

PLANNING HARVESTING OPERATIONS IN FOREST ENVIRONMENT: REMOTE SENSING FOR DECISION SUPPORT

M. Piragnolo^{1,2}, S. Grigolato¹, F. Pirotti^{1,2*}

¹ TESAF Department , University of Padova, Via dell'Università, 16, 35020 Legnaro (PD), Italy

² CIRGEO – Interdepartmental Research Center in Geomatics - University of Padova, Via dell'Università, 16, 35020 Legnaro (PD), Italy (marco.piragnolo, stefano.grigolato, francesco.pirotti)@unipd.it

Commission III, WG III/10

KEY WORDS: Machine learning, forest harvesting, skyline vs. forwarder, digital terrain model, morphology, classification

ABSTRACT:

The goal of this work is to assess a method for supporting decisions regarding identification of most suitable areas for two types of harvesting approaches in forestry: skyline vs. forwarder. The innovative aspect consists in simulating the choices done during the planning in forestry operations. To do so, remote sensing data from an aerial laser scanner were used to create a digital terrain model (DTM) of ground surface under vegetation cover. Features extracted from the DTM are used as input for several machine learning predictors. Features are slope, distance from nearest roadside, relative height from nearest roadside and roughness index. Training and validation is done using areas defined by experts in the study area. Results show a K value of almost 0.92 for the classifier with best results, random forest. Sensibility of each feature is assessed, showing that both distance and height difference from nearest road-side are more significant than overall DTM value.

1. INTRODUCTION

In the last decades, the interest in forest planning has been renewed with the use of high-resolution remote sensing data and with the use of advanced spatial analysis techniques (Baskent and Keles, 2005; Li et al., 2007). Spatial analysis has also been applied successfully to wood harvesting and wood transportation both at the tactical and operational level (Grigolato et al., 2017). However, planning wood harvesting operations in complex situations, such as in wet-area (Murphy et al., 2007), roughness terrain (Duka et al., 2017), steep terrain (Bont et al., 2012; Kuhmaier and Stampfer, 2010) and low density of forest road network (Cavalli and Grigolato, 2010; Najafi and Richards, 2013) is challenging. Considering wood harvesting operations, the main phases are the following: (i) felling, (ii) tree processing, (iii) yarding and (iv) loading and hauling. Generally, in term of transportation, wood harvesting can be split into primary transportation and haul (or secondary transportation) (Owende, 2004). The primary transportation consists of extracting logs from the stump site to the nearest road along which logs are loaded into a truck for long transportation to the final wood processing sites. In the context of primary transportation, steep slope terrain together with a widespread terrain roughness force the choice of forest utilization systems based on highly specialized machines (Mologni et al., 2016). The main forest machines used in very steep slope terrain are an overhead system of the cinch-driven cable such as cable cranes (Proto et al., 2016) and articulated chassis tractors such as forwarders (Strandgard and Mitchell, 2015). Recently, the introduction of winch assisted machine systems lets also forwarders expand to work also in steeper terrain up to 70–80% (Mologni et al., 2018).

Setting up spatial analysis models for supporting forest technicians in the medium-term planning of wood harvesting is a very important task. The current availability of high-resolution

earth surface information such as the one derived from laser identification detection and ranging (lidar) can give an advantage in terms of awareness also in forest environment (Pirotti et al., 2017). Certainly, the complexity of the spatial analysis (Grigolato et al., 2017) depends on a large number of variables, data sets and software. In fact, in addition to the slope and roughness of terrain, other important factors are the forest growing stock distribution, stand parameters and distance from the forest roads.

Remote sensing plays an important role in supporting forestry operation plans. Geographic information systems (GIS) can validly support decision-makers with data analysis (Martin et al., 2016; Piragnolo et al., 2014). The information derived from close- and far- range sensing using lidar and/or optical data is used for modelling vegetation, geomorphology(Rutzinger et al., 2018), erosion, hazards (Pirotti et al., 2015). The combination of lidar-derived features and machine learning can obtain positive results that exceed the traditional approaches. Machine learning methods have been applied in several studies, in particular in land use / land cover (Pirotti et al., 2016). It can be effective using geological maps, lithology, digital elevation model (DEM) and morphological information e.g. altitude, slope, distance from road (Pradhan, 2013). Positive results in many fields of research led to choosing this type of approach in the presented work.

The goal of this research is to assess classification methods that use remote sensing derived data to support decision-makers. Morphological information is processed and fed to machine learners to identify suitable areas for two classes of forestry harvesting methods: skyline and forwarder. Both techniques aim at carrying trees or part of trees from cutting from the forest to the side of the road. Skyline uses cables to slide trees to the side of the road, whereas forwarders are vehicles on wheels with mechanized arms to load trees and carry them out by driving.

Defining which of the two methods is best using geospatial data and spatial modelling can bring innovation in the field of forestry planning and management, simulating decisions in the field of forestry harvesting operations.

2. MATERIALS AND METHOD

2.1 Study area

The study area is in the southern part of the Altopiano dei Sette Comuni in the Veneto region in Northern Italy, (Figure 2). The study area is 34.32 km². The Pre-Alps area is largely covered by forests, which are an important economical asset of the region. Two types of approaches are used for wood harvesting, forwarders and skylines. The road network is an important factor affecting choice of best approach (Figure 2), but is not the only one.

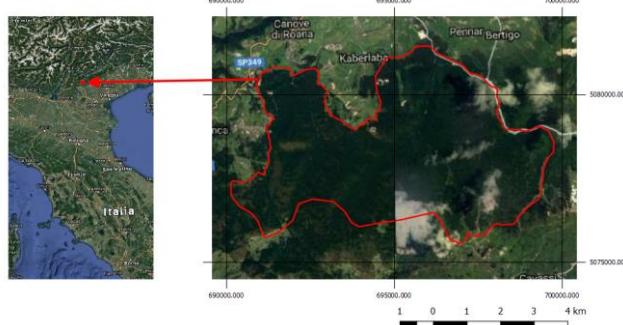


Figure 1. Study area.

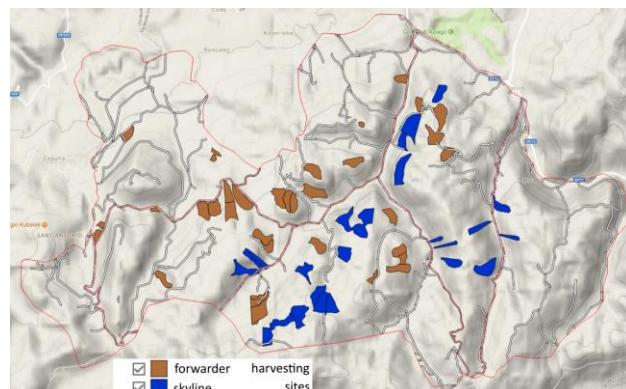


Figure 2. Location of harvesting sites used for training and validation, divided by wood harvesting approach.

2.2 Dataset

The digital terrain model (DTM) is derived from an aerial laser scanning mission that surveyed with Optech ALTM GEMINI 1064 nm wavelength, 167 kHz max pulse repetition frequency (PRF), 0.25 mrad (1/e) beam divergence. The point cloud average density was about 2 points per square meter. The points were processed using Terrasolid® software, which provides a progressive triangulation densification algorithm for ground points classification (Axelsson, 1999).

The final DTM resolution was set to 1 m x 1 m, which covered the area with more than 2 million cells. Based on the DTM, a slope and roughness map were calculated. Forestry forwarders can work only with slope lower than 40% and a maximum roughness of 70 centimetres. Otherwise, a skyline is preferable. Up to now, slope and roughness a map was used for indicating

preference between forwarder and skyline solutions. In our case we have added two extra layers to test, via machine learners, if a better result can be obtained. The two layers are the elevation difference between each cell the nearest road, and the distance from the nearest road. The reason is to simulate decisions taken during the forestry operations. In our simulation, the decision maker is inside the harvesting site, and defining for each cell the best choice for harvest type. We assume that it is more natural to focus the attention on the accessibility of the area rather than to estimate the morphological aspect, such as slope or roughness. Therefore, the decision maker will use the most intuitive solution by identifying the minimum elevation difference between the place where he is standing (cell) and the nearest road.

The distance and elevation difference from the nearest road has been recognized for each cell in our area, which represents the position. Then, the road elevation has been extracted for all the pixels over the road network, and the elevation difference has been calculated. The choice of nearest road cell was done using fast k-nearest neighbour (KNN) search, using cell centers as points.

The final dataset used for input in machine learners is composed of six layers: the DTM, the CHM, the slope, the roughness, the distance from between the roads, and the elevation difference between each cell and the nearest cell of the road (see Table 1). To clarify the terminology, in the following part of this article we use "layer" to identify the single raster of the stack, whereas use "input" to identify the raster layer used as input in the sensitivity analysis.

Raster base layer	Raster Stack layer (1x1 m resolution)
DTM	DTM
	Slope
	Roughness
CHM	Elevation difference from DTM and the nearest road
	CHM
Road's network	Distance from nearest road

Table 1. Summary of the layers that compose the raster stack.

2.3 Machine learning

The six layers described in the previous section were used as input features for several machine learning algorithms. The algorithms tested are conditional inference trees (Ctree), k-nearest neighbours (KNN), linear discriminant analysis (LDA), logistic regression (LR), multi-layered perceptron (MLP), multi-layered perceptron ensemble (MLPE), naïve, naivebayes (NB), and random forest (RF). Training and validation was done with areas defined by an expert in the field (see Figure 2), which knows the area very well and for this specific research, provided a careful evaluation of ideal areas for the two types of harvesting methods (forwarder and skyline).

The Ctree uses a tree to partition recursively the covariate M that influence the variable Y. M is an m-dimensional covariate vector X=(X₁,..., X_m). The algorithm fits a learning sample, where the learning sample is a random sample composed of n observations vector. Consequently, the algorithm works recursively on the vector. The steps of binary partitioning are: testing the null hypothesis between the m covariate and Y and

selecting the covariate X_i with the strongest association with Y . Measuring the associated P-values. Running the algorithm until the hypothesis is accepted. Splitting the variable into two subgroups, and repeating the previous steps. In case of the covariate X_i is missing, a new split can be calculated leading the same division of original split. When the null hypothesis is accepted, the algorithm stops (Hothorn et al., 2015)

The KNN is based on a distance metric. Then, it finds the k nearest neighbours of the sample, and it assigns the sample with a majority vote among the K -nearest Neighbors. The method advantage is to adapt to a new training data set, but the computational request and storage space is significant. The key to avoiding overfitting and the underfitting problem is the right choice of k , for example, using the standardized Euclidean distance. The steps are: choose the number of k nearest neighbours and a distance metric. Determine the k nearest neighbours. Count the majority vote and assign the class label (Raschka, 2015). The LDA is an algorithm for solving multi-class classification problems. The LDA is based on the dataset normality assumption and the independence feature assumption. The goal is to find the best variable combination for dividing the space into regions using a linear boundary. (Venables and Ripley, 2002). The LR is an algorithm for linear and binary classification based on neural network, where the classes are linearly separable. The binary classification is extended to multiclass classification using the One-vs-Rest (OvR) method. The OvR trains one classifier for a class. Then, it assigns a positive value to the class membership, and all other classes are considered as negative values. Iterating n times, where n is the number of classes a specific sample is assigned to a class considering the probability of membership. Logistic regression tries to maximize the likelihoods of the training data (Cheng et al., 2006). The MLP and the MLPE are neural networks composed of two perceptron layers. The perceptron algorithm has been developed in the 50s, and it is a single layer neural network with a threshold activation function, where the inputs are connected directly output using connected with adaptive weights. The advantage of the perceptron is to find a solution in a finite number of steps as studied by several authors (Arbib, 1987)(Duda and Hart, 1973) (Hand, 1985) (Hertz et al., 1991)(Minsky and Papert, 1988)(Van Der Malsburg, 1986). However, when the dataset is not linearly separable data the learning algorithm cycle to infinite and never terminate. The multi-layer perceptron is a perceptron network composed of layers with adaptive weights. Thus, the input units of the first layer are connected to the output layer through intermediate layers (Atkinson and Tatnall, 1997) (Benz et al., 2004). The naive and NB algorithm tries to solve the classification problem using a simple and direct way. It could be described as a query that starts from the source to the end of the dataset. The NB uses the probability theorem to classify a normally distributed dataset, where the class feature is independent. Nevertheless, in the case of weak violation of independence assumptions in a small sample, the NB still tend to perform well. In case of strong violation of independence assumptions or non-linear classification, very poor performance is obtained (Raschka, 2014). The RF is a decision tree based on a bootstrap technique that creates a large number of training sets to compute the statistics. Hence, it draws randomly a sample of size n with no replacement, and the value of n controls the bias and the variance. Replacement means an element can appear multiple times in the sample. On one hand, large values of n reduce randomness and increase the overfitting risk, on the other hand, small values of n reduce risk and model performance. Then, a decision tree grows, selecting d feature without replacement, so an element can appear only once in the sample. The tree nodes

are split and the procedure is repeated k times. Consequently, the class label is assigned by majority vote to the aggregated classification (Breiman, 2001).

2.4 Training and validation

The training and validation cells are picked randomly, to avoid spatial autocorrelation. To obtain a robust accuracy assessment, a stratified sample of 10% of the total cells was taken. Previous literature (Piragnolo et al., 2017) shows that using a lower number of cells for training can decrease accuracy, and that more does not improve results significantly. This, of course, depends on the case, the number of features used, on the sensitivity of each feature and on the number of classes, which in our case is two. The 10% was used as an indicative amount for choosing the training set size. The machine learning framework has been tested using a similar approach to the framework described in (Pirotti et al., 2016). The training phase uses regions of interest (ROIs) that correspond to two types of harvesting site where forwarder and skyline machines have been used. During the training and the testing phase, the accuracy has been evaluated using three metrics applying k-fold cross-validation. The metrics are: kappa (K), accuracy (ACC) and classification error (CE), and the execution time. The k-fold cross-validation method splits the subset in 10 folds, thus using 90% for training and the remaining information for testing, and helps to reduce overfitting problems.

2.5 Sensitivity analysis

The trained model has been tested for sensitivity analysis (SA). The SA is a technique to interpret a black box model, which was used in this work to analyse and interpret our six input layers. Indeed, a black box model produces one output (y) using M inputs. Applying the sensitivity samples to the fitted model, the model responses are obtained as equation below:

$$\hat{y} = P(x) \quad (1)$$

where \hat{y} is the value predicted by the model, $P()$ is the function used to build the predicted value and x is the data sample. Consequently, the input variable $\{x_a; a \in \{1, M\}\}$ ranges through minimum to maximum with L levels. Calculating the input sensitivity for each input variable, the most relevant input can be identified through the variation of the input level. In this method all inputs are kept at their average value, and one is changed, assuming the L value (Kewley et al., 2000). For example, using a One-dimensional SA (1D-SA) with 7 levels, the x_a ranges [0, 1], and the values are $x_{aj} \in (0.0, 0.14, 0.29, 0.57, 0.71, 0.86, 1.0)$. The following formula describes the model (2):

$$\hat{y}_{aj} = \{\hat{y}_{aj} : j \in \{1, \dots, L\}\} \quad (2)$$

where \hat{y}_{aj} is the x_{aj} response.

The input importance can be calculated using the input sensitivity. The sensitivity measure is calculated using the Average Absolute Distance (AAD) from the median (s_d) proposed by (Cortez and Embrechts, 2013). A relevant input can be identified through the variation of the input level, and a variation of input levels produce output changes. The AAD is a robust method few sensitive to out layers, where a high value indicates high input relevance, as stated in equation (3).

$$s_d = \sum_{j=1}^L |\hat{y}_{aj} - \widetilde{y}_a| / L \quad (3)$$

The sensitivity measure is calculated from the responses of x_a , and a high value indicates high relevance. The relative importance (4) is:

$$r_a = \xi_a / \sum_{i=1}^M \xi_i \quad (4)$$

r_a = relative importance

ξ_a = sensitivity for input x_a

The sensitivity values have been plotted using the Variable Effect Curve (VEC) that plot input x_{aj} versus the \hat{y}_{aj} responses. In the paper, the input variables are the layers of the raster stack previous described. Consequently, I1 is the DTM layer, I2 is the CHM, I3 is the slope, I4 is the roughness, I5 is the distance from DTM and the roads and I6 is elevation difference between a DTM cell and nearest roads cell.

3. RESULTS AND DISCUSSION

3.1 Collinearity

It is well known in literature that correlated features are not helpful in classification methods, and might decrease accuracy. It is true that tree-based machine learning methods are very robust with respect of collinearity in input features. Nevertheless a check on correlation between the input features was done and resulted in very weak correlation, thus proving that using all six features will result in a diverse set of information that the machine learning algorithms can used to fit the prediction model. The statistical analysis evidences weak and very weak correlation, so it reasonable think to a limited effect in the classification process. The scatterplot (Figure 3) shows a very low correlation between DTM and stack's layers.

3.2 Classification result

The accuracy metrics range between 0-100, and are reported in Table 2 and Table 3. The three best algorithms in terms of K and process time are the RF, the KNN, and the Ctree, and they have been used for the testing. The K value for RF is 95.49, for KNN is 91.04, and for the Ctree is 84.36. In contrast, the faster algorithm is the KNN, whereas the slower algorithm is the RF.

Model	Train. time (s)	K	ACC	CE	Classif. time (s)
RF	292.35	95.49	97.60	2.40	9690.19
KNN	0.01	91.04	95.23	4.77	0.33
Ctree	34.87	84.36	91.70	8.30	1155.80
Mlpe	136.64	63.56	80.60	19.40	4529.05
Mlp	136.76	62.80	80.20	19.80	4533.03
Lr	9.59	54.99	75.95	24.05	317.87
NB	0.36	54.65	75.61	24.39	11.93
Lda	0.37	52.04	73.67	26.33	12.26
Naive	0.01	0.01	52.64	47.36	0.33

Table 2. Accuracy metric for the training set. The table reports also the time estimation for the classification for the full dataset.

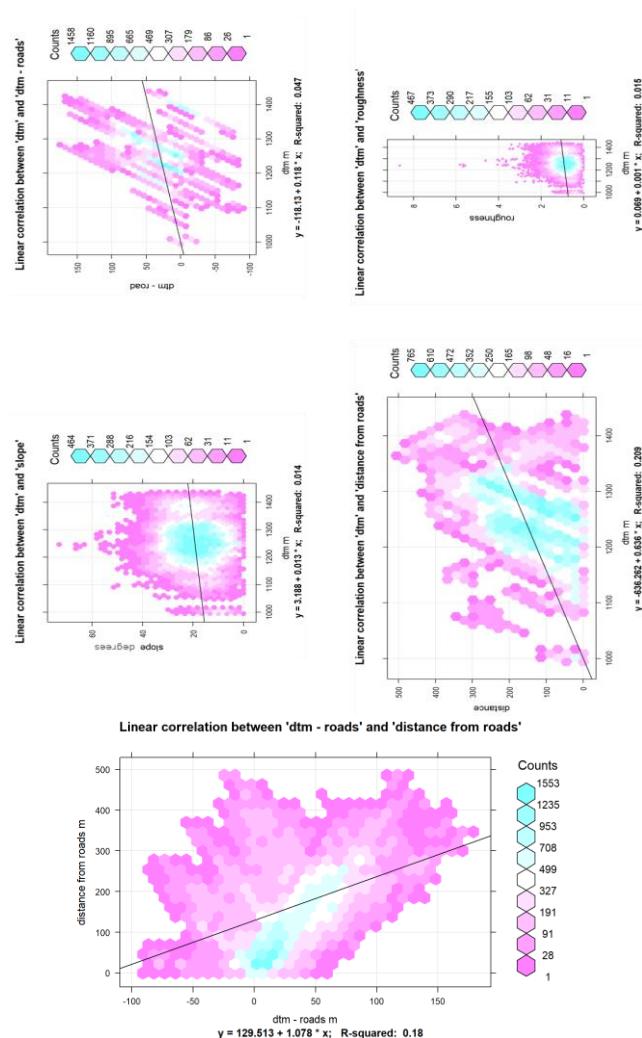


Figure 3. Scatterplot to visualize large dataset.

The result of machine learning using the test set is reported in the following Table 3 and Figure 4. Likewise, in the training phase, RF has the higher K value, and the Ctree has the lowest K value. In comparison to the training phase, the values have been decreased.

Model	Time	K	ACC	CE
RF	770.2	91.85	95.67	4.33
KKN	57.45	87.07	93.13	6.87
Ctree	50.52	77.93	88.33	11.67

Table 3. Accuracy metric for the test set

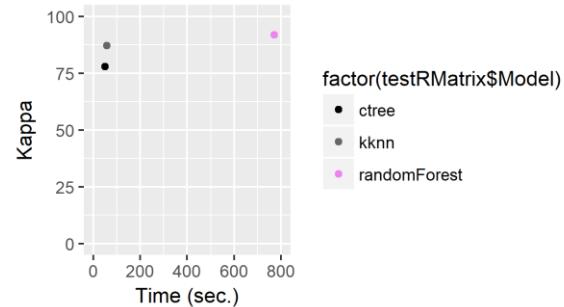


Figure 4. Time vs K accuracy for the test set.

Specifically, the RF is the slower algorithm, but it has a high K score of 91.85. KNN has K of 87.07. In contrast to the training set result, the KNN is slower than the Ctree. The Ctree is the fastest algorithm with a K of 77.93.

3.3 Sensitivity results

Finally, the sensitivity analysis has been conducted for the three algorithms, RF KNN and CTree, using the AAD metric. The AAD metric measures the relative importance of RF inputs as reported in Figure 5. Results show that the most important inputs are the elevation difference between the DTM and the roads, followed by the DTM values, and the distance between road and cell. In contrast, the CHM, the slope and the roughness have less importance. To understand the contribution of each input in the classification process, a detailed analysis of the inputs has been done using the VEC plot as a report in Figure 6. The values are summarized in Table 4.

The elevation difference, which ranges between negative values and 50 metres, has a high sensitivity in the forwarder class. In contrast, the skyline label is assigned when the elevation difference is greater than 100 metres. Between 0 and 50 metres, the sensitivity is 0.5 for both classes.

The DTM has a threshold of 1300 metres that clearly discriminates the forwarder and the skyline classes.

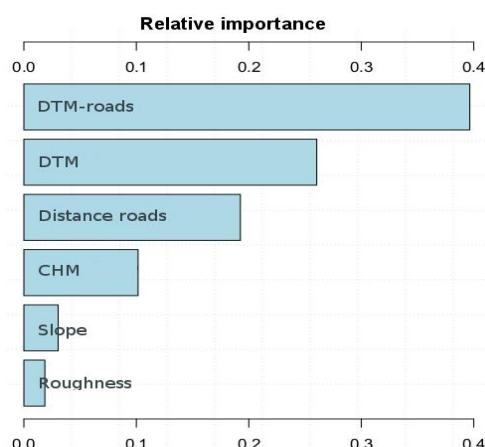


Figure 5. Random Forest relative importance of the input variables. The x-axis reports the AAD metric.

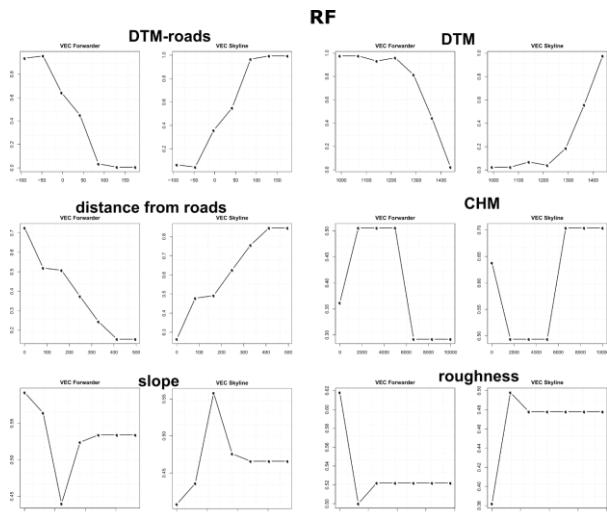


Figure 6. Sensitivity for the two classes of the RF classification.

The distance from roads affects the forwarder class for values lower than 100 metres. In contrast, the skyline is recognized for values greater than 200 metres. Between 100 and 200 metres, there classification is not clear, and both classes have a value of 0.5. The CHM sets a threshold of 6 metres that discriminates well the forwarder and the skyline classes.

The slope of 20 degrees is the threshold that discriminates the forwarder and the skyline. Indeed, the forwarder sensitivity ranges between 0 and 20 degrees, whereas the skyline ranges from 20 degrees to 40 degrees.

The roughness sensitivity can be found for values lower than 1 metre, but the flat trend suggests a low importance of this input for the classification.

Input	Forwarder Sensitivity	Skyline Sensitivity
Elevation difference	< 50 m	> 100 m
DTM	< 1300 m	> 1300 m
Distance from roads	< 100 m	> 200 m
CHM	< 6 m	> 6 m
Slope	< 20°	20° - 40°
Roughness	-	-

Table 4. Class sensitivity for the input layers using RF

The KKN has a K score of 87.07, and it is faster than RF. The inputs relative importance for the KNN is reported in Figure 7. The most important input is the elevation difference between the DTM and the roads. Then, the distance between roads and CHM. The slope, the roughness and the DTM are fewer important for the classification process.

The specific VEC plots are reported in Figure 8, and the values are summarized in Table 5. The sensitivity of the elevation difference is similar to the RF. The forwarder label is assigned for values lower than 50 metres, whereas the skyline is assigned for values greater than 50 metres.

The distance from the roads shows a saw shape curve. The forwarder sensitivity ranges between 0.2 and 1, and the skyline sensitivity ranges between 0.1 and 0.7. Still, the peak of one class corresponds to the minimum of the other class. Consequently, in this range, there is a high sensitivity for all the classes, but the forwarder has slight higher values. However, the skyline sensitivity reaches the top when the distance is greater than 300 metres.

The CHM trend highlight the value of 6 metres is significant to define the two classes. Below 6 metres of the tree height, the forwarder has the maximum sensitivity. Above 6 metres, the skyline is defined.

The slope has a clear peak at 40 degrees. The sensitivity for the forwarder reaches the maximum below the threshold of 40 degrees. In addition, above 60 degrees, forwarder gets a high score, but the result is not significant. Indeed, the limit operation of the forestry machine is 40 degrees. Instead, the skyline has a high value between 20 degrees and 60 degrees, where the maximum is 40 degrees.

Input	Forwarder Sensitivity	Skyline Sensitivity
Elevation difference	< 50 m	> 50 m
Distance from roads	< 300 m	> 300 m
CHM	< 6 m	> 6 m
Slope	< 40°	> 40°
Roughness	< 1	> 1
DTM	-	-

Table 5. Class sensitivity for the input layers using KNN.

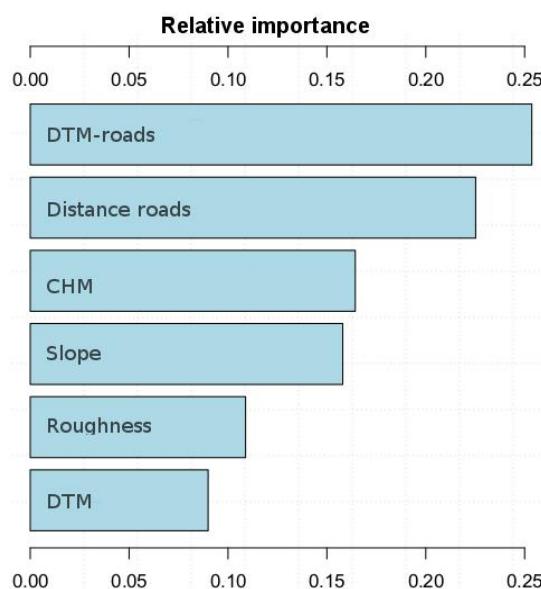


Figure 7. KNN relative importance of the input variables. The x-axis reports the AAD metric.

The roughness shows a saw shape trend, where the minimum and the maximum are alternated. The sensitivity is high, but the roughness values do not match with the real operation limits. Finally, the DTM has high sensitivity without variation for both classes, which implicates a low relative importance.

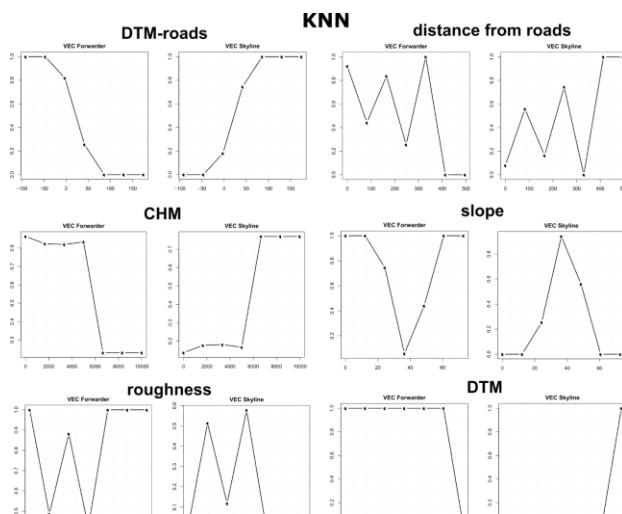


Figure 8. Inputs sensitivity analysis for the two classes of the KNN classification.

The Cree classifier has a K score of 77.93, and it is the fastest algorithms used in the test. The AAD metric and the relative importance are reported in Figure 9. The VEC plots for the two classes, forwarder and skyline, are reported in Figure 10, and the values are summarized in Table 6.

A high sensitivity has been found for the distance from roads. Both classes range between 0.2 and 1. Thus, the threshold can be set at 300 metres distance.

The forwarder is defined below the 300 metres and the skyline above the 300 metres distance. Regarding the elevation difference input, the forwarder has high sensitivity when the elevation difference ranges between -100 metres and 50 metres.

In contrast, the skyline gets a high score only when the elevation difference is greater than 50 metres. Analysing the DTM, forwarder has high sensitivity when the elevation is lower than 1300 metres, whereas the skyline above the 1300 metres of elevation.

The slope threshold for the two classes is defined at 20 degrees. Indeed, below 20 degrees forwarder has a score of 1, whereas above 20 degrees the skyline sensitivity increases. The sensitivity of tree height is high until 2 metres, where forwarder sensitivity is high at 0 metres and forwarder at 2 metres. Two metres is a low height for a coppice, and it suggests that CHM can be omitted from the dataset. Finally, roughness does not produce relative important output.

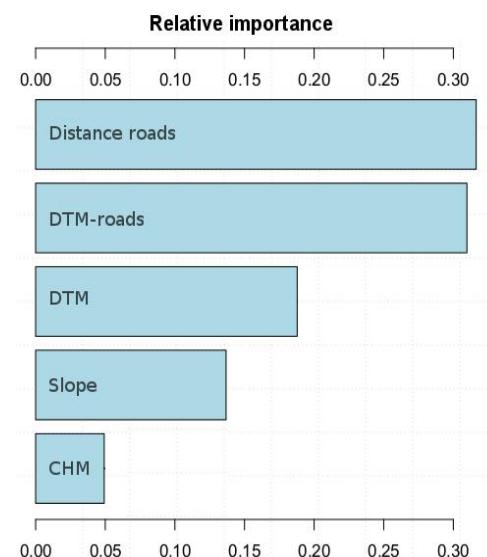


Figure 9. Relative Importance of the input variables using Ctree. The x-axis reports the AAD metric.

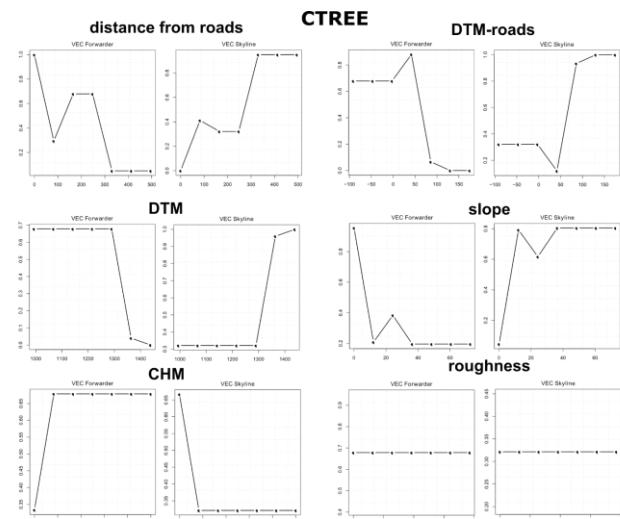


Figure 10. Inputs sensitivity analysis for the two classes of the Ctree classification.

Input	Forwarder Sensitivity	Skyline Sensitivity
Distance from roads	< 300 m	> 300 m
Elevation difference	< 50 m	> 50 m
DTM	< 1300 m	> 1300 m
Slope	< 10° m	10° - 40°
CHM	-	-
Roughness	-	-

Table 6. Class sensitivity for the input layers using Ctree.

4. CONCLUSIONS

In this paper, we propose to use spatial and morphological information, such as the DTM, the CHM, the slope, the roughness, the elevation difference and the distance from roads using machine learning for testing several algorithms and classifying a suitability map for two forestry machine types: forwarder and skyline. Forestry machine operates in a defined range of technical specifications and parameters, and the parameters are considered by the decision maker for defining the harvesting area. The novel approach consists to integrate the information on the accessibility of the area for simulating the worker choices. The benchmark shows the RF has the best performance in terms of K, but it is the slower in comparison to the KNN and the Ctree. However, the KNN process has a high accuracy, and the processing is very fast. Finally, Ctree is fastest, but it has the lower accuracy. To support a decision-maker, understanding and interpretability of the data model is a key issue. For this purpose, the SA analysis has been conducted. The SA evidences the inputs relative importance for the RF, the KNN and the Ctree, where the accessibility have higher relative importance rather than the morphological aspect. The accessibility is described through of the harvesting site describes through the distance from the roads, and the elevation difference between the roads. Consequently, the slope and the roughness are less important for the classification. The forestry machine has some operation limits, specified in terms of maximum slope or roughness of the harvesting site. However, the choice of the route based on the accessibility of the area is a human decision. In addition, the VEC plots report the sensitivity of each input. Then, the RF uses the elevation and the distance from the roads network for making the prediction, and all inputs can give a contribution to the classification. The KNN is similar to the RF, and it considers the CHM as an important input. On the other hand, the roughness and the DTM do not add significant information.

Finally, the Ctree takes into consideration the accessibility of the harvesting site, but the CHM provides an information that has no evidence in the real domain. Moreover, the roughness does not influence the classification. Considering the low relative importance of some inputs, more studies can be conducted using only the important inputs as a strategy to improve the performance and reducing the computing time.

REFERENCES

- Arbib, M.A., 1987. Brains, Machines, and Mathematics (2Nd Ed.). Springer-Verlag, Berlin, Heidelberg.
- Atkinson, P.M., Tatnall, A.R.L., 1997. Introduction Neural networks in remote sensing. *International Journal of Remote Sensing*, 18, 699–709. doi:10.1080/014311697218700
- Axelsson, P., 1999. Processing of laser scanner data—algorithms and applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54, 138–147. doi:10.1016/S0924-2716(99)00008-8
- Baskent, E.Z., Keles, S., 2005. Spatial forest planning: A review. *Ecological Modelling*, doi:10.1016/j.ecolmodel.2005.01.059
- Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I., Heynen, M., 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing*, doi:10.1016/j.isprsjprs.2003.10.002
- Bont, L.G., Heinemann, H.R., Church, R.L., 2012. Concurrent optimization of harvesting and road network layouts under steep terrain. *Annals of Operations Research*, 232, 41–64. doi:10.1007/s10479-012-1273-4
- Breiman, L., 2001. Random forests. *Machine Learning*, 45, 5–32. doi:10.1023%2FA%3A1010933404324
- Cavalli, R., Grigolato, S., 2010. Influence of characteristics and extension of a forest road network on the supply cost of forest woodchips. *Journal of Forest Research*, 15, 202–209. doi:DOI 10.1007/s10310-009-0170-4
- Cheng, Q., Varshney, P.K., Arora, M.K., 2006. Logistic Regression for Feature Selection and Soft Classification of Remote Sensing Data. *IEEE Geoscience and Remote Sensing Letters*, 3, 491–494. doi:10.1109/LGRS.2006.877949
- Cortez, P., Embrechts, M.J., 2013. Using sensitivity analysis and visualization techniques to open black box data mining models. *Information Sciences*, 225, 1–17. doi:10.1016/j.ins.2012.10.039
- Duda, R.O., Hart, P.E., 1973. Pattern classification and scene analysis / Richard O. Duda, Peter E. Hart. Wiley New York.
- Duka, A., Grigolato, S., Papa, I., Pentek, T., Poršinsky, T., 2017. Assessment of timber extraction distance and skid road network in steep karst terrain. *iForest - Biogeosciences and Forestry*, 10, 886–894. doi:10.3832/ifor2471-010
- Grigolato, S., Mologni, O., Cavalli, R., 2017. GIS Applications in Forest Operations and Road Network Planning: an Overview over the Last Two Decades. *Croatian Journal of Forest Engineering*, 38, 175–186.
- Hand, D. J., 1985. Discrimination and Classification. *Biometrical Journal*, 27, 148. doi:10.1002/bimj.4710270204
- Hertz, J., Krogh, A., Palmer, R.G., 1991. Introduction to the Theory of Neural Computation. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.
- Hothorn, T., Hornik, K., Zeileis, A., 2015. ctree: Conditional Inference Trees [WWW Document]. *The Comprehensive R Archive Network*.
- Kewley, R.H., Embrechts, M.J., Breneman, C., 2000. Data strip mining for the virtual design of pharmaceuticals with neural networks. *IEEE Transactions on Neural Networks*, 11, 668–679. doi:10.1109/72.846738
- Kuhmaier, M., Stampfer, K., 2010. Development of a Multi-

- Attribute Spatial Decision Support System in Selecting Timber Harvesting Systems. *Croatian Journal of Forest Engineering*, 31, 75–88.
- Li, R., Bettinger, P., Danskin, S., Hayashi, R., 2007. A Historical Perspective on the Use of GIS and Remote Sensing in Natural Resource Management, as Viewed through Papers Published in North American Forestry Journals from 1976 to 2005. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 42, 165–178. doi:10.3138/carto.42.2.165
- Martin, A., Botequim, B., Oliveira, T.M., Ager, A., Pirotti, F., 2016. Temporal optimization of fuel treatment design in blue gum (*Eucalyptus globulus*) plantations. *Forest Systems*, 25, eRC09. doi:10.5424/fs/2016252-09293
- Minsky, M.L., Papert, S.A., 1988. Perceptrons: Expanded Edition. MIT Press, Cambridge, MA, USA.
- Mologni, O., Dyson, P., Amishev, D., Proto, A.R., Zimbalatti, G., Cavalli, R., Grigolato, S., 2018. Tension monitoring on large winch-assist forwarders operating in British Columbia. *Croatian Journal of Forest Engineering*, 39, 95–105.
- Mologni, O., Grigolato, S., Cavalli, R., 2016. Harvesting systems for steep terrain in the Italian Alps: state of the art and future prospects. *Contemporary Engineering Sciences*, 9, 1229–1242. doi:10.12988/ces.2016.68137
- Murphy, P.N.C., Ogilvie, J., Castonguay, M., Meng, F.R., Arp, P.A., 2007. Verifying calculated flow accumulation patterns of mapped and unmapped forest streams by culvert location. *Forestry Chronicle*, 83, 198–206. doi:10.5558/tfc83198-2
- Najafi, A., Richards, E.W., 2013. Designing a Forest Road Network Using Mixed Integer Programming. *Croatian Journal of Forest Engineering*, 34, 17–30.
- Owende, P.M.O., 2004. Wood Delivery, in: Burley, J., Evans, J., Youngquist, J.A. (Eds.), *Encyclopedia of Forest Sciences*. Elsevier, Oxford, pp. 269–279. doi:10.1016/B0-12-145160-7/00007-7
- Piragnolo, M., Masiero, A., Pirotti, F., 2017. Open source R for applying machine learning to RPAS remote sensing images. *Open Geospatial Data, Software and Standards*, 2, 16. doi:10.1186/s40965-017-0033-4
- Piragnolo, M., Pirotti, F., Guarneri, A., Vettore, A., Salogni, G., 2014. Geo-Spatial Support for Assessment of Anthropic Impact on Biodiversity. *ISPRS International Journal of Geo-Information*, 3, 599–618.
- Pirotti, F., Guarneri, A., Masiero, A., Vettore, A., 2015. Preface to the special issue: the role of geomatics in hydrogeological risk. *Geomatics, Natural Hazards and Risk*, 6, 357–361. doi:10.1080/19475705.2014.984248
- Pirotti, F., Neteler, M., Rocchini, D., 2017. Preface to the special issue “Open Science for earth remote sensing: latest developments in software and data.” *Open Geospatial Data, Software and Standards*, doi:10.1016/B978-008044335-5/50026-7
- Pirotti, F., Sunar, F., Piragnolo, M., 2016. Benchmark Of Machine Learning Methods for Classification of a Sentinel-2 Image. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B7, 335–340. doi:10.5194/isprs-archives-XLI-B7-335-2016
- Pradhan, B., 2013. A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS. *Computers and Geosciences*, 51, 350–365. doi:10.1016/j.cageo.2012.08.023
- Proto, A.R., Skoupy, A., Macri, G., Zimbalatti, G., 2016. Time consumption and productivity of a medium size mobile tower yarder in downhill and uphill configurations: a case study in Czech Republic. *Journal of Agriculture Engineering*, 67, 216–221. doi:10.4081/jae.2016.551
- Raschka, S., 2014. Naive Bayes and Text Classification I – Introduction and Theory 1–20. doi:10.13140/2.1.2018.3049
- Raschka, S., 2015. Python Machine Learning, First editi. ed, Igarss 2014. Birmingham.
- Rutzinger, M., Bremer, M., Höfle, B., Häggerle, M., Lindenbergh, R., Oude Elberink, S., Pirotti, F., Scaioni, M., Wujanz, D., Zieher, T., 2018. Training in Innovative Technologies for Close-Range Sensing in Alpine Terrain. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-2, 239–246. doi:10.5194/isprs-annals-IV-2-239-2018
- Strandgard, M., Mitchell, R., 2015. Automated time study of forwarders using GPS and a vibration sensor. *Croatian Journal of Forest Engineering*, 36, 175–184.
- Van Der Malsburg, C., 1986. Frank Rosenblatt: Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms, in: Palm, G., Aertsen, A. (Eds.), *Brain Theory*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 245–248.
- Venables, W.N., Ripley, B.D., 2002. MASS: modern applied statistics with S. doi:10.1198/tech.2003.s33