada et al. (2017) studied the differences between the average intensities of different tree species as seen in multi-temporal laser data taken using two wavelengths (1064 nm and 1550 nm). They found that the average intensities of the deciduous trees at both wavelengths are higher than those for conifers during the leaf-on period, while the opposite is true during the leaf-off period. When they studied the differences between the intensities classified as the first or the only reflections and the average intensities of all reflections, they found that the latter are lower during the leaf-on period, while they are higher than the intensities of the first or the only reflections during the leaf-off period. Ørka et al. (2009) and Suranto et al. (2009) found that the proportion of different orders of laser returns (first, middle, last return) depends on the tree species, due to the different canopy structure of the individual tree species, i.e., the density of the needles or foliage and their distribution and the distribution of the branches.

In this investigation, we added a study of the geometry of the uppermost 3 m of the individual tree-crown point cloud to a study of the average intensities of different tree species in a different part of the year in which the laser scanning was performed. We chose the top of the tree crown because this part of the tree is mostly not intermingled with neighbouring trees, and as Korpela et al. (2013) already noted, the differences in the distribution of tree branches with regard to the tree species can most easily be determined at the top of the trees. By analysing the average intensities of the individual trees' point clouds and analysing selected geometric parameters at different times of the year, we developed several predictive models based on machine learning – the decision-tree method was used. Based on multi-temporal laser datasets, the decision trees were developed to distinguish between the conifers and deciduous trees and to identify each studied tree species. In addition, we tested the performance of the decision-tree method on only one laser dataset, i.e., the national aerial laser scanning of Slovenia (LSS), as this would allow us to transfer this method from the test environment to a countrywide area.

2 DATA AND METHODOLOGY

2.1 Data

The test area is located near Podljubel, Slovenia, at an altitude of 700–800 m and is covered by mixed forests. The six tree species, which are the most frequent tree species in this test area, were considered in our study: Norway spruce (*Picea abies*), European larch (*Larix decidua*), Scots pine (*Pinus sylvestris*), European beech (*Fagus sylvatica*), sycamore (*Acer pseudoplatanus*) and ash (*Fraxinus excelsior*). For each tree species, we identified 40 trees in the field, being as similar as possible to each other in terms of canopy shape and size (Figure 1).

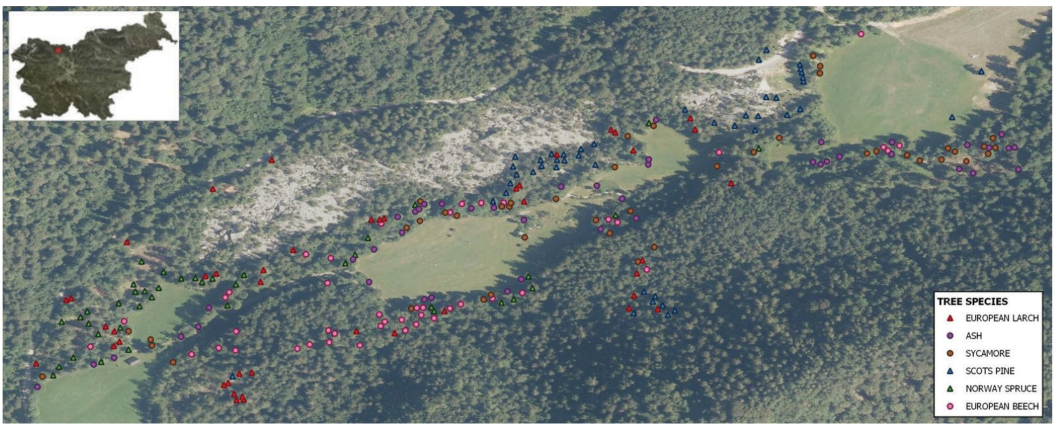


Figure 1: Locations of the selected trees (the source of the overview map of Slovenia: Geopedia; source of the base map: orthophoto taken at the same time as the laser scanning dataset 1).

We used the data from four datasets of aerial laser scanning (Table 1), taken in different years and at different times of the year. The first two datasets were scanned with the same Riegl LMS-Q560 scanner; the second two with the same type of scanner, i.e., the Riegl LMS-Q780. The first two datasets were scanned with a wavelength of 1550 nm; the second two, with 1064 nm. Three datasets were scanned during the leaf-on season and one dataset during the leaf-off season. These datasets also differ in terms of the average laser point density and the height of the flight above the ground. The data used in the analysis were in discrete pulse format, even though the full waveform laser scanner recorded the original data. In parallel with the first two laser datasets, an aerial photogrammetric survey was performed from the same platform and orthophotos were produced as well.

Table 1: Aerial laser scanning dataset characteristics.

	Dataset 1	Dataset 2	Dataset 3	Dataset 4 (LSS)
Date	15 May 2012	18 September 2012	5 March 2013	17 July 2014
Point density [point/m^2]	8	8	15	5
Wavelength [nm]	1550	1550	1064	1064
Height of flight [m]	700	700	1000	1200–1400
Leaf-on	Yes	Yes	No	Yes
Orthophoto	Yes	Yes	No	No

2.2 Preparation of aerial laser datasets

Before the field measurements of the individual trees were performed, we analysed an orthophoto taken at the same time as the laser dataset 1 in *ArcMap* (https://desktop.arcgis.com/en/arcmap/). Together with the orthophoto, we displayed the points of local maxima from the digital canopy model (DCM) created from the laser dataset 1. The DCM is the difference between the digital surface model (DSM) and the digital terrain model (DTM). In it the local maxima represent the treetops of the dominant trees. The combined orthophoto with the local maxima was exported as a georeferenced TIFF and imported into the *Avenza Maps* mobile application (https://www.avenza.com/avenza-maps/), which allows location and navigation on maps prepared in advance on a smartphone. In the field, we identified the location and tree species of each selected tree and added in the application a photograph of the trunk and the top of the tree crown. We selected trees that differed in terms of the maximum height from other trees in the forest by at least 3 m or the tree canopy at the upper 3 m did not intermingle with the tree crowns of other trees.

The field-measured tree locations were related in *ArcMap* to local maximum points, determined from the DCM. The trees were divided by tree species (Figure 1). For each tree, we defined a polygon in the shape of a circle centred on the local maximum point of the tree and with a radius of 5 m. These polygons were used to crop all the laser datasets into smaller point clouds representing each tree, i.e., individual tree-crown point clouds. The cropping was performed using the *Extract LAS* tool in *ArcMap*. To obtain only the data of a single tree, without intermingled tree canopies from other trees, we used only the upper 3 m of the tree crown. The tree crowns themselves were determined in *CloudCompare* (https://www.danielgm.net/cc/), and the tree crowns were cropped to only the top 3 m in *RStudio* (https://rstudio.com/), where all further analyses were performed as well.

2.3 Determining the single tree-canopy mean intensity and geometry

For an individual tree crown's point cloud, we calculated the average value of its intensity for each dataset separately. Here, we will compare only the relative ratios between the different trees in every dataset separately, as without an intensity normalisation the data scanned at different wavelengths cannot be directly compared.

For an individual treetop's point cloud we also calculated the following geometric features: the width of the narrowest and the width of the widest cross-sections of the crown top, the area of the narrowest and widest cross-sections, the floor area and the volume of the crown top. The widths of the narrowest and widest sections of the canopy top were determined by drawing an ellipse around all the points in the tree-crown point cloud in the plan view. In the floor plan view the shortest and longest semi-axes of the ellipses define the narrowest and the widest sections (Figure 2). The areas of the narrowest and widest cross-sections of the crown tops were determined by calculating the area of the figure represented by the convex hull of laser points mapped onto the vertical plane through the previously determined narrower and wider axes of the ellipse, respectively (Figure 3). The floor area of the tree-crown point cloud was determined by calculating the area of the figure representing the projection of the convex hull around all the points in the treetop-canopy point cloud onto the horizontal plane. The volume of the treetop crown was calculated as the volume of the solid formed by the convex hull around all the points that form the treetop point cloud.

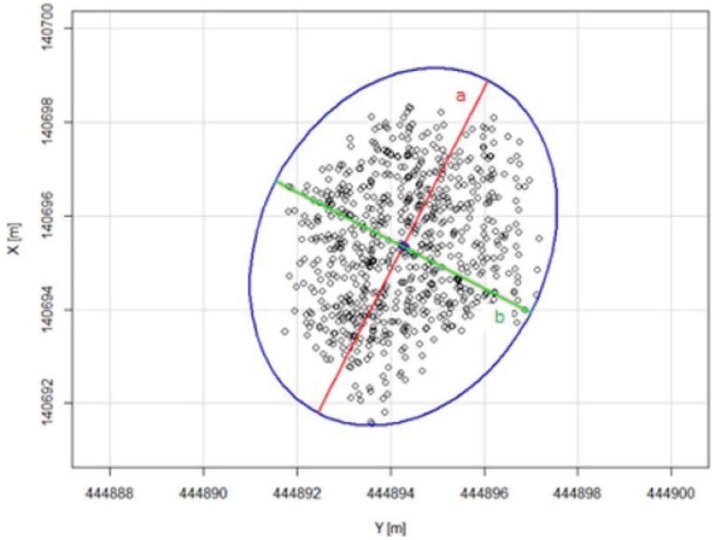


Figure 2: Outline of the ellipse around the tree-crown point cloud in the plan view and the narrowest and widest axes.

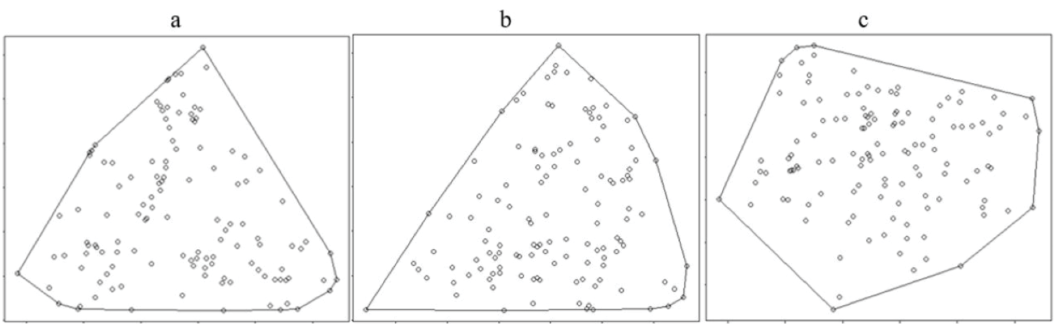


Figure 3: Plotting a convex hull for the Norway spruce tree-crown point cloud: a) narrow cross-section, b) wide cross-section, c) floor area.

For the successful discrimination between the values of the above-mentioned geometric parameters representing the individual trees or their treetop-crown point clouds, the original densities in the datasets with 5–15 points/m^2 were too low. Therefore, we increased the point-cloud density of each treetop point cloud by merging datasets 1, 2 and 3. The first three datasets were recorded in a 10-month period with the same spatial accuracy. We assumed that the size and shape of the trees did not change significantly over this time. Due to the larger period between the recordings of the first three datasets and dataset 4 (LSS), we did not use the latter to produce the merged individual treetop-canopy point clouds.

2.4 Statistical analysis

For a statistical analysis of the differences in the average intensities and geometric parameters of the treetop point clouds we used a simple analysis of variance (ANOVA). In the case of statically significant ($p < 0.05$) differences, we used Tukey's honestly significant difference test to analyse which of the mean values differ statistically significantly. The statistical analyses were performed with *RStudio*.

2.5 Building decision trees

The mean-intensity values and the mean values of the geometric parameters of the treetops were combined into a single classification model for tree-species classification. The model was built using the decision-tree method, a versatile machine-learning algorithm. We used *RStudio* to build the decision tree, enabling the decision steps to be displayed graphically and the results weighted by the probability that each result truly belongs to a particular category.

The decision-tree thresholds were generated from a classification tree. The algorithm first divides the initial subset of the training set into subsets that are as class pure as possible based on the attributes of the data. The process is repeated until the subsets are clean or until a cleaner subset is obtained by repeated breaking. The threshold setting was based on recursive binary splitting. In this study we used the Gini Index as a criterion for the binary splitting, which shows how much of the training-set data belongs to a particular class. If the index value is small, most of the data are part of one class.

Eight different decision trees were constructed based on analyses of the mean values of the intensities and analyses of the mean values of the geometric parameters of the treetop point clouds (representing the uppermost 3 m of the tree crown). In the first four we built a recognition model for dividing between a conifer and deciduous tree, in the other four a recognition model for classifying selected tree species. The input data were divided into training and test data, which were independent of each other. The training data contained 80% of the randomly selected trees (192 trees) and the test data the remaining 20% of the trees (48 trees). The training data were used for the basic tuning of the model parameters.

Based on statistical analyses of the geometric and intensity values, we determined the input parameters for the decision trees. Different decision trees were constructed from the following input parameters:

– the tree-crown volumes,
– the intensity values of datasets 2, 3 and 4,
– the tree-crown volumes and intensities of datasets 2, 3 and 4,
– the intensities of dataset 4, the freely available national laser scanning of Slovenia.

For the geometry differentiation we chose volume, as this data captures the average values of all the other geometry parameters being considered.

The decision-tree model was evaluated on the basis of its performance, defined as the proportion of correctly classified trees in the training or test set of the control data.

3 RESULTS

3.1 Comparison of tree species by geometry

The tree-species crowns tended to differ statistically significantly ($p < 0.05$) when we compared the following parameters: narrow cross-section width, wide cross-section width, area of the narrow cross-section, area of the wide cross-section, floor area and volume (Figure 4). However, the following pairs of tree species do not differ statistically significantly ($p > 0.05$) when comparing the same parameters: European larch – Norway spruce, European larch – Scots pine and sycamore – European beech.

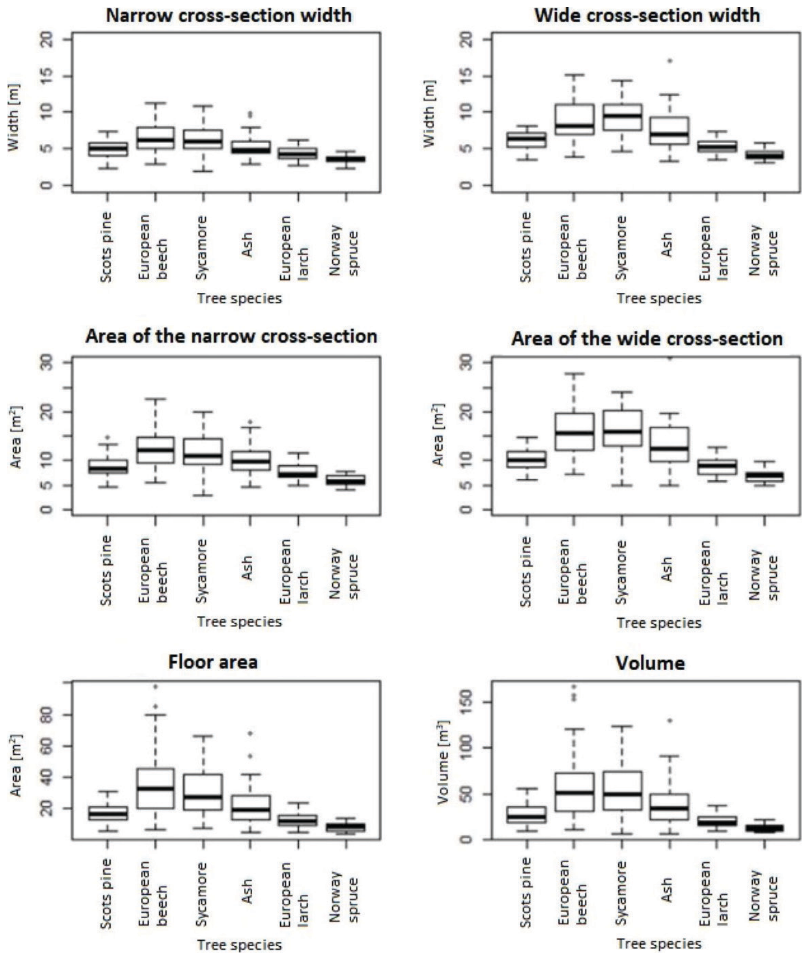


Figure 4: Comparison of tree-crown geometry by tree species.

3.2 Comparison of tree species using the intensity of laser points

The tree species' mean intensities differ statistically significantly ($p < 0.05$) when comparing the values between all the laser scanning datasets (Figure 5). The following tree-species-pair mean intensities were found to differ statistically significantly ($p < 0.05$): European beech – Scots pine, sycamore – Scots pine, European larch – Scots pine and Norway spruce – Scots pine. On the other hand, the statistically insignificant differences ($p > 0.05$) between the mean intensities of the following species pairs were found: ash – sycamore.

Comparing all three datasets taken during the leaf-on season (datasets 1, 2, 4), we observed lower intensities for the coniferous tree species, while the opposite is true during the leaf-off season (dataset 3).

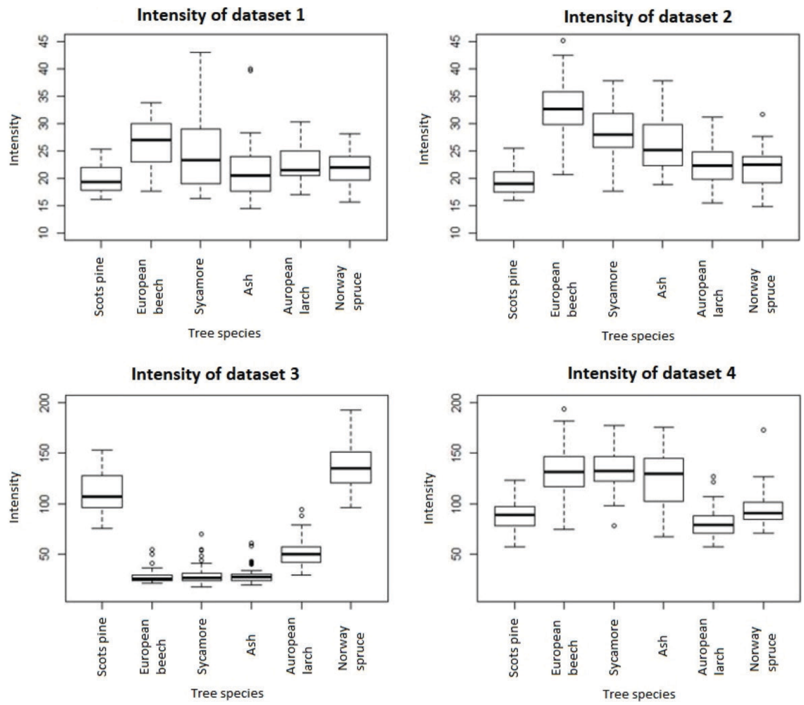


Figure 5: Comparison of tree-crown mean intensities presented by tree species and laser scanning dataset.

3.3 Decision trees for the prediction of coniferous and deciduous trees

Using a conifer/deciduous-prediction decision tree, where only the treetop crown's volume was considered, we successfully identified 79% of the trees of the training data (Table 2). In the test data we correctly identified 15 out of 21 conifers and 21 out of 27 deciduous trees. We misclassified 6 out of the 21 conifers, wrongly identified them as deciduous, and 6 of the 27 deciduous trees, which we misidentified as coniferous. Thus, we correctly classified 75% of the trees in the test data as conifers or deciduous.

Using a conifer/deciduous-prediction decision tree, where only the mean values of the intensities in datasets 2, 3 and 4 were considered, 93% of the trees in the training data were correctly identified being a coniferous or a deciduous tree. The success rate for the identification of coniferous and deciduous trees in the test data was 90%.

Using a prediction decision tree, where we used both the average values of tree crown volume and the average values of the intensities of datasets 2, 3 and 4 (Figure 6), we correctly identified 96% of the trees in the training data. The identification success rate of the conifers and deciduous trees in the test data was 95%.

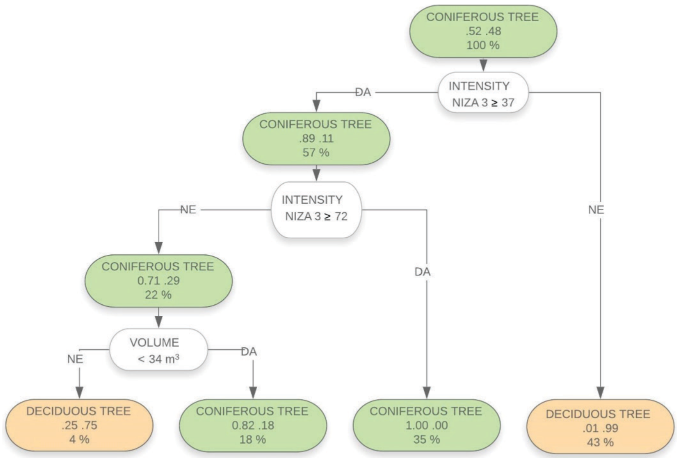


Figure 6: Decision tree for predicting conifers and deciduous trees from tree-crown volumes and average intensities of laser scanning datasets 2, 3 and 4.

Table 2: Results of conifer and deciduous-tree identification using different decision trees.

Training data				Test data			
	Coniferous	**Deciduous**	**Success rate**		**Coniferous**	**Deciduous**	**Success rate**
Volume							
Coniferous	**86**	13	**79%**	**Coniferous**	**15**	6	**75%**
Deciduous	27	**66**		**Deciduous**	6	**21**	
Intensity of datasets 2, 3 in 4							
Coniferous	**98**	1	**93%**	**Coniferous**	**19**	2	**90%**
Deciduous	12	**81**		**Deciduous**	3	**24**	
Volume and intensity in datasets 2, 3 in 4							
Coniferous	**96**	3	**96%**	**Coniferous**	**19**	2	**95%**
Deciduous	6	**87**		**Deciduous**	0	**27**	
Intensity of dataset 4 (LSS)							
Coniferous	**87**	12	**85%**	**Coniferous**	**15**	6	**81%**
Deciduous	16	**77**		**Deciduous**	3	**24**	

With the conifer and deciduous prediction decision tree, using only the average intensities of the laser scanning dataset 4 (LSS), we successfully classified 85% of the trees in the training data as coniferous or deciduous. The identification-success rate of the conifers and deciduous trees in the test data was 81%.

From the results it is clear that the least-successful decision tree on the test data is the prediction tree where only the treetop volume is considered (75% success rate). The decision tree where both the treetop volume and the intensity were used has the highest prediction-success rate (95%). The decision tree considering only the intensities of the laser scanning dataset 4 (LSS) had a success rate of 81% and the decision tree combining the intensities of the three datasets (2, 3, 4) had a success rate of 90%.

3.4 Decision trees for the prediction of separate tree species

Using a prediction decision tree where only the mean values of the volumes were considered, we correctly identified 48% of the tree species in the training data (Table 3). In the test data, we correctly identified 2 out of 5 Scots pines, 1 out of 13 European beeches, 2 out of 7 sycamores, 0 out of 8 ashes, 3 out of 8 European larches and 2 out of 7 Norway spruces. The identification-success rate of the tree species in the test data was 21%.

Using a decision tree where only the average intensities of sets 2, 3 and 4 were considered, we correctly identified 70% of the tree species in the training data. The identification-success rate of the tree species in the test data was 54%.

Using a decision tree, where we used the average treetop volumes and the average intensities of datasets 2, 3 and 4 (Figure 7), we correctly classified separate tree species for 74% of the trees in the training data. The identification-success rate of tree species in the test data was 60%.

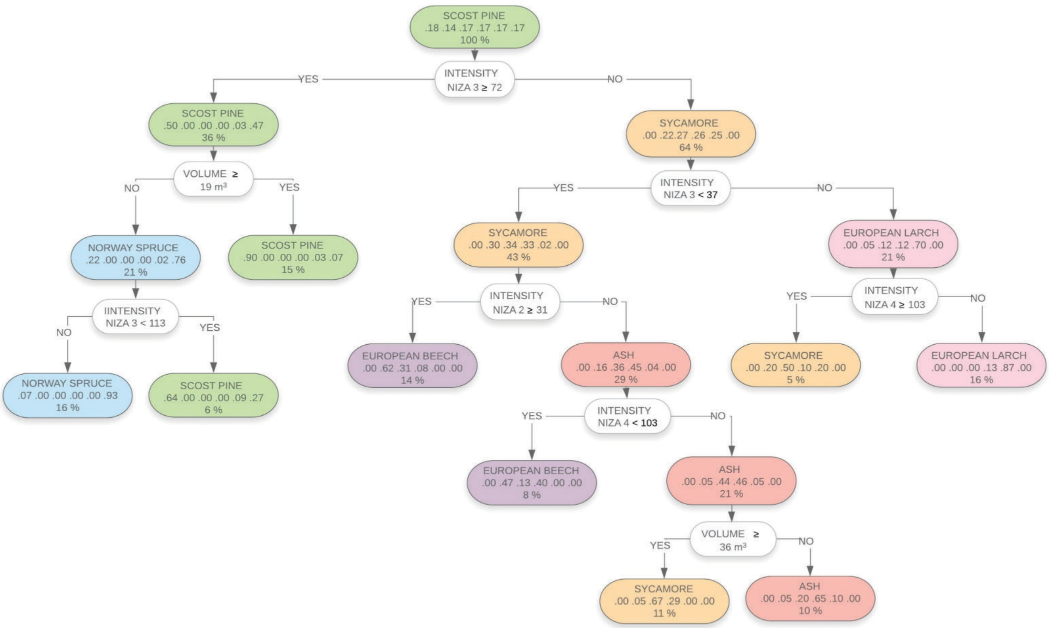


Figure 7: Decision tree for predicting tree species from treetop volumes and average intensities of laser scanning datasets 2, 3 and 4.

The decision tree where we used only the average intensities of dataset 4 (LSS) correctly classified the tree species with a 44% success rate in the training data. The identification-success rate of the tree species in the test data was 31%. The decision tree for tree-species separation based only on the geometry of the treetops is as ineffective as when trying to separate deciduous or coniferous trees, achieving an identification-success rate of only 21% in the test data. The decision tree using only the LSS intensities gives the second-worst results, with a 31% success rate for the test data. More than half of the tree species are correctly predicted by the following two decision trees: the first, which includes a combination of intensities from datasets 2, 3 and 4 (54%) and the second, which uses both volumes and a combination of those three intensities (60%).

Looking separately at the individual species and comparing the results of both the training and test data, we see that decision trees give poorer results when trying to identify ash trees. Most often, an ash tree was wrongly classified as a sycamore or a Scots pine. The European beech also has a poor identification-success rate: it cannot be distinguished from other tree species by the decision tree when only the intensities of dataset 4 (LSS) were used. The European beech identification is also as ineffective when using a decision tree where we compare only the mean volumes. Here, the European beech is mainly misclassified as a Scots pine or a sycamore.

Table 3: Tree species identification results using different decision trees.

	Training data								Test data						
	Scots pine	**European beech**	**Sycamore**	**Ash**	**European larch**	**Norway spruce**	**Success rate**		**Scots pine**	**European beech**	**Sycamore**	**Ash**	**European larch**	**Norway spruce**	**Success rate**
							Volume								
Scots pine	**22**	0	2	0	6	5	**48%**	**Scots pine**	**2**	1	0	0	1	1	**21%**
European beech	8	**14**	3	0	1	1		**European beech**	1	**1**	10	0	1	0	
Sycamore	8	4	**18**	0	2	1		**Sycamore**	2	3	**2**	0	0	0	
Ash	15	4	7	**0**	4	2		**Ash**	3	1	3	**0**	1	0	
European larch	14	0	0	0	**16**	2		**European larch**	3	0	0	0	**3**	2	
Norway spruce	1	0	0	0	9	**23**		**Norway spruce**	0	0	0	0	5	**2**	
							The intensity of datasets 2, 3 in 4								
Scots pine	**32**	0	0	0	0	3	**70%**	**Scots pine**	**5**	0	0	0	0	0	**54%**
European beech	0	**23**	3	1	0	0		**European beech**	0	**8**	3	2	0	0	
Sycamore	0	10	**19**	4	0	0		**Sycamore**	0	4	**1**	2	0	0	
Ash	0	8	10	**10**	4	0		**Ash**	0	4	3	**1**	0	0	
European larch	2	0	2	2	**26**	0		**European larch**	1	0	0	0	**7**	0	
Norway spruce	9	0	0	0	0	**24**		**Norway spruce**	3	0	0	0	0	**4**	
							Volume and intensity of datasets 2, 3 in 4								
Scots pine	**33**	0	0	0	0	2	**74%**	**Scots pine**	**5**	0	0	0	0	0	**60%**
European beech	0	**23**	3	1	0	0		**European beech**	0	**8**	4	1	0	0	
Sycamore	0	10	**19**	4	0	0		**Sycamore**	0	4	**2**	1	0	0	
Ash	0	8	7	**13**	4	0		**Ash**	0	4	2	**2**	0	0	
European larch	2	0	2	2	**26**	0		**European larch**	1	0	0	0	**7**	0	
Norway spruce	5	0	0	0	0	**28**		**Norway spruce**	2	0	0	0	0	**5**	
							Intensity of dataset 4 (LSS)								
Scots pine	**26**	0	1	2	2	4	**44%**	**Scots pine**	**5**	0	0	0	0	0	31%
European beech	6	**0**	17	2	2	0		**European beech**	1	**0**	10	1	0	1	
Sycamore	1	0	**26**	2	1	3		**Sycamore**	1	0	**5**	1	0	0	
Ash	6	0	16	**6**	3	1		**Ash**	0	0	7	**0**	0	1	
European larch	13	0	1	0	**15**	3		**European larch**	5	0	0	0	**3**	0	
Norway spruce	16	0	2	1	3	**11**		**Norway spruce**	4	0	0	1	0	**2**	

4 DISCUSSION AND CONCLUSION

The tree species considered in this research can be distinguished from each other based on comparisons of the tree crown's geometry parameters. Based on the mean treetop-volume changes only the following pairs of tree species cannot be statistically significantly ($p > 0.05$) distinguished: ash – Scots pine, European larch – Scots pine, sycamore – European beech and Norway spruce – European larch. Due to their growth physiology, deciduous trees have a larger floor area, crown width and treetop volume than coniferous trees.

The intensities obtained with a wavelength of 1064 nm range between 0 and 200, and with a wavelength of 1550 nm, they range between 0 and 500. Since the intensity is described as an integer, a larger range between the smallest and highest values has an advantage. In this way it is easier to distinguish between the individual tree species by comparing the laser scanning datasets of 3 and 4, which were taken at 1550 nm, since one was taken during the leaf-off season (dataset 3) and the other during the leaf-on season (dataset 4). Therefore, in those two datasets, we can clearly recognise the differences between the intensities of tree species in leaf-on and leaf-off seasons.

In terms of mean intensity, we statistically significantly ($p < 0.05$) distinguished the following pairs of tree species in all datasets: European beech – Scots pine, sycamore – Scots pine, European larch – European beech and Norway spruce – European beech. The pair ash – sycamore cannot be statistically significantly ($p < 0.05$) distinguished in any dataset. Thus, based on the mean intensities in datasets 2, 3 and 4, conifers can be distinguished by prediction deciduous trees. In dataset 1, we statistically significantly separate fewer tree species; therefore, we did not use this dataset for the decision-tree development.

Norway spruce, Scots pine and European larch can only be distinguished from each other in dataset 3, which was recorded in early March, i.e., in the leaf-off season. European larch is distinctly different from Norway spruce and Scots pine in terms of its intensities, as it is the only coniferous tree that bears no needles during the leaf-off season. European beech, sycamore and ash can only be distinguished from each other in dataset 2, which was taken in mid-September, when early autumn already appears at this altitude. As we can see from those examples, the laser datasets taken in the leaf-off season make it easier to distinguish between tree species. Shi et al. (2018) came to the same conclusions on a German example, where they found it easier to distinguish between beech, birch and spruce during the leaf-off season. Triglav Čekada et al. (2017) came to the same conclusion on a Slovenian example.

In the decision-tree development described in detail in the research of Kranjec (2020), she has considered the implementation of normalised intensities as well, but those did not give more useful results. Therefore, in this paper we only discuss the non-normalised intensities due to there being fewer processing steps.

The best-performing decision tree for conifer and deciduous tree separation is based on the simultaneous consideration of the treetop volumes and the overall mean values of the intensities from datasets 2, 3 and 4. The conifer and the deciduous separation-success rate is 96% for the training data and 95% for the test data. In total, 115 out of 120 conifers and 114 out of 120 deciduous trees were successfully classified in the training and test data. Similarly, the decision tree that simultaneously uses the treetop volumes and the total average intensities of the three datasets (2, 3 and 4) performs best in identifying different tree species. In the training data, the correct tree species were identified for 74% of the trees, and in the test data for 60% of the trees. In total, out of 40 trees of each tree species, we correctly identified the species of 38 Scots pines, 33 European larches, 33 Norway spruces, 31 European beeches, 21 sycamores and 15 ash trees if we combine the results from the training and test data. In general, the worst performing tree-species prediction was for ash, where all four prediction models gave equally poor results. European beech prediction performs even worse, but only in models where just the volumes or just the intensities of one dataset (LSS) were used.

All the decision trees perform better at identifying coniferous tree species, but less well at identifying deciduous tree species. Thus, among the conifers, in a few cases we could not distinguish between Norway spruce and Scots pine, while for deciduous trees we had difficulty to distinguish all the deciduous tree species between themselves.

The decision tree for conifer and deciduous tree identification, using only intensities from dataset 4, the freely available national laser scanning from Slovenia, achieved a success rate of 85% on the training data and 81% on the test data. In total, we correctly classified 102 out of 120 conifers and 101 out of 120 deciduous trees. These results can be used as a starting point for the separation of deciduous and coniferous tree mixtures at the level of forest stands anywhere in Slovenia. With additional calibration of the model by individual forest site-management classes that describe different growing conditions (Torresan et al., 2016), this will allow the methodology to be applied to forest-management units and forest-management areas in the future. The models performed less well in single-tree species identification, where we achieved 44% success on the training data and 31% on the test data. In total, out of 40 trees for each species, we correctly classified 31 Scots pines and 31 sycamores, 18 European larches, 13 Norway spruces, 6 ash trees and 0 European beeches.

When separating deciduous from coniferous trees and predicting the individual tree species, the best decision trees are those where we can apply multi-temporal laser scanning data, as these can describe the annual cycles on the trees, i.e., the changes in foliage between the leaf-on and leaf-off seasons. Therefore, when planning new laser scanning in areas where we already have one laser scanning dataset, it would be reasonable to consider the planning of a new laser scanning in the opposite foliage season to the one that we already have. This would give us a possibility to distinguish between the individual tree species with greater confidence in the future.

ACKNOWLEDGEMENTS

We would like to thank the Geodetic Institute of Slovenia for the data used in this study, which was commissioned within the framework of the cross-border project Slovenia-Austria 2011–2014: Natural Disasters without Borders (NH-WF). For the national laser scanning of Slovenia, we would like to thank the Ministry of Environment and Spatial Planning of the Republic of Slovenia.

Literature and references:

Benčina, A., Kobal, M. (2019). Uporaba laserskega skeniranja za vrednotenje poškodovanosti dreves zaradi žledoloma. Gozdarski vestnik, 77 (4), 159–169.

Budei, B. C., St-Onge, B., Hopkinson, C., Audet, F. (2018). Identifying the genus or species of individual trees using a three-wavelength airborne lidar system. Remote Sensing of Environment, 204, 632–647. DOI: https://doi.org/10.1016/j.rse.2017.09.037

Burt, A., Disney, M., Calders, K. (2018). Extracting individual trees from lidar point cloud using treseg. Methods in Ecology and Evolution, 10, 438–445. DOI: https://doi.org/10.1111/2041-210X.13121

Eysn, L., Hollaus, M., Lindberg, E., Berger, F., Monnet, J.-M., Dalponte, M., Kobal, M., Pellegrini, M. A., Lingua, E., Mongus, D., Pfeifer, N. (2015). A benchmark of lidar-based single tree detection methods using heterogeneous forest data from the alpine space. Forests, 6 (5), 1721–1747, DOI: https://doi.org/10.3390/f6051721

Holmgren, J., Persson, Å. (2004). Identifying species of individual trees using airborne laser scanner. Remote Sensing of Environment, 90 (4), 415–423. DOI: https://doi.org/10.1016/S0034-4257(03)00140-8

Kansanen, K., Vauhkonen, J., Lähivaara, T., Seppänen, A., Maltamo, M., Mehtätalo, L. (2019). Estimating forest stand density and structure using Bayesian individual tree detection, stochastic geometry, and distribution matching. ISPRS Journal of Photogrammetry and Remote Sensing, 152, 66–78. DOI: https://doi.org/10.1016/j.isprsjprs.2019.04.007

Karna, Y. K., Penman, T. D., Aponte, C., Hinko-Najera, N., Bennett, L. T. (2020). Persistent changes in the horizontal and vertical canopy structure of fire-tolerant forests after severe fire as quantified using multi-temporal airborne lidar data. Forest Ecology and Management, 472, 118255. DOI: https://doi.org/10.1016/j.foreco.2020.118255

Kobal, M., Triplat, M., Krajnc, N. (2014). Pregled uporabe zračnega laserskega skeniranja površja v gozdarstvu. Gozdarski vestnik, 72 (5-6), 235–248.

Korpela, I., Ørka, H. O., Maltamo, M., Tokola, T., Hyyppä, J. (2010). Tree Species Classification Using Airborne LiDAR – Effects of Stand and Tree Parameters, Downsizing of Training Set, Intensity Normalization, and Sensor Type. Silva Fennica, 44. DOI: https://doi.org/10.14214/sf.156

Korpela, I., Hovi, A., Korhonen, L. (2013). Backscattering of individual LiDAR pulses from forest canopies explained by photogrammetrically derived vegetation structure. ISPRS Journal of Photogrammetry and Remote Sensing, 83, 81–93. https://doi.org/10.1016/j.isprsjprs.2013.06.002

Kranjec, N. (2020). Preučevanje gozdnih sestojev in posameznih drevesnih vrst iz normalizirane intenzitete laserskih točk. Magistrsko delo. Ljubljana: Univerza v Ljubljani. Fakulteta za gradbeništvo in geodezijo.

Mongus, D., Vilhar, U., Skudnik, M., Žalik, B., Jesenko, D. (2018). Predictive analytics of tree growth based on complex networks of tree competition. Forest Ecology and Management, 425, 164–176. DOI: https://doi.org/10.1016/j.foreco.2018.05.039

Mongus, D., Žalik, B. (2015). An efficient approach to 3D single tree-crown delineation in LiDAR data. ISPRS journal of photogrammetry and remote sensing, 108, 219–233. DOI: https://doi.org/10.1016/j.isprsjprs.2015.08.004

Ørka, H. O., Næsset, E., Bollandsås, O. M. (2009). Classifying species of individual trees by intensity and structure features derived from airborne laser scanner data. Remote Sensing of Environment, 113, 1163–1174. DOI: https://doi.org/10.1016/j.rse.2009.02.002

Shi, Y., Wang, T., Skidmore, A. K., Heurich, M. (2018). Important LiDAR metrics for discriminating forest tree species in Central Europe. ISPRS Journal of Photogrammetry and Remote Sensing, 137, 163–174. DOI: https://doi.org/10.1016/j.isprsjprs.2018.02.002

Suranto, A., Seielstad, C., Queen, L. (2009). Tree species identification in mixed coniferous forest using airborne laser scanning. ISPRS Journal of Photogrammetry and Remote Sensing, 64, 683–693. DOI: https://doi.org/10.1016/j.isprsjprs.2009.07.001

Šturm, T., Pisek, R., Repnik, B., Matjašić, D. (2017). Pregledovalnik podatkov o gozdovih. Geodetski vestnik, 61 (1), 125–131.

Torresan, C., Corona, P., Scrinzi, G., Marsal, J. V. (2016). Using classification trees to predict forest structure types from LIDAR data. Annals of Forest Research, 59 (2), 281–298. DOI: https://doi.org/10.15287/afr.2016.423

Triglav Čekada, M., Lavrič, M., Kosmatin Fras, M. (2017). Ločevanje iglavcev in listavcev na podlagi neobdelane intenzitete laserskih točk. Geodetski vestnik, 61 (1), 23–34. DOI: https://doi.org/10.15292//geodetski-vestnik.2017.01.23-34

Triglav Čekada, M. (2011). Možnosti uporabe zračnega laserskega skeniranja (lidar) za geomorfološke študije. Geografski vestnik, 83 (2), 81–93.

Kranjec N., Triglav Čekada M., Kobal M. (2021). Predicting tree species based on the geometry and intensity of aerial laser scanning point cloud of treetops. Geodetski vestnik, 65 (2), 234-259.
DOI: https://doi.org/10.15292/geodetski-vestnik.2021.02.234-259

NAPOVEDOVANJE DREVESNIH VRST IZ GEOMETRIJE IN INTENZITETE OBLAKA AEROLASERSKIH TOČK VRHOV DREVESNIH KROŠENJ

OSNOVNE INFORMCIJE O ČLANKU:

GLEJ STRAN 234

1 UVOD

Aerolasersko skeniranje omogoča podrobno preučevanje gozda tako na krajinski ravni kot na ravni gozdnih sestojev in posameznih dreves, saj se laserski žarek večkrat odbije, in sicer od vrha drevesa, vej, debla in nazadnje od tal. Zato lahko na podlagi teh podatkov preučujemo ne le horizontalno, ampak tudi vertikalno strukturo gozda, kar omogoča oceno deleža iglastih in listnatih dreves, gostote dreves v sestoju, zdravstvenega stanja in poškodovanosti dreves (Kobal et al., 2014; Mongus et al., 2018; Benčina in Kobal, 2019; Karna et al., 2020). Lasersko skeniranje Slovenije v letih 2014–2015 je tako zelo uporaben vir podatkov tudi za preučevanje slovenskih gozdov (Šturm et al., 2016).

Laserske podatke lahko preučujemo zapisane v vektorski obliki, to je tako imenovano preučevanje oblaka točk, ter izvedeni rastrski obliki, kjer analiziramo razmerja med različnimi rastrskimi digitalnimi modeli višin (relief, površje, krošnje) oziroma drugimi rastrskimi tematskimi izvedenimi izdelki (na primer voksli), v katerih prikazujemo vrednosti izbranih atributov iz oblaka laserskih točk (na primer intenzitete).

Posamezna drevesa lahko v oblaku laserskih točk razmeroma uspešno ločimo z metodami, ki preučujejo lokalno geometrijo v oblaku točk, pri čemer razmejujemo drevesa na podlagi prepoznavanja njihovih vrhov, odkrivanja robov med krošnjami, iskanja senc in združevanja regij (Kobal et al., 2014; Mongus in Žalik, 2015; Eysn et al., 2016; Burt et al., 2018; Kansanen et al., 2019). Zahtevnejše in še ne popolnoma raziskano področje je prepoznavanje posameznih drevesnih vrst na podlagi samostojnih laserskih podatkov. Postopke za ločevanje drevesnih vrst lahko razdelimo na (Shi et al., 2018):

- metode, s katerimi se preučujejo geometrične lastnosti posamezne drevesne vrste (na primer oblika in gostota krošnje, razporejenost listov in vej, odprtost krošnje), ali
- metode, s katerimi se preučujejo radiometrične lastnosti laserskih odbojev (na primer intenziteta posameznega odboja).

Sama intenziteta laserskega odboja je odvisna od oblike listov ali iglic, njihove orientacije in velikosti, gostote v drevesni krošnji ter razporeditve drevesnih vej v krošnji (Holmgren in Person, 2004; Korpela et al., 2010 2013) ter od tipa laserskega skenerja in parametrov leta (Triglav Čekada, 2011).

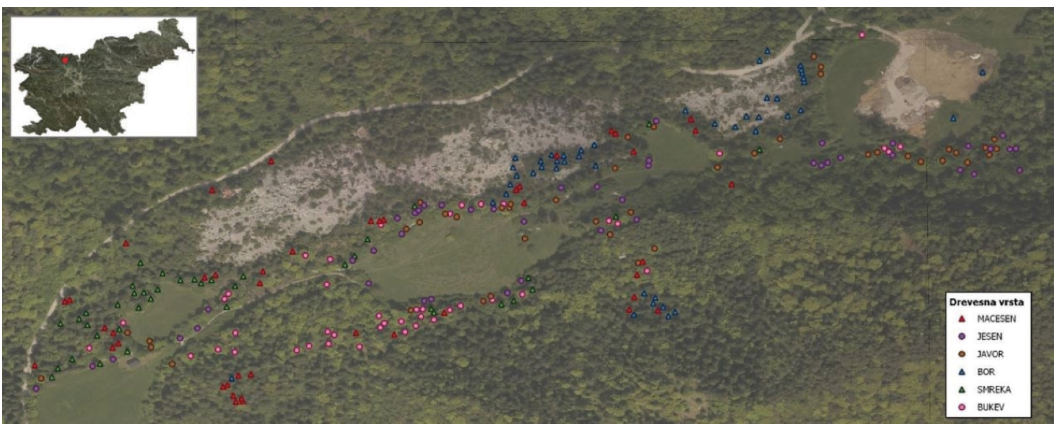
Pri nas so se s preučevanjem povprečnih intenzitet različnih drevesnih vrst v veččasovnih laserskih podatkih, posnetih z dvema valovnima dolžinama (1064 nm in 1550 nm) ukvarjale že Triglav Čekada et al. (2017). Ugotovile so, da so povprečne intenzitete listavcev v obeh valovnih dolžinah višje od iglavcev v času olistanosti, medtem ko je v času neolistanosti ravno nasprotno. Ko so ločile med sabo intenzitete prvih in edinih odbojev in intenzitete vseh odbojev, so ugotovile, da so slednje nižje v času olistanosti, medtem ko so v času neolistanosti višje od intenzitet prvih in edinih odbojev. Delež različnih redov odbojev (prvi, srednji, zadnji odboj) je odvisen od strukture krošnje posamezne preučevane vrste, torej od gostote iglic ali listja ter njihove razporeditve in razporeditve vej, kar so predhodno ugotovili že Ørka et al. (2009) in Suranto et al. (2009).

V naši raziskavi smo k preučevanju intenzitet odboja glede na različno drevesno vrsto in del leta, v katerem je bilo izvedeno lasersko skeniranje, dodali še preučevanje geometrije vrhnjih 3 metrov oblaka točk posameznega drevesa. Vrh krošnje smo si izbrali, ker se ta del drevesa večinoma ne prepleta s sosednjimi drevesi, prav tako so že Korpela et al. (2013) ugotovili, da je mogoče razlike v razporeditvi drevesnih vej glede na vrsto najlažje določiti na vrhu dreves. Z analizo povprečnih intenzitet oblakov točk posameznih dreves in analizo izbranih geometrijskih parametrov v različnih obdobjih leta smo razvili več napovedovalnih modelov na podlagi strojnega učenja – uporabili smo metodo odločitvenih dreves (angl. *decision tree*). Z uporabo veččasovnih laserskih podatkov smo razvili odločitvena drevesa za ločitev med iglavci in listavci ter za določitev posamezne drevesne vrste. Dodatno smo preverili uspešnost odločitvenega drevesa samo na enem nizu laserskih podatkov, in sicer na podatkih laserskega skeniranja Slovenije, saj bi tako lahko metodo prenesli iz testnega okolja na območje, ki pokriva celotno državo.

2 PODATKI IN METODOLOGIJA

2.1 Podatki

Testno območje stoji v bližini Podljubelja na nadmorski višini od 700 do 800 metrov in je poraščeno z mešanim gozdom. V raziskavi smo obravnavali šest drevesnih vrst, ki se na izbranem območju pojavljajo najpogosteje: navadno smreko (*Picea abies*), evropski macesen (*Larix decidua*), rdeči bor (*Pinus sylvestris*), navadno bukev (*Fagus sylvatica*), gorski javor (*Acer pseudoplatanus*) in veliki jesen (*Fraxinus excelsior*). Za vsako drevesno vrsto smo na terenu določili 40 dreves, ki so si čim bolj podobna po obliki krošnje in velikosti (slika 1).



Slika 1: Lokacije izbranih dreves (vir pregledne karte Slovenije: Geopedia, http://www.geopedia.si), vir podlage: ortofoto, posnet skupaj z laserskim skeniranjem niza 1).

Uporabili smo podatke štirih nizov aerolaserskih skeniranj (preglednica 1), ki so bili posneti v različnih letih in različnih letnih časih. Prva dva niza sta bila posneta z istim skenerjem Riegl LMS-Q560, druga dva z enakim tipom skenerja Riegl LMS-Q780. Prva dva niza sta bila posneta z valovno dolžino 1550 nm, druga dva z valovno dolžino 1064 nm. Trije nizi so bili posneti med olistanostjo ter en niz v času neolistanosti. Laserski nizi se med seboj razlikujejo tudi po gostoti laserskih točk in višini leta nad terenom. V analizi smo uporabili podatke, zapisane v pulznem načinu, originalni podatki pa so bili posneti s polnovalovnim laserskim skenerjem. Pri prvih dveh nizih so poleg aerolaserskega skeniranja izvedli še aerofotografiranje ter iz njih izdelali ortofote.

Preglednica 1: Lastnosti aerolaserskega skeniranja po nizih

	NIZ 1	**NIZ 2**	**NIZ 3**	**NIZ 4 (LSS)**
Datum snemanja	15. 5. 2012	18. 9. 2012	5. 3. 2013	17. 7. 2014
Gostota oblaka točk [točk/m²]	8	8	15	5
Valovna dolžina [nm]	1550	1550	1064	1064
Višina leta nad terenom [m]	700	700	1000	1200–1400
Olistanost	DA	DA	NE	DA
Ortofoto	DA	DA	NE	NE

2.2 Priprava podatkov aerolaserskega skeniranja

Pred terenskim snemanjem posameznih dreves smo v programu *ArcMap* (https://desktop.arcgis.com/en/arcmap/) prikazali ortofoto, ki je bil posnet istočasno kot laserski niz 1. Nanj smo dodali točke lokalnih maksimumov iz digitalnega modela krošenj (DMK), ki smo ga izdelali iz laserskih podatkov niza 1. DMK je razlika med modelom površja (DMP) in modelom reliefa (DMR). V njem lokalni maksimumi predstavljajo vrhove dominantnih dreves. Ortofoto z lokalnimi maksimumi smo izvozili kot georeferenciran TIFF in tega uvozili v mobilno aplikacijo *Avenza Maps* (https://www.avenza.com/avenza-maps/), ki omogoča določitev lokacije in navigacijo na predpripravljeni karti kar na pametnem telefonu. Na terenu smo vsakemu izbranemu drevesu določili lokacijo, vrsto in dodali fotografijo debla ter vrha krošnje. Na terenu smo izbirali drevesa, ki so se po višini za vsaj tri metre ločila od ostalih dreves v gozdu, oziroma se njihove krošnje v zgornjih metrih niso prepletale s krošnjami ostalih dreves.

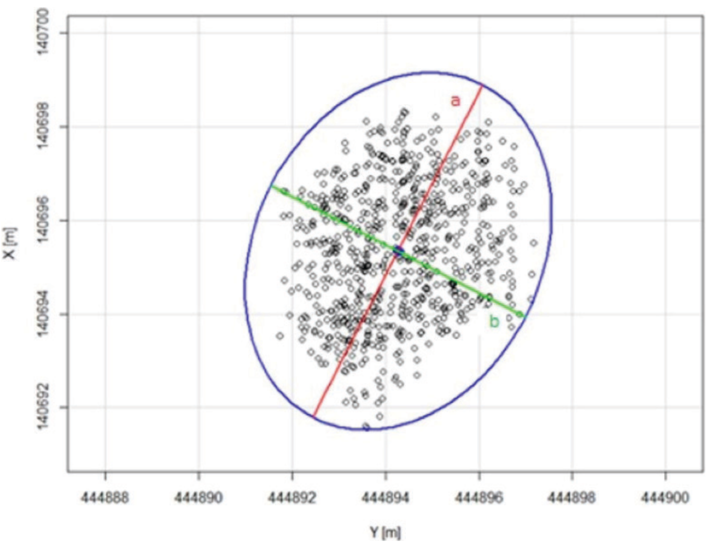
Terensko izmerjene lokacije dreves smo v programu *ArcMap* premaknili na točko lokalnega maksimuma, določenega na podlagi DMK, ki predstavlja na terenu izmerjeno drevo. Drevesa smo razdelili po drevesnih vrstah (slika 1). Vsakemu drevesu smo določili poligon v obliki kroga, s središčem v točki lokalnega maksimuma drevesa in polmerom pet metrov. Te poligone smo uporabili za obrez vseh nizov laserskih podatkov na manjše oblake točk, ki so prikazovali posamezno drevo. Obrez smo izvedli z orodjem *Extract LAS* (*ArcMap*). Da bi dobili le podatke o posameznem drevesu, brez prepletanja drevesnih krošenj z drugimi drevesi, smo uporabili le zgornje tri metre vrhov drevesnih krošenj. Same drevesne krošnje smo določili v programu *CloudCompare* (https://www.danielgm.net/cc/), obrez drevesne krošnje na le zgornje tri metre pa smo izdelali v programu *RStudio* (https://rstudio.com/), kjer smo izvedli tudi vse nadaljnje analize.

2.3 Določanje povprečnih vrednosti intenzitet in geometrije posamezne drevesne krošnje

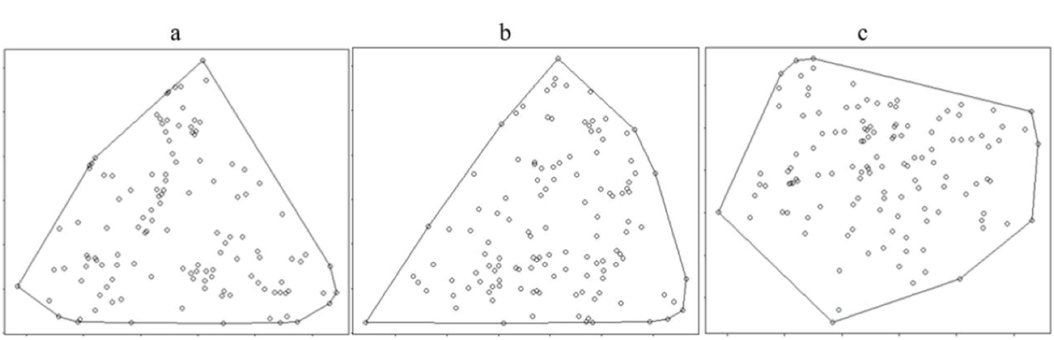
Iz oblaka točk vrha posamezne drevesne krošnje smo izračunali povprečno vrednost njegove intenzitete za vsak podatkovni niz posebej. Tu primerjamo le relativna razmerja v posameznem nizu, saj nenormaliziranih intenzitet, narejenih z različnimi valovnimi dolžinami, med sabo ne moremo neposredno primerjati.

Za vsak vrh drevesne krošnje smo izračunali naslednje geometrijske značilnosti: širino njegovega ožjega in širšega prereza, površino njegovega ožjega in širšega prereza ter površino tlorisa in volumen vrha krošnje. Širino ožjega in širino širšega prereza vrha krošnje smo določili tako, da smo laserskim točkam v pogledu tlorisa očrtali elipso ter določili njeno najkrajšo in najdaljšo polos (slika 2).

Površino ožjega in širšega prereza vrha krošnje smo določili z izračunom površine lika, ki ga predstavlja konveksna množica laserskih točk, preslikanih na vertikalno ravnino skozi prej določeno ožjo oziroma širšo os elipse (slika 3). Površino vrha krošnje v prerezu tlorisa smo določili z izračunom površine lika, ki ga predstavlja konveksna množica laserskih točk, preslikanih na horizontalno ravnino. Volumen vrha krošnje smo izračunali kot volumen telesa, ki ga tvori konveksna možica laserskih točk vrha krošnje.



Slika 2: Oris elipse laserskim točkam v tlorisu ter določitev ožje osi in širše osi elipse.



Slika 3: Izris konveksne množice na drevesni krošnji smreke: a) ožji prečni prerez, b) širši prečni prerez, c) prerez v tlorisu.

Za uspešno razlikovanje med vrednostmi navedenih geometrijskih parametrov med posameznimi drevesi oziroma njihovimi posameznimi oblaki točk so originalne gostote posameznih nizov od 5 do 15 točk/m^2 premalo. Zato smo gostoto v oblaku točk posameznega drevesa povečali tako, da smo združili točke nizov 1, 2 in 3. Snemanja prvih treh nizov so bila izvedena v razponu desetih mesecev z enako prostorsko točnostjo. Predpostavili smo, da se velikost in oblika dreves v tem času nista bistveno spremenila. Zaradi večje časovne razlike med snemanjem prvih treh nizov in nizom 4 (LSS) slednjih nismo uporabili za pripravo združenih oblakov točk.

2.4 Statistična analiza

Za statistično analizo razlik v intenziteti odbojev laserskih točk in geometriji vrha krošenj smo uporabili enostavno analizo variance ANOVA, pri statično značilnih ($p < 0,05$) razlikah pa Tukeyjev test mnogoterih primerjav (angl. *Tukey's honestly significant difference test)*, kjer smo preverili, med katerimi skupinami se povprečne vrednosti med seboj statistično značilno razlikujejo. Statistično analizo smo izvedli v programu *RStudio*.

2.5 Izdelava odločitvenih dreves

Povprečne vrednosti intenzitet in parametre geometrije drevesnih krošenj smo združili v enoten model prepoznavanja drevesnih vrst. Model smo izdelali z metodo odločitvenih dreves, vsestranskim algoritmom strojnega učenja. Za izdelavo odločitvenega drevesa smo uporabili program *RStudio*, kjer lahko korake odločanja prikažemo grafično, rezultate pa ovrednotimo z verjetnostjo, da posamezen rezultat resnično spada v določeno kategorijo.

Mejne vrednosti odločitvenega drevesa so bile izdelane na podlagi klasifikacijskega drevesa (angl. *classification tree*). Algoritem v osnovi najprej začetno podmnožico učnega niza razdeli glede na atribute podatkov na razredno čim bolj čiste podmnožice. Postopek se ponavlja, dokler podmnožice niso čiste oziroma s ponovnim razbitjem ne pridobimo čistejših podmnožic. Določitev mejnih vrednosti je bila izdelana na podlagi binarnega ločevanja (angl. *recursive binary splitting*). V članku smo kot kriterij pri binarnem ločevanju uporabili Ginijev indeks (angl. *Gini index*), ki prikaže, koliko podatkov učnega niza pripada nekemu razredu. Če je vrednost indeksa majhna, je večina podatkov del enega razreda.

Na podlagi analiz povprečnih vrednosti intenzitet in analiz povprečnih vrednosti geometrij zgornjih treh metrov vrhov drevesnih krošenj smo izdelali osem različnih odločitvenih dreves. V prvih štirih smo izdelali model prepoznavanja iglavcev in listavcev, v drugih štirih model prepoznavanja izbranih drevesnih vrst. Vhodne podatke smo delili na učne in testne podatke, ki so bili med seboj neodvisni. Učni podatki so vsebovali 80 % naključno izbranih dreves (192 dreves), testni podatki pa preostalih 20 % dreves (48 dreves). Učne podatke smo uporabili za osnovno nastavitev parametrov modela.

Na podlagi statističnih analiz geometrijskih vrednosti in vrednosti intenzitet smo določili vhodne parametre za odločitvena drevesa. Različna odločitvena drevesa so bila izdelana iz naslednjih vhodnih parametrov:

– volumnov drevesnih krošenj,
– intenzitet niza 2, 3 in 4,

– volumnov drevesnih krošenj in intenzitet niza 2, 3 in 4 ter
– intenzitet niza 4, prosto dostopnih podatkov laserskega skeniranja Slovenije.

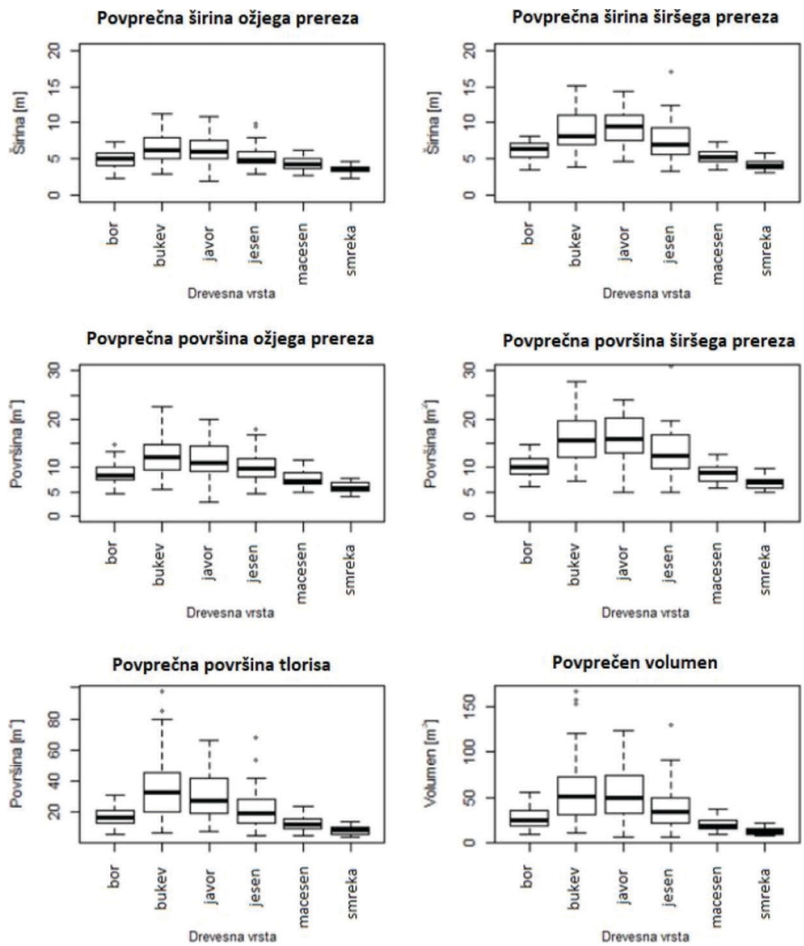
Pri razlikovanju geometrije smo se odločili za volumen, saj ta podatek zajema povprečne vrednosti vseh ostalih obravnavanih parametrov geometrijskih parametrov.

Model odločitvenega drevesa smo ocenili z uspešnostjo, ki je definirana kot delež pravilno klasificiranih dreves učnega oziroma testnega niza.

3 REZULTATI

3.1 Primerjava drevesnih vrst po geometriji

Statistično značilno ($p < 0{,}005$) se med seboj razlikujejo vrhovi krošenj drevesnih vrst po ožji širini prereza, širši širini prereza, površini ožjega prereza, površini širšega prereza, površini tlorisa in volumnu (*slika 4*), vendar se po geometriji vrha krošenj statistično značilno ne razlikujejo ($p > 0.05$) naslednji pari drevesnih vrst: navadna smreka – evropski macesen, evropski macesen – rdeči bor in gorski javor – navadna bukev.

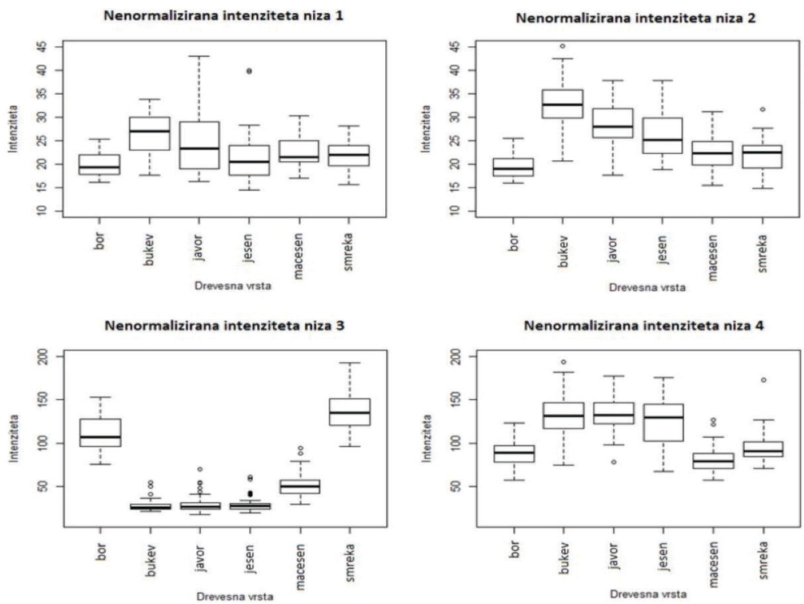


Slika 4: Primerjava geometrije drevesnih krošenj po drevesnih vrstah.

3.2 Primerjava drevesnih vrst po intenziteti

Statistično značilno ($p < 0{,}05$) se po drevesnih vrstah razlikujejo tudi povprečne vrednostih intenzitet v vseh nizih (slika 5), med seboj pa se statistično značilno razlikujejo ($p < 0{,}05$) naslednji pari drevesnih vrst: navadna bukev – rdeči bor, gorski javor – rdeči bor, evropski macesen – navadna bukev in navadna smreka – navadna bukev. Statistično značilno ($p > 0{,}05$) v nobenem nizu ne moremo razlikovati povprečnih vrednostih intenzitet para veliki jesen – gorski javor.

Če primerjamo vse tri nize iz časa olistanosti (nizi 1, 2, 4), opazimo nižje vrednosti intenzitet pri iglavcih, v času neolistanosti (niz 3) pa ravno nasprotno.



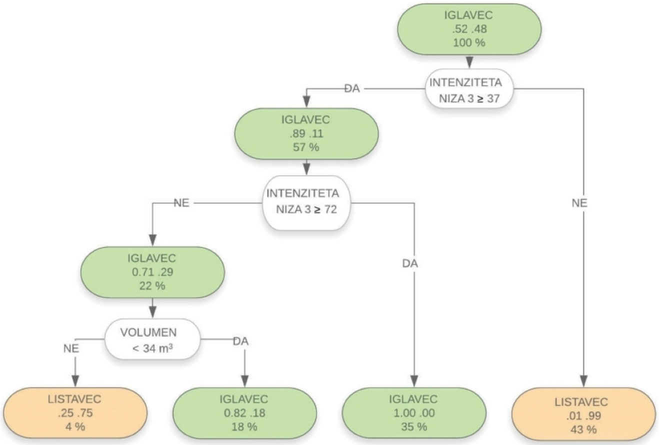
Slika 5: Primerjava povprečnih vrednosti intenzitet drevesnih krošenj po drevesnih vrstah v posameznih nizih.

3.3 Odločitveno drevo napovedovanja iglavcev oziroma listavcev

Z odločitvenim drevesom napovedovanja iglavcev oziroma listavcev, v katerem smo upoštevali samo volumen vrha krošnje, smo v učnih podatkih uspešno določili 79 % dreves (preglednica 2). V testnih podatkih smo pravilno prepoznali 15 od 21 iglavcev in 21 od 27 listavcev. Napačno smo razvrstili 6 od 21 iglavcev in jih prepoznali kot listavce ter 6 od 27 listnatih dreves, ki smo jih napačno prepoznali kot iglasta. Tako smo pravilno med iglavce in listavce v testnih podatkih razvrstili 75 % dreves.

Z odločitvenim drevesom napovedovanja iglavcev oziroma listavcev, v katerem smo upoštevali le povprečne vrednosti intenzitet niza 2, 3 in 4, smo v učnih podatkih pravilno določili, ali je drevo iglasto ali listnato, 93 % dreves. Uspešnost prepoznavanja iglavcev in listavcev v testnih podatkih je 90 %.

Z odločitvenim drevesom napovedovanja, kjer smo uporabili tako povprečne vrednosti volumna krošnje kot povprečne vrednosti intenzitet niza 2, 3 in 4 (slika 6), smo v učnih podatkih pravilno prepoznali 96 % dreves. Uspešnost prepoznavanja iglavcev in listavcev v testnih podatkih je 95 %.



Slika 6: Odločitveno drevo napovedovanja iglavcev oziroma listavcev iz vrednosti volumnov in povprečnih vrednosti intenzitet niza 2, 3 in 4.

Preglednica 2: Rezultati prepoznavanja iglavcev in listavcev z odločitvenim drevesom

Učni podatki				Testni podatki			
	iglavec	**listavec**	**uspešnost**		**iglavec**	**listavec**	**uspešnost**
volumen							
iglavec	**86**	13	**79 %**	**iglavec**	**15**	6	**75 %**
listavec	27	**66**		**listavec**	6	**21**	
intenziteta niza 2, 3 in 4							
iglavec	**98**	1	**93 %**	**iglavec**	**19**	2	**90 %**
listavec	12	**81**		**listavec**	3	**24**	
volumen in intenziteta niza 2, 3 in 4							
iglavec	**96**	3	**96 %**	**iglavec**	**19**	2	**95 %**
listavec	6	**87**		**listavec**	0	**27**	
intenziteta niza 4 (LSS)							
iglavec	**87**	12	**85 %**	**iglavec**	**15**	6	**81 %**
listavec	16	**77**		**listavec**	3	**24**	

Z odločitvenim drevesom napovedovanja iglavcev oziroma listavcev, kjer smo uporabili samo povprečne intenzitete niza 4 (LSS), smo v učnih podatkih uspešno razvrstili med iglavce in listavce 85 % dreves. Uspešnost prepoznavanja iglavcev in listavcev v testnih podatkih je 81 %.

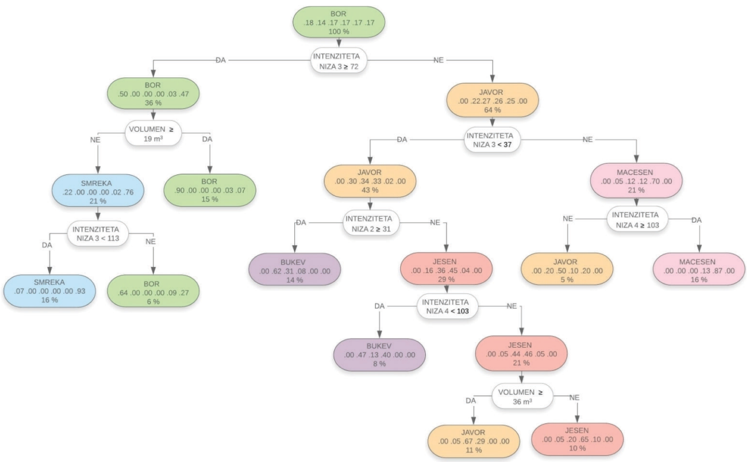
Razvidno je, da je na testnih podatkih najmanj uspešno odločitveno drevo napovedovanja, kjer upoštevamo samo volumen vrha krošenj (uspešnost 75 %). Najvišjo uspešnost (95 %) napovedovanja ima odločitveno drevo, v katerem upoštevamo tako volumen kot intenziteto. Odločitveno drevo, kjer upoštevamo samo intenzitete niza 4 (LSS), je uspešno v 81 %, odločitveno drevo s kombinacijo intenzitet treh nizov pa v 90 %.

3.4 Odločitvena drevesa napovedovanja drevesnih vrst

Z odločitvenim drevesom napovedovanja, kjer upoštevano samo povprečne vrednosti volumnov, smo v učnih podatkih pravilno prepoznali 48 % drevesnih vrst (preglednica 3). V testnem nizu smo pravilno prepoznali 2 od 5 rdečih borov, 1 od 13 navadnih bukev, 2 od 7 gorskih javorjev, 0 od 8 velikih jesenov, 3 od 8 evropskih macesnov in 2 od 7 navadnih smrek. Uspešnost prepoznavanja drevesnih vrst v testnih podatkih je 21 %.

Z odločitvenim drevesom napovedovanja, kjer upoštevamo le povprečne vrednosti intenzitete nizov 2, 3 in 4, smo v učnih podatkih pravilno prepoznali 70 % drevesnih vrst. Uspešnost prepoznavanja drevesnih vrst v testnih podatkih je 54 %.

Z odločitvenim drevesom napovedovanja, kjer smo uporabili povprečne vrednosti volumna in povprečne vrednosti intenzitet nizov 2, 3 in 4 (slika 7), smo v učnih podatkih pravilno po drevesnih vrstah razvrstili 74 % dreves. Uspešnost prepoznavanja drevesnih vrst v testnih podatkih je 60 %.



Slika 7: Odločitveno drevo napovedovanja drevesnih vrst iz vrednosti volumnov in povprečnih vrednosti intenzitet niza 2, 3 in 4.

Z odločitvenim drevesom napovedovanja, pri katerem smo uporabili samo povprečne intenzitete niza 4 (LSS), smo v učnih podatkih pravilno med drevesne vrste razvrstili 44 % dreves. Uspešnost prepoznavanja drevesnih vrst v testnih podatkih je 31 %. Enako kot pri napovedovanju, ali je drevo listavec ali iglavec, so tudi pri napovedi drevesne vrste najslabši rezultati napovedovanja, ki temeljijo samo na geometriji vrhov dreves, saj uspešnost na testnih podatkih znaša 21 %. Drugo najslabše odločitveno drevo je tisto, ki uporabi samo intenzitete LSS, in sicer z uspešnostjo 31 %. Več kot polovico drevesnih vrst pravilno napovesta odločitveno drevo, ki vključuje kombinacijo treh intenzitet (54 %), ter odločitveno drevo, ki uporablja tako volumne kot kombinacijo treh intenzitet (60 %).

Če pogledamo ločeno po posameznih vrstah ter primerjamo rezultate, tako na učnih kot testnih podatkih, vidimo, da odločitvena drevesa dajejo slabše rezultate pri prepoznavanju velikega jesena, saj je največkrat uvrščen med gorski javor ali rdeči bor. Sledi navadna bukev, ki jo z odločitvenim drevesom, kjer uporabimo samo intenzitete niza 4 (LSS), ne moremo ločiti od drugih drevesnih vrst. Slabše rezultate napovedovanja navadne bukve dobimo tudi pri odločitvenem drevesu, ki uporabi samo volumne. Največkrat navadno bukev odločitveno drevo napačno razvrsti med rdeči bor ali gorski javor.

Preglednica 3: Rezultati prepoznavanja drevesnih vrst z odločitvenim drevesom.

	Učni podatki								**Testni podatki**						
	rdeči bor	**navadna bukev**	**gorski javor**	**veliki jesen**	**evropski macesen**	**navadna smreka**	**uspešnost**		**rdeči bor**	**navadna bukev**	**gorski javor**	**veliki jesen**	**evropski macesen**	**navadna smreka**	**uspešnost**
volumen															
rdeči bor	**22**	0	2	0	6	5	**48 %**	**rdeči bor**	**2**	1	0	0	1	1	**21 %**
navadna bukev	8	**14**	3	0	1	1		**navadna bukev**	1	**1**	10	0	1	0	
gorski javor	8	4	**18**	0	2	1		**gorski javor**	2	3	**2**	0	0	0	
veliki jesen	15	4	7	**0**	4	2		**veliki jesen**	3	1	3	**0**	1	0	
evropski macesen	14	0	0	0	**16**	2		**evropski macesen**	3	0	0	0	**3**	2	
navadna smreka	1	0	0	0	9	**23**		**navadna smreka**	0	0	0	0	5	**2**	
intenziteta niza 2, 3 in 4															
rdeči bor	**32**	0	0	0	0	3	**70 %**	**rdeči bor**	**5**	0	0	0	0	0	**54 %**
navadna bukev	0	**23**	3	1	0	0		**navadna bukev**	0	**8**	3	2	0	0	
gorski javor	0	10	**19**	4	0	0		**gorski javor**	0	4	**1**	2	0	0	
veliki jesen	0	8	10	**10**	4	0		**veliki jesen**	0	4	3	**1**	0	0	
evropski macesen	2	0	2	2	**26**	0		**evropski macesen**	1	0	0	0	**7**	0	
navadna smreka	9	0	0	0	0	**24**		**navadna smreka**	3	0	0	0	0	**4**	
volumen in intenziteta niza 2, 3 in 4															
rdeči bor	**33**	0	0	0	0	2	**74 %**	**rdeči bor**	**5**	0	0	0	0	0	**60 %**
navadna bukev	0	**23**	3	1	0	0		**navadna bukev**	0	**8**	4	1	0	0	
gorski javor	0	10	**19**	4	0	0		**gorski javor**	0	4	**2**	1	0	0	
veliki jesen	0	8	7	**13**	4	0		**veliki jesen**	0	4	2	**2**	0	0	
evropski macesen	2	0	2	2	**26**	0		**evropski macesen**	1	0	0	0	**7**	0	
navadna smreka	5	0	0	0	0	**28**		**navadna smreka**	2	0	0	0	0	**5**	
intenziteta niza 4 (LSS)															
rdeči bor	**26**	0	1	2	2	4	**44 %**	**rdeči bor**	**5**	0	0	0	0	0	**31 %**
navadna bukev	6	**0**	17	2	2	0		**navadna bukev**	1	**0**	10	1	0	1	
gorski javor	1	0	**26**	2	1	3		**gorski javor**	1	0	**5**	1	0	0	
veliki jesen	6	0	16	**6**	3	1		**veliki jesen**	0	0	7	**0**	0	1	
evropski macesen	13	0	1	0	**15**	3		**evropski macesen**	5	0	0	0	**3**	0	
navadna smreka	16	0	2	1	3	**11**		**navadna smreka**	4	0	0	1	0	**2**	

4 RAZPRAVA IN SKLEP

Obravnavane drevesne vrste lahko med seboj ločimo s primerjavami geometrije vrhov drevesnih krošenj, vendar statistično značilno (p > 0,05) iz povprečnih vrednosti volumnov ne moremo ločiti naslednjih parov drevesnih vrst: veliki jesen – rdeči bor, evropski macesen – rdeči bor, gorski javor – navadna bukev in navadna smreka – evropski macesen. Zaradi fiziologije rasti listnata drevesa dosegajo večjo površino prereza, širino krošnje in volumen vrha drevesne krošnje kot iglavci.

Vrednosti intenzitet, pridobljene z valovno dolžino 1064 nm, znašajo med 0 in 200, z valovno dolžino 1550 nm pa med 0 in 500. Ker se intenziteta beleži kot celo število, je večji razpon med vrednostmi prednost. Zato lažje ločimo posamezne drevesne vrste, če primerjamo niza 3 in 4, ki sta bila posneta z valovno dolžino 1550 nm. Ker pa je eden posnet v času neolistanosti (niz 3) in drugi v času olistanosti (niz 4), na njiju lahko lepo prepoznamo tudi to razliko.

Glede na povprečne intenzitete lahko statistično značilno (p < 0,05) v vseh nizih ločimo pare navadna bukev – rdeči bor, gorski javor – rdeči bor, evropski macesen – navadna bukev in navadna smreka – navadna bukev. Statistično značilno (p < 0,05) v nobenem nizu ne moremo razlikovati povprečnih vrednostih intenzitet para veliki jesen – gorski javor. Tako iz povprečnih vrednosti intenzitet v nizih 2, 3 in 4 med seboj lahko ločimo iglavce od listavcev. V nizu 1 med seboj statistično ločimo najmanj drevesnih vrst, zato povprečnih vrednosti intenzitet iz niza 1 v odločitvenih drevesih nismo uporabili.

Smreko, bor in macesen lahko med seboj ločimo le v nizu 3, ki je bil posnet začetek marca, torej v neolistanem delu leta. Evropski macesen se pri tem po vrednostih intenzitete izrazito loči od navadne smreke in rdečega bora, saj v času neolistanosti med iglavci edini nima iglic. Navadna bukev, gorski javor, veliki jesen lahko med seboj ločimo le v nizu 2, ki je bil posnet sredi septembra, ko se na tej nadmorski višini že pojavi zgodnja jesen. Tako drevesne vrste lažje med seboj ločimo v času neolistanosti. Do enakih ugotovitev so v Nemčiji prišli Shi et al. (2018), ki so med seboj lažje ločili bukev, brezo in smreko v času neolistanosti, pri nas pa Triglav Čekada et al. (2017).

Opazimo, da imajo v času olistanosti nižje vrednosti intenzitet iglavci, v času neolistanosti pa je ravno nasprotno, kar so ugotovili že Holmgren in Persson (2004), Ørka et al. (2009), Triglav Čekada et al. (2017), Shi et al. (2018). V času olistanosti, v nizu 2, ima med listavci najvišjo vrednost intenzitete navadna bukev, sledi gorski javor, najnižjo povprečno vrednost ima veliki jesen. V času neolistanosti, v nizu 3, ima med iglavci najvišjo vrednost intenzitete navadna smreka, sledi rdeči bor, najnižjo, podobno listavcem, ima evropski macesen. Da ima smreka višjo povprečno intenziteto v času neolistanosti, sta v raziskavi na Švedskem ugotovila tudi Holmgren in Persson (2004), pri nas pa Triglav Čekada et al. (2017).

V raziskavi smo se ukvarjali tudi z možnostjo uporabe normaliziranih intenzitet (Kranjec, 2020), ki pa niso dale uporabnejših rezultatov, zato zaradi manj korakov obdelave v članku uporabljamo le nenormalizirane intenzitete.

Najuspešnejše odločitveno drevo prepoznavanja iglavcev in listavcev temelji na hkratnem upoštevanju volumnov in skupnih povprečnih vrednosti intenzitet iz nizov 2, 3 in 4. Uspešnost prepoznavanja iglavcev in listavcev v učnih podatkih znaša 96 %, v testnih pa 95 %. Skupaj smo v učnih in testnih podatkih uspešno pojasnili 115 od 120 iglavcev in 114 od 120 listavcev. Enako je odločitveno drevo napovedovanja, ki hkrati uporablja volumne in povprečne skupne vrednosti intenzitet treh nizov (2, 3 in 4), najuspešnejše pri razpoznavanju posamezne drevesne vrste. V učnih podatkih smo pravilno drevesno vrsto določili 74 % drevesom, v testnih podatkih 60 % drevesom. Skupaj smo v učnih in testnih podatkih od 40 dreves posamezne drevesne vrste pravilno določili vrsto 38 rdečim borom, 33 evropskim macesnom, 33 navadnim smrekam, 31 navadnim bukvam, 21 gorskim javorjem in 15 velikim jesenom. V splošnem se najslabše obnese napovedovanje velikega jesena, kjer dajo vsi štirje napovedovalni modeli enako slabe rezultate. Slabše se obnese še napovedovanje navadne bukve, vendar le v modelih, kjer uporabimo samo volumne ali samo intenzitete enega niza (laserskega skeniranja Slovenija).

Vsa odločitvena drevesa dosegajo večjo uspešnost pri prepoznavanju drevesnih vrst med iglavci, manj uspešni so pri prepoznavanju drevesnih vrst listavcev. Tako med iglavci v nekaj primerih nismo ločili med navadno smreko in rdečim borom, pri listavcih pa smo imeli težave z medsebojnim ločevanjem vseh drevesnih vrst.

Z odločitvenim drevesom prepoznavanja iglavcev in listavcev, kjer smo uporabili samo intenzitete iz prosto dostopnih podatkov laserskega skeniranja Slovenije, smo na učnih podatkih dosegli uspešnost 85 %, na testnih pa 81 %. Skupaj smo pravilno razporedili 102 od 120 iglavcev ter 101 od 120 listavcev. Ti rezultati odločitvenega drevesa so zagotovo lahko izhodišče za ločevanje mešanosti oziroma listavcev in iglavcev na ravni gozdnih sestojev kjerkoli v Sloveniji, kar ob dodatni kalibraciji modela po posameznih rastiščno-gojitvenih razredih, ki upoštevajo različne rastiščne razmere (Torresan et al., 2016), omogoča uporabo metodologije na ravni gozdnogospodarskih enot in gozdnogospodarskih območij. Manj uspešni so se modeli izkazali pri prepoznavanju posamezne drevesne vrste, kjer smo v učnih podatkih dosegli uspešnost 44 %, v testnih pa 31 %. Skupaj smo od 40 dreves posamezne drevesne vrste pravilno razvrstili 31 rdečih borov in 31 gorskih javorjev, 18 evropskih macesnov, 13 navadnih smrek, 6 velikih jesenov in 0 navadnih bukev.

Pri ločevanju listavcev od iglavcev in pri napovedovanju posameznih drevesnih vrst so najboljša odločitvena drevesa napovedovanja, v katerih obravnavamo veččasovne laserske podatke, saj le-ti lahko opišejo največje letne cikle na drevesih, to je spremembe med olistanostjo in neolistanostjo. Zato bi bilo pri načrtovanjih novih laserskih skeniranj na območjih, kjer že imamo en laserski podatkovni niz, smiselno razmisliti o podrobnem časovnem načrtovanju novega laserskega skeniranja, da zajamemo sezono olistanosti/neolistanosti, ki je še nimamo skenirane. Tako bomo omogočili večjo zanesljivost določitve posameznih drevesnih vrst v prihodnjih analizah.

ZAHVALA

Za podatke, ki smo jih uporabili v raziskavi, se najlepše zahvaljujemo Geodetskemu inštitutu Slovenije, ki jih je naročil v okviru čezmejnega projekta Slovenija-Avstrija 2011–2014: *Naravne nesreče brez meja (NH-WF)*. Za podatke laserskega skeniranja Slovenije se zahvaljujemo Ministrstvu za okolje in prostor Republike Slovenije.

Literatura in viri:

Glej stran 246

Kranjec N., Triglav Čekada M., Kobal M. (2021). Napovedovanje drevesnih vrst iz geometrije in intenzitete oblaka aerolaserskih točk vrhov drevesnih krošenj. Predicting tree species based on the geometry and intensity of aerial laser scanning point cloud of treetops. Geodetski vestnik, 65 (2), 234-259. DOI: https://doi.org/10.15292/geodetski-vestnik.2021.02.234-259

Nina Kranjec, mag. inž. geod. geoinf.
Dobrovita d.o.o.
Tbilisijska ulica 87, SI-1000 Ljubljana, Slovenija
e-naslov: nina@dobrovita.com

doc. dr. Mihaela Triglav Čekada, univ. dipl. inž. geod.
Geodetski inštitut Slovenije in
Univerza v Ljubljani, Fakulteta za gradbeništvo in geodezijo,
Jamova cesta 2, SI-1000 Ljubljana, Slovenija
e-naslov: mihaela.triglav@gis.si

izr. prof. dr. Milan Kobal, univ. dipl. inž. gozd.
Univerza v Ljubljani, Biotehniška fakulteta,
Oddelek za gozdarstvo in obnovljive gozdne vire
Večna pot 83, SI-1000 Ljubljana, Slovenija
e-naslov: milan.kobal@bf.uni-lj.si