

# FEASIBILITY OF MACHINE LEARNING METHODS FOR SEPARATING WOOD AND LEAF POINTS FROM TERRESTRIAL LASER SCANNING DATA

D. Wang<sup>a,\*</sup>, M. Hollaus<sup>a</sup>, N. Pfeifer<sup>a</sup>

<sup>a</sup> Department of Geodesy and Geoinformation, TU Wien, Vienna 1040, Austria  
(di.wang, markus.hollaus, norbert.pfeifer)@geo.tuwien.ac.at

**KEY WORDS:** Terrestrial Laser Scanning, Wood-leaf classification, Machine Learning, Feature selection

## ABSTRACT:

Classification of wood and leaf components of trees is an essential prerequisite for deriving vital tree attributes, such as wood mass, leaf area index (LAI) and woody-to-total area. Laser scanning emerges to be a promising solution for such a request. Intensity based approaches are widely proposed, as different components of a tree can feature discriminatory optical properties at the operating wavelengths of a sensor system. For geometry based methods, machine learning algorithms are often used to separate wood and leaf points, by providing proper training samples. However, it remains unclear how the chosen machine learning classifier and features used would influence classification results. To this purpose, we compare four popular machine learning classifiers, namely Support Vector Machine (SVM), Naïve Bayes (NB), Random Forest (RF), and Gaussian Mixture Model (GMM), for separating wood and leaf points from terrestrial laser scanning (TLS) data. Two trees, an *Erytrophleum fordii* and a *Betula pendula* (silver birch) are used to test the impacts from classifier, feature set, and training samples. Our results showed that RF is the best model in terms of accuracy, and local density related features are important. Experimental results confirmed the feasibility of machine learning algorithms for the reliable classification of wood and leaf points. It is also noted that our studies are based on isolated trees. Further tests should be performed on more tree species and data from more complex environments.

## 1. INTRODUCTION

Quantifying forest structure is of broad importance. For example, understanding forest foliage profile can be of particular interest for biodiversity conservation and climate adaptation, as it affects the photosynthesis and evapotranspiration processes (Ma et al., 2016). Monitoring carbon stocks in forested ecosystems requires accurate quantification of the spatial distribution of wood volume (Levick et al., 2016). Moreover, description of 3D structure helps to investigate species competition, wood production, and ecosystem and agro-ecosystem dynamics (Béland et al., 2014). For mapping forest structure, laser scanning is widely used in past decades. Laser scanning technique, also known as light detection and ranging (lidar), acquires 3D coordinates of objects over a large scale. In addition, full-waveform laser scanners are able to measure the scattering properties of vegetation in a quantitative way (Wagner et al., 2008). Therefore, laser scanning generates a high potential for forest related studies.

Assessment of canopy structure at tree or branch scale can be difficult with laser scanning data acquired from satellite and airborne platforms (Tao et al., 2015). Terrestrial Laser Scanning (TLS), on the other hand, has been established as an efficient tool for acquiring 3D data used for a range of fine-scale forest studies (Liang et al., 2016), including stem mapping (Liang et al., 2012), tree height measurement (Olofsson et al., 2014), diameter estimation (Wang et al., 2017), stem curve retrieval (Wang et al., 2016), biomass calculation (Kankare et al., 2013), and leaf area index (LAI) estimation (Zheng et al., 2013). To better retrieve forest ecological attributes, it is often necessary to separate wood and leaf components of trees (Tao et al., 2015). For example, estimation of LAI requires to screen out wood points, otherwise the wood returns will artificially increase the apparent foliage content (Béland et al., 2014).

Wood-leaf point separation for TLS data is challenging. In general, existing methods can be categorized into two groups; inten-

sity based and geometry based. Intensity based methods (Pfenigbauer and Ullrich, 2010; Béland et al., 2014) use radiometric information of objects captured by a laser scanner. The assumption is that wood and leaf components have different optical properties at the operating wavelength of the laser scanner (Tao et al., 2015). By determining a proper intensity threshold, wood and leaf points can be separated. However, intensity captured by a laser scanner needs an instrument specific radiometrical calibration before including it in further processing (Calders et al., 2017). Recently developed multi-wavelength (e.g., hyperspectral) scanners can help to better solve such a task (Li et al., 2013; Hakala et al., 2012; Vauhkonen et al., 2013). However, these scanners are still in an early development stage, and not yet widely available. Geometry based methods only use 3D coordinates of objects captured by a laser scanner. Local structure-related saliency information are derived from 3D points and supervised machine learning methods such as Support Vector Machine (SVM) (Yun et al., 2016) and Gaussian Mixture Model (GMM) (Ma et al., 2016) are often employed to classify wood and leaf points. Some direct geometric methods were also reported (Tao et al., 2015). Nevertheless, geometry based machine learning methods are rarely systematically examined for wood-leaf classification, although it is a well-known and widely adapted technique for other classification tasks (Weinmann et al., 2013, 2017; Brodu and Lague, 2012). There is a vast need to exploit 3D geometry based approaches for separating wood and leaf points, as 3D coordinates are the most fundamental information acquired by any laser scanners. For machine learning methods, various classifiers were used in previous studies (Yun et al., 2016; Ma et al., 2016). The lack of comparable studies calls for a specific examination on how the chosen machine learning classifier and features used would influence classification results.

This study aims to examine four machine learning algorithms, Support Vector Machine (SVM), Naïve Bayes (NB), Random Forest (RF), and Gaussian Mixture Model (GMM), in geometry-based wood and leaf points separation using TLS data. In the following section 2 the used data are described, in section 3 the ma-

\*Corresponding author

chine learning models used for separating wood and leaf points are presented. Finally, in section 4 the results are presented and discussed in section 5. Conclusion is given in section 6.

## 2. MATERIALS

### 2.1 *Erytrophleum fordii*

TLS data of an evergreen sub tropical tree, *Erytrophleum fordii*, were provided by Hackenberg et al. (2015). The data were acquired in October 2013 from eight scan positions. The acquired point cloud was further manually cleaned, as the tree crown interacts with other trees. Therefore, points from adjacent trees' foliage need to be removed. The cleaned point cloud for the *Erytrophleum fordii* tree contains  $\sim 3.9$  million points. The average distance between two adjacent points is  $\sim 5$  mm (Figure 1).

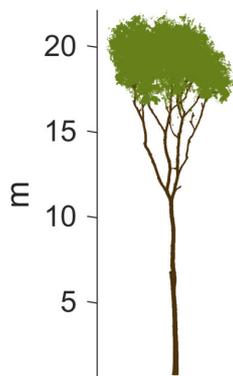


Figure 1. Point cloud of the *Erytrophleum fordii*.

### 2.2 *Betula pendula*

Hyperspectral TLS data of another silver birch tree (*Betula pendula*, Figure 2) were provided by Puttonen et al. (2016). The single-scan data feature radiometrical information of the scanned tree, in addition to the 3D XYZ coordinates. The average distance between two adjacent points is  $\sim 1$  cm. Measurements were carried out using a Hyperspectral Laser Scanner (HSL) from the Finnish Geospatial Research Institute (FGI) (Hakala et al., 2012). Laser radiometry was calibrated by setting up an external reference plate. For more information about the HSL data, readers are referred to Puttonen et al. (2016).

## 3. METHODS

### 3.1 Feature Calculation

Twenty-six 2D and 3D geometry-based point cloud features extracted for each data set were described in Table 1. The features were originally proposed and used in Weinmann et al. (2015) for urban area scene analysis. Local 3D features are inferred from the distribution of neighboring points of every point. Structure saliency such as planar, linear, and scattering can be inferred from the eigenvalues of the decomposed covariance matrix (Equation 1).

$$Cov_p = \frac{\sum_{i=1}^K (p_i - \bar{p})(p_i - \bar{p})^T}{K}, \quad (1)$$

where  $p_i = \{x_i, y_i, z_i\}^T$  is a 3D point and  $\bar{p}$  is the barycenter of the  $K$  nearest neighboring points. 2D feature calculation involves

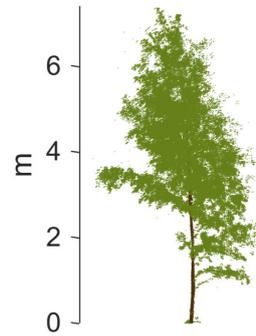


Figure 2. Point cloud of the *Betula pendula*, silver birch.

a projection of points onto the horizontal plane. For details of feature extraction procedures, the readers are referred to Weinmann et al. (2015).

### 3.2 Feature Selection

The high dimensionality of the input data may exhibit redundancy and can be potentially reduced by various feature selection algorithms. Moreover, feature selection may attenuate the over-fitting problem in multivariate classification methods (Geiß et al., 2015). Feature selection methods can be grouped into three categories; wrappers, embedded, and filters (Guyon et al., 2008). Wrappers methods evaluate a subset of features by accuracy estimates and require trained classifiers. Embedded methods embed the selection process into the classifier learning. On the other hand, filter methods explore the intrinsic properties of the data, and thus operate independently with respect to classifiers. In this study, we employ the filter method for feature selection for its simplicity and efficiency, although more robust and concrete methods are used in previous studies (Weinmann et al., 2013).

A fast and effective filter method is the *Fisher* method (Gu et al., 2012). This method computes a score (*Fisher score*) according to a ratio of interclass separation and intraclass variance for each feature and ranks them. The scores reflect the discriminative power of each feature. In this study, we apply the *Fisher* method for both *Erytrophleum fordii* and silver birch datasets. The resultant rankings are given in Table 1. Consequently, the classifier learning was performed for 5, 10, 15, 20, and 26 features accordingly, based on the rankings.

### 3.3 Machine Learning Classifiers

Wood-leaf separation is a binary classification problem. Given  $m$  training samples,  $(y_i, \mathbf{x}_i)$   $i = 1, \dots, m$  with labels  $y_i \in \{1, -1\}$  and  $n$  dimensional feature vectors,  $\mathbf{x}_i \in \mathbf{R}^n$ , the objective is to find a function  $f(\cdot; \alpha) : \mathbf{x} \mapsto y$  that represents the classifier  $y = f(\mathbf{x}; \alpha)$ , where  $\alpha$  are all the parameters of the classifier.

In this study, we examine the feasibility of four machine learning algorithms, SVM, NB, RF, GMM, for wood-leaf separation. In this section, the fundamentals and principles of the four used machine learning algorithms are briefly summarized.

**3.3.1 Support Vector Machine** SVM was proposed by Vapnik (1995). For a binary classification problem, it finds a hyperplane  $\mathbf{w} \cdot \mathbf{x} + b = 0$ , which maximizes the distance of the closest

Table 1. Features extracted from the point cloud. *EV* denotes the eigenvalue and *NV* is the normal vector. *EVs* are sorted in a descend manner. "R-" is the abbreviation of "Ranking" form feature selection. "R-Combined" means that the feature set is a combination of data from both trees. Top 5 ranked features are underlined.

No.	Feature	Description	R- <i>E. fordii</i>	R-Birch	R-Combined
1	linearity $L_{\lambda,3D}$	linear saliency $(EV_1 - EV_2)/EV_1$ .	23	19	22
2	planarity $P_{\lambda,3D}$	planar saliency $(EV_2 - EV_3)/EV_1$ .	17	26	17
3	scattering $S_{\lambda,3D}$	volumetric saliency $EV_3/EV_1$ .	14	14	15
4	omnivariance $O_{\lambda,3D}$	variance of the neighborhoods $\sqrt[3]{EV_1 * EV_2 * EV_3}$ .	16	20	16
5	anisotropy $A_{\lambda,3D}$	$(EV_1 - EV_3)/EV_1$ .	15	15	14
6	eigenentropy $E_{\lambda,3D}$	$-\sum_{n=1}^3 EV_n * \log(EV_n)$	18	22	19
7	sum_ $EV_{3D}$	$\sum_{n=1}^3 EV_n$ .	20	10	20
8	surface_variation $_{3D}$	change of curvature $EV_{min}/(\sum EV)$ .	12	17	12
9	Z value $Z_{\lambda,3D}$	the height of the point.	<u>1</u>	<u>2</u>	<u>1</u>
10	radius_knn $R_{knn,3D}$	radius of local neighborhood.	9	<u>4</u>	8
11	density $_{3D}$	local point density.	<u>3</u>	<u>1</u>	<u>3</u>
12	verticality $V_{\lambda,3D}$	$1 - NV_{s_z}$ .	24	13	24
13	$\Delta Z_{knn,3D}$	height difference of local neighborhood.	10	8	10
14	$\sigma Z_{knn,3D}$	standard deviation of heights of local neighborhood.	11	7	11
15	radius_knn $R_{knn,2D}$	radius of local neighborhood.	8	<u>5</u>	7
16	density $_{2D}$	local point density.	<u>2</u>	9	<u>2</u>
17	sum_ $EV_{2D}$	$\sum_{n=1}^2 EV_n$ .	21	11	21
18	$EV\_ratio_{2D}$	$EV_2/EV_1$ .	7	25	9
19	cell_density $_{2D}$	density of projected 2D cells.	<u>4</u>	6	<u>4</u>
20	$\Delta cell_{2D}$	height difference of points in each cell.	25	<u>3</u>	25
21	$\sigma cell_{2D}$	standard deviation of heights of points in each cell.	22	12	23
22	$EV\_1_{3D}$	first eigenvalue of 3D covariance matrix.	26	16	26
23	$EV\_2_{3D}$	second eigenvalue of 3D covariance matrix.	19	21	18
24	$EV\_3_{3D}$	third eigenvalue of 3D covariance matrix.	13	18	13
25	$EV\_1_{2D}$	first eigenvalue of 2D covariance matrix.	6	24	<u>5</u>
26	$EV\_2_{2D}$	second eigenvalue of 2D covariance matrix.	<u>5</u>	23	6

vectors (i.e., margin) in both classes.  $\mathbf{w}$  is the  $n$ -dimensional vector perpendicular to the hyper-plane, and  $b$  is the distance of the closest point on the hyper-plane to the origin. The classifier is then

$$f(\mathbf{x}) = \text{sgn} \left( \sum_{i=1}^m \lambda_i y_i K(\mathbf{x}_i, \mathbf{x}_j) + b \right), \quad (2)$$

where  $\lambda$  is the weight and  $K(\mathbf{x}_i, \mathbf{x}_j)$  is a kernel function  $K(\mathbf{x}_i, \mathbf{x}_j) = \Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j)$ , subjects to  $y_i(\langle \mathbf{w}, \mathbf{x}_i \rangle + b) - 1 \geq 0$ .

**3.3.2 Naïve Bayes** NB is a statistical approach based on Bayes's theorem (Marcot et al., 2006). It assumes that the features are conditionally independent given the class,

$$p(x|y) = \prod_{i=1}^m p(x_i|y). \quad (3)$$

Therefore, from the Bayes's theorem, the posterior probability of a feature vector to be part of a certain class is

$$p(y|x) = \frac{p(y) \prod_{i=1}^m p(x_i|y)}{p(x)}, \quad (4)$$

where  $p(y)$  is the prior probability of the class. A point will be labeled as the class with the highest probability.

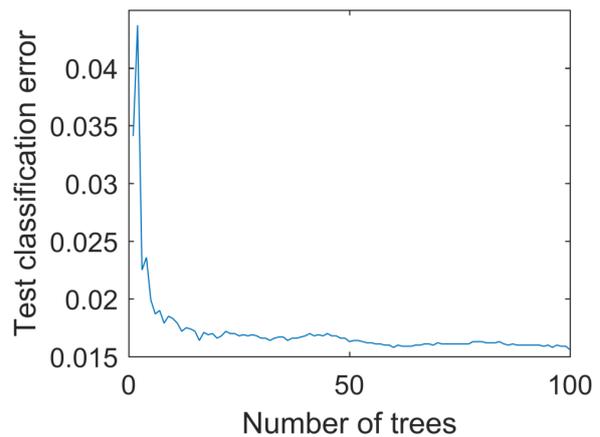


Figure 3. Evaluation of number of classification trees to be grown.

**3.3.3 Random Forest** RF is a decision tree based ensemble learning method that was proposed by Breiman (2001). The learned model is a collection of weak models. Multiple decision trees are grown on random subsets of training data. The class determina-

Table 2. Statistical evaluation of machine learning classifiers for wood-leaf separation.

Features	<i>E. fordii</i>				birch		
	Classifier	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy
5	SVM	0.94	1.00	0.97	0.88	0.93	0.93
	NB	0.91	0.99	0.95	0.90	0.90	0.90
	RF	0.96	1.00	0.98	0.93	0.94	0.94
	GMM	0.90	0.99	0.95	0.88	0.94	0.94
10	SVM	0.93	1.00	0.97	0.92	0.98	0.97
	NB	0.91	0.98	0.95	0.92	0.89	0.89
	RF	0.96	1.00	0.98	0.96	0.98	0.98
	GMM	0.91	0.99	0.95	0.87	0.97	0.97
15	SVM	0.93	1.00	0.97	0.94	0.97	0.97
	NB	0.91	0.97	0.94	0.93	0.86	0.86
	RF	0.95	1.00	0.98	0.96	0.98	0.98
	GMM	0.91	0.99	0.95	0.87	0.97	0.97
20	SVM	0.94	1.00	0.97	0.94	0.97	0.97
	NB	0.92	0.96	0.94	0.93	0.86	0.87
	RF	0.95	1.00	0.98	0.96	0.98	0.98
	GMM	0.91	0.99	0.95	0.87	0.97	0.97
26	SVM	0.94	1.00	0.97	0.94	0.97	0.97
	NB	0.93	0.95	0.94	0.93	0.86	0.87
	RF	0.97	1.00	0.98	0.95	0.98	0.98
	GMM	0.91	0.99	0.95	0.87	0.97	0.97
NDVI threshold					0.78	0.97	0.96

tion is based on a majority votes fashion. RF has proven to be an accurate and robust classification and regression approach, even on noisy data (Geiß et al., 2015).

When employing RF, two necessary parameters need to be specified; the number of classification tree  $n_{trees}$  and the number of input features  $m_{ft}$  used at each node (Geiß et al., 2015). A higher number of  $n_{trees}$  increases model accuracy until convergence. We used our data with all features to train models. We observe that in our study, the model performance converges at the point of approximate 60 trees (Figure 3). However, since our data set is not large enough for us to consider a trade off for computation time, we keep the number as 100. We set another parameter,  $m_{ft} = \sqrt{p}$ , where  $p$  denotes the number of input feature, as suggested by Breiman (2001).

**3.3.4 Gaussian Mixture Model** GMM is a modeling technique that uses a probability distribution to estimate the likelihood of a given feature vector. The assumption is that classes obey a normally distributed density function. For a binary classification problem, the continuous probability density function can be approximated as a linear combination of two probability density functions (Ma et al., 2016),

$$p(x) = \sum_{k=1}^m w_k p(x|k) \quad (5)$$

where  $w_k$  is the weight for each probability density function.  $p(x|k)$  is the conditional probability of a point  $x$  belonging to the  $k$ th density function. The probability that a point  $x_i$  lies within the a distribution with parameters  $\mu$  and  $\Sigma$  is given by

$$N(\mu_k, \Sigma_k) = \frac{e^{-\frac{1}{2}(x_i - \mu_k)^T \Sigma^{-1} (x_i - \mu_k)}}{\sqrt{|2\pi\Sigma|}} \quad (6)$$

In this study, manually delineated training points are used to train the GMM model. The *Expectation-Maximization algorithm (EM)* is used to estimate the  $\mu$  and  $\Sigma$  of each class. Consequently, a point will be labeled as the class with the highest probability.

### 3.4 Evaluation

The performance of each classifier is evaluated based on three statistical indexes; sensitivity, specificity, and accuracy. Sensitivity measures the correctly classified positive samples (true positive rate,  $TP$ ). In this study, it represents that the correct rate for wood points. Specificity gives the true negative rate ( $TN$ ), thus it measures the correct rate for leaf points. Accuracy ( $ACC$ ) gives the overall correctness by

$$ACC = \frac{TP + TN}{P + N}, \quad (7)$$

where  $P$  and  $N$  are the number of real positive (wood) and negative (leaf) samples.

## 4. EXPERIMENTS AND RESULTS

We manually selected approximate 10% points from each tree as the training data for the machine learning classifiers. These training points are evenly distributed from the bottom to the top of each tree. Consequently, four machine learning classifiers were trained accordingly with different feature sets. The statistical performance indices are summarized in Table 2.

For both trees, RF model resulted in best performance invariably, while NB model gave least accuracy. The accuracy of four classifiers with various feature sets are demonstrated in Figures 4 and 5. Feature sets are selected based on the ranking lists resulted from the *Fisher* filter method (section 3.2). For the *Erythropileum fordii* tree, all classifiers' performances remained similar, indicating little effects from the number of feature used. All four classifiers showed promising results with more than 94% accuracy, which are comparable to e.g., Ma et al. (2016). Although, it is noted that Ma et al. (2016) worked on a more complex and littery scene. For the silver birch tree, the accuracy of the SVM, RF, and GMM model became stable when number of features used reached 10. The NB model performed less well in this study, and its performance reduced with increased feature sets. A similar trend also can be observed from the study of the *Erythropileum*

*fordii* tree, although in this case the trend was weak. The reason may be that the assumption in NB that a particular feature is independent of the value of any other feature was violated when more features were involved. In such a case, the Bayesian Network model (Friedman et al., 1997) will be more suitable. In addition, NB is known to have difficulties when dealing with unbalanced data.

We observed that the high classification accuracy of the *Erytrophleum fordii* might be caused by the fact that the distributions of its stem and crown are essentially very well distinct. To assess the performances of machine learning classifiers in regions where leaf and wood components are heavily interacted, we selected a subset point cloud between 16 and 20 m above ground of the *Erytrophleum fordii* (Figure 6), and ran the experiments on this subset. The results are given in Table 3. The accuracy remained almost identical compared to those from the whole point cloud, indicating that machine learning algorithms can commendably separate leaf and wood components by providing proper training samples.

For the silver birch, calibrated spectral attributes exist. Therefore, leaf and wood can be separated from the spectral information of each point as well. This is based on the fact that different components of a tree feature discriminatory optical properties at the operating wavelengths of the laser scanning system (Tao et al., 2015). In this study, the birch leaf and wood were separated with a hard normalized difference vegetation index (NDVI) threshold value of 0.2. All points that have NDVI value less than 0.2 were labeled as wood components, and vice versa. The accuracy of the spectral method is included in Table 2. The sensitivity (i.e., accuracy for wood identification) is lower than those from machine learning algorithms, mainly because some higher parts of the stem were misclassified as leaves.

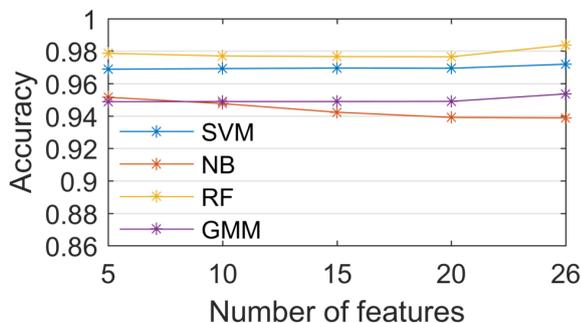


Figure 4. Performance of four classifiers for the *Erytrophleum fordii* as a function of the different feature sets. Feature sets were determined based on the *Fisher* filter method described in section 3.2.

## 5. DISCUSSION

### 5.1 Classifier Performance

As summarized in the Tables 2 and 3, the performances of selected machine learning classifiers are comparable to and surpassing published studies (Ma et al., 2016; Tao et al., 2015). In our tests, RF model produced best results, proving that RF might be very well suitable for wood-leaf classification. This can also be

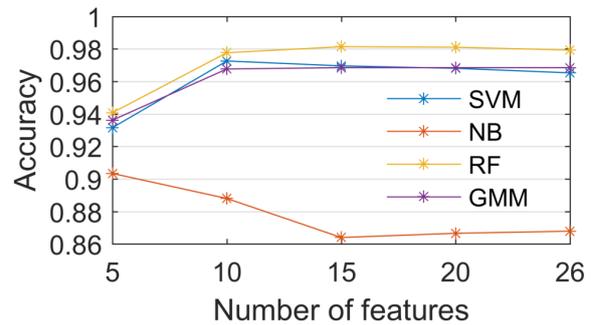


Figure 5. Performance of four classifiers for the silver birch as a function of the different feature sets. Feature sets were determined based on the *Fisher* filter method described in section 3.2.



Figure 6. A crown subset (i.e., 16 - 20 m) of the *Erytrophleum fordii*. Branches and leaves heavily interact.

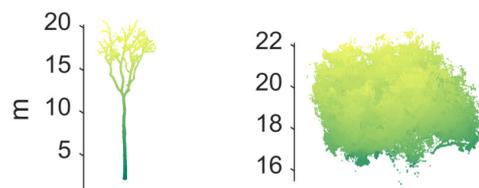


Figure 7. Classification results of RF model for *Erytrophleum fordii*. Left part shows the wood components and right part shows the leaf points.

justified by visualizations of the classification results in this study (Figure 7 and 8). The popular SVM model also gave promising results, however, its model training time were much longer than

Table 3. Statistical evaluation of machine learning classifiers for the crown subset of the *Erytrophleum fordii* with 5, 10, 15, 20, and 26 features.

Features	Classifier	Sensitivity	Specificity	Accuracy
5	SVM	0.93	0.97	0.95
	NB	0.95	0.95	0.95
	RF	0.95	0.98	0.96
	GMM	0.92	0.95	0.94
10	SVM	0.93	0.97	0.95
	NB	0.95	0.92	0.94
	RF	0.96	0.98	0.97
	GMM	0.91	0.96	0.94
15	SVM	0.94	0.98	0.96
	NB	0.96	0.90	0.93
	RF	0.96	0.98	0.97
	GMM	0.91	0.96	0.94
20	SVM	0.94	0.98	0.96
	NB	0.96	0.90	0.93
	RF	0.96	0.98	0.97
	GMM	0.91	0.96	0.94
26	SVM	0.94	0.98	0.96
	NB	0.96	0.90	0.93
	RF	0.96	0.98	0.97
	GMM	0.91	0.96	0.94

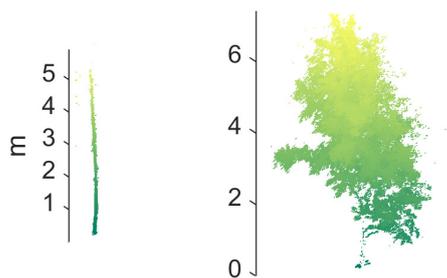


Figure 8. Classification results of RF model for silver birch. Left part shows the wood components and right part shows the leaf points.

others'. NB model performed worse in this study and might not be suitable for leaf-wood separation, unlike its high efficiency in text classification (Kim et al., 2006). GMM model is typically used in unsupervised classification problems (Koo et al., 2014), although it was previously used in separating leaf, wood, and ground points (Ma et al., 2016). We briefly tested the performance of the GMM classifier without training data, so that the data were clustered into two groups in the feature space. We obtained an accuracy of 93% and 91% for the *Erytrophleum fordii* and silver birch, respectively, which are lower than that of the supervised GMM.

## 5.2 Feature Importance

In this study, features were ranked based the *Fisher* filter feature selection method (Table 1). Furthermore, feature sets with different sizes based on the rankings are tested. For both trees, point height and local density seem to be the most vital features, as

they were both ranked as top 5. This indicates that local density characteristics might play a vital role in leaf-wood separation. However, both of them are bound to perform worse in a more complex scene. Commonly used structure inferring features such as linearity and planarity turned out to be less important as they were ranked as non-significant (e.g., 50% in the case of the latter). This can also be justified from the performances of various feature sets. For both trees, the first 10 best features according to the ranking are enough to stabilize the model accuracy, meaning that features such as linearity and planarity are not necessary to be included in such a wood-leaf classification issue. However, we note that feature selection should consider the local tree structure characteristics, such as tree species. In addition, more feature selection approaches should be tested, possibly in connected with the chosen machine learning model. Such methods are known as *wrappers*.

## 5.3 Training Sample Delineation

In this study, training samples were manually and evenly selected from the bottom to the top of each tree. The selected training data take up around 10% of the whole point cloud. In order to assess the influences of training samples, we re-selected a different training sample set with 1m height intervals for the crown subset of the *Erytrophleum fordii* (Figure 6). The re-selected training sample only occupies ~1% of the whole data. The classification results are compared in Table 4. It is noted that the accuracy decreased when less and unevenly distributed training data were used. In particular, model sensitivities reduced drastically, meaning that some wood points were misclassified as leaf points. This implies that the local geometry properties of branch points are not well represented by a small set and vertically spaced training data.

In addition, we trained all classifiers with training data from both trees, meaning that half training data are from the *Erytrophleum fordii* and left are from the silver birch. The results for the *Erytrophleum fordii* remain identical compared to those classifiers trained with only *Erytrophleum fordii* data (Table 5 and Table 2). However, the results for the silver birch are worse, especially in terms of the sensitives, except the RF model. This indicates that the wood parts of the silver birch are severely misclassified as leaf points when using the classifiers trained with a combined training set. RF is immune from this situation, again indicating its efficiency and capability for such as task.

Table 4. Comparison of performances with different training data on the crown subset. Sample *f* denotes the manually selected 10% training set. Sample *s* refers to a training set with 1m height intervals.

Classifier	Sample	Sensitivity	Specificity	Accuracy
SVM	<i>f</i>	0.94	0.98	0.96
	<i>s</i>	0.76	0.96	0.91
NB	<i>f</i>	0.96	0.90	0.93
	<i>s</i>	0.89	0.84	0.85
RF	<i>f</i>	0.96	0.98	0.99
	<i>s</i>	0.84	0.96	0.93
GMM	<i>f</i>	0.91	0.96	0.94
	<i>s</i>	0.68	0.97	0.89

## 6. CONCLUSION

In this study, we compared four machine learning algorithms, namely Support Vector Machine, Naïve Bayes, Random Forest, and Gaussian Mixture Model, for separating wood and leaf points

Table 5. Performances of all classifiers trained with a combined training set from both trees.

Data set	Classifier	Sensitivity	Specificity	Accuracy
Birch	SVM	0.67	1.00	0.98
	NB	0.48	1.00	0.97
	RF	0.88	1.00	0.99
	GMM	0.53	1.00	0.97
<i>E. fordii</i>	SVM	0.90	0.90	0.95
	NB	0.94	0.90	0.92
	RF	0.95	1.00	0.97
	GMM	0.89	0.96	0.93

from TLS data. In general, there is a lack of comparative studies of machine learning algorithms for such problems. Our study highlighted the feasibility of the methodology. Specifically, two trees were tested, an *Erytrophleum fordii* and a silver birch. Twenty-six geometry-based features were extracted and individually ranked by a filter feature selection method. Various feature sets and training data were tested. Our results show that machine learning algorithms can efficiently separate wood and leaf point from TLS data with an accuracy of, in general, more than 95%. Evenly distributed training data are recommended, as sparse training data can reduce the classification accuracy especially for branches inside the tree crown. It is noted that our studies were performed on purer data sets. More tests on tree data from more complex natural conditions should be carried out in the future. In addition, more tree species should be tested.

#### ACKNOWLEDGEMENTS

The study was supported by the project "The influence of Biomass and its change on landSLIDE activity" (BioSLIDE) within the research program Earth System Sciences (ESS) of the Austrian Academy of Science (Österreichische Akademie der Wissenschaften, ÖAW) and financed by the Federal Ministry of Science and Research (BMWF). The authors also thank the Austrian Research Promotion Agency (FFG) for providing financial support via the project no. 860021. Special thanks are due to three referees for their constructive reviews that improved the manuscript significantly.

#### REFERENCES

Béland, M., Baldocchi, D. D., Widlowski, J.-L., Fournier, R. A. and Verstraete, M. M., 2014. On seeing the wood from the leaves and the role of voxel size in determining leaf area distribution of forests with terrestrial lidar. *Agricultural and Forest Meteorology* 184, pp. 82–97.

Breiman, L., 2001. Random forests. *Machine learning* 45(1), pp. 5–32.

Brodu, N. and Lague, D., 2012. 3d terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: Applications in geomorphology. *ISPRS Journal of Photogrammetry and Remote Sensing* 68, pp. 121–134.

Calders, K., Disney, M. I., Armston, J., Burt, A., Brede, B., Origo, N., Muir, J. and Nightingale, J., 2017. Evaluation of the range accuracy and the radiometric calibration of multiple terrestrial laser scanning instruments for data interoperability. *IEEE Transactions on Geoscience and Remote Sensing* 55(5), pp. 2716–2724.

Friedman, N., Geiger, D. and Goldszmidt, M., 1997. Bayesian network classifiers. *Machine learning* 29(2-3), pp. 131–163.

Geiß, C., Pelizari, P. A., Marconcini, M., Sengara, W., Edwards, M., Lakes, T. and Taubenböck, H., 2015. Estimation of seismic building structural types using multi-sensor remote sensing and machine learning techniques. *ISPRS Journal of Photogrammetry and Remote Sensing* 104, pp. 175–188.

Gu, Q., Li, Z. and Han, J., 2012. Generalized fisher score for feature selection. *arXiv preprint arXiv:1202.3725*.

Guyon, I., Gunn, S., Nikravesh, M. and Zadeh, L. A., 2008. Feature extraction: foundations and applications. Vol. 207, Springer.

Hackenberg, J., Wassenberg, M., Spiecker, H. and Sun, D., 2015. Non destructive method for biomass prediction combining tIs derived tree volume and wood density. *Forests* 6(4), pp. 1274–1300.

Hakala, T., Suomalainen, J., Kaasalainen, S. and Chen, Y., 2012. Full waveform hyperspectral lidar for terrestrial laser scanning. *Optics express* 20(7), pp. 7119–7127.

Kankare, V., Holopainen, M., Vastaranta, M., Puttonen, E., Yu, X., Hyypä, J., Vaaja, M., Hyypä, H. and Alho, P., 2013. Individual tree biomass estimation using terrestrial laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing* 75, pp. 64–75.

Kim, S.-B., Han, K.-S., Rim, H.-C. and Myaeng, S. H., 2006. Some effective techniques for naive bayes text classification. *IEEE transactions on knowledge and data engineering* 18(11), pp. 1457–1466.

Koo, S., Lee, D. and Kwon, D.-S., 2014. Unsupervised object individuation from rgb-d image sequences. In: *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on, IEEE*, pp. 4450–4457.

Levick, S. R., Hessenmöller, D. and Schulze, E.-D., 2016. Scaling wood volume estimates from inventory plots to landscapes with airborne lidar in temperate deciduous forest. *Carbon Balance and Management* 11(1), pp. 7.1 – 7.14.

Li, Z., Douglas, E., Strahler, A., Schaaf, C., Yang, X., Wang, Z., Yao, T., Zhao, F., Saenz, E. J., Paynter, I. et al., 2013. Separating leaves from trunks and branches with dual-wavelength terrestrial lidar scanning. In: *Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International, IEEE*, pp. 3383–3386.

Liang, X., Kankare, V., Hyypä, J., Wang, Y., Kukko, A., Haggren, H., Yu, X., Kaartinen, H., Jaakkola, A., Guan, F. et al., 2016. Terrestrial laser scanning in forest inventories. *ISPRS Journal of Photogrammetry and Remote Sensing* 115, pp. 63–77.

Liang, X., Litkey, P., Hyypä, J., Kaartinen, H., Vastaranta, M. and Holopainen, M., 2012. Automatic stem mapping using single-scan terrestrial laser scanning. *Geoscience and Remote Sensing, IEEE Transactions on* 50(2), pp. 661–670.

Ma, L., Zheng, G., Eitel, J. U., Moskal, L. M., He, W. and Huang, H., 2016. Improved salient feature-based approach for automatically separating photosynthetic and nonphotosynthetic components within terrestrial lidar point cloud data of forest canopies. *Geoscience and Remote Sensing, IEEE Transactions on* 54(2), pp. 679–696.

Marcot, B. G., Steventon, J. D., Sutherland, G. D. and McCann, R. K., 2006. Guidelines for developing and updating bayesian belief networks applied to ecological modeling and conservation. *Canadian Journal of Forest Research* 36(12), pp. 3063–3074.

Olofsson, K., Holmgren, J. and Olsson, H., 2014. Tree stem and height measurements using terrestrial laser scanning and the ransac algorithm. *Remote Sensing* 6(5), pp. 4323–4344.

Pfennigbauer, M. and Ullrich, A., 2010. Improving quality of laser scanning data acquisition through calibrated amplitude and pulse deviation measurement. In: SPIE Defense, Security, and Sensing, International Society for Optics and Photonics, pp. 76841F–76841F.

Puttonen, E., Briese, C., Mandlbürger, G., Wieser, M., Pfennigbauer, M., Zlinszky, A. and Pfeifer, N., 2016. Quantification of overnight movement of birch (*Betula pendula*) branches and foliage with short interval terrestrial laser scanning. *Frontiers in Plant Science* 7, pp. 222.1 – 222.13.

Tao, S., Guo, Q., Su, Y., Xu, S., Li, Y. and Wu, F., 2015. A geometric method for wood-leaf separation using terrestrial and simulated lidar data. *Photogrammetric Engineering & Remote Sensing* 81(10), pp. 767–776.

Vapnik, V., 1995. *The Nature of Statistical Learning Theory*. Springer.

Vauhkonen, J., Hakala, T., Suomalainen, J., Kaasalainen, S., Nevalainen, O., Vastaranta, M., Holopainen, M. and Hyypä, J., 2013. Classification of spruce and pine trees using active hyperspectral lidar. *IEEE Geoscience and Remote Sensing Letters* 10(5), pp. 1138–1141.

Wagner, W., Hollaus, M., Briese, C. and Ducic, V., 2008. 3d vegetation mapping using small-footprint full-waveform airborne laser scanners. *International Journal of Remote Sensing* 29(5), pp. 1433–1452.

Wang, D., Hollaus, M., Puttonen, E. and Pfeifer, N., 2016. Automatic and self-adaptive stem reconstruction in landslide-affected forests. *Remote Sensing* 8(12), pp. 974.1 – 974.23.

Wang, D., Kankare, V., Puttonen, E., Hollaus, M. and Pfeifer, N., 2017. Reconstructing stem cross section shapes from terrestrial laser scanning. *IEEE Geoscience and Remote Sensing Letters* 14(2), pp. 272–276.

Weinmann, M., Jutzi, B. and Mallet, C., 2013. Feature relevance assessment for the semantic interpretation of 3d point cloud data. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences II-5/W2*, pp. 313 – 318.

Weinmann, M., Urban, S., Hinz, S., Jutzi, B. and Mallet, C., 2015. Distinctive 2d and 3d features for automated large-scale scene analysis in urban areas. *Computers & Graphics* 49, pp. 47–57.

Weinmann, M., Weinmann, M., Mallet, C. and Brédif, M., 2017. A classification-segmentation framework for the detection of individual trees in dense mms point cloud data acquired in urban areas. *Remote Sensing* 9(3), pp. 277.1 – 277.28.

Yun, T., An, F., Li, W., Sun, Y., Cao, L. and Xue, L., 2016. A novel approach for retrieving tree leaf area from ground-based lidar. *Remote Sensing* 8(11), pp. 942.1 – 942.21.

Zheng, G., Moskal, L. M. and Kim, S.-H., 2013. Retrieval of effective leaf area index in heterogeneous forests with terrestrial laser scanning. *IEEE Transactions on Geoscience and Remote Sensing* 51(2), pp. 777–786.