WOOD-LEAF UNSUPERVISED CLASSIFICATION OF SILVER BIRCH TREES FOR BIOMASS ASSESSMENT USING OBLIQUE POINT CLOUDS

WOOD-LEAF UNSUPERVISED CLASSIFICATION OF SILVER BIRCH TREES FOR BIOMASS ASSESSMENT USING OBLIQUE POINT CLOUDS

C. Spadavecchia [1]*, M. B. Campos [2], M. Piras [1], E. Puttonen [2], A. Shcherbacheva [2],

1 DIATI, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy - (claudio.spadavecchia, marco.piras)@polito.it
2 Department of Photogrammetry and Remote Sensing, Finnish Geospatial Research Institute, National Land Survey of Finland, 02150 Espoo, Finland - (mariana.campos, eetu.puttonen, anna.shcherbacheva)@nls.fi

KEY WORDS: Remote Sensing, Machine Learning, Wood-Leaf Classification, Biomass Assessment, LiDAR Point Clouds.

ABSTRACT

Forests play a fundamental role in carbon stocking since about a third of the carbon dioxide produced by activities of human origin is absorbed by forests. Forest biomass is an essential indicator of carbon dioxide absorption, enabling an understanding the interaction between forest dynamics and climate change effects. However, biomass and wood material changes are challenging to quantify in forest stands. Nowadays, recent 3D remote sensing technologies, such as laser scanning systems, have allowed accurate measures of single trees. This study evaluates three approaches to classify wood and non-wood materials and quantify biomass based on LiDAR data, aiming at biomass change detection. Specifically, we propose an automated methodology for estimating the single tree-level biomass of a portion of forest monitored through a LiDAR oblique acquisition. The classification of wood and foliage points was performed with machine learning algorithms, while the tree modelling was conducted rigorously through a Quantitative Structure Model (QSM). The purpose of this study is to evaluate (1) two different unsupervised and one semi-supervised classification approaches for wood and foliage separation and (2) the accuracy of the biomass assessment performed on a QSM-based approach on innovative LiDAR acquisitions. The results are promising; the wood-leaf classification performs effectively in all 20 silver birches considered; as regards the biomass, when the noise is limited, it is estimated in a manner consistent with the reference values calculated using an appropriate allometric equation. Higher values are found mainly in dense undergrowth, which negatively affects the modelling of the tree. The research is ongoing, and further tests will be performed to generalize the methodology on different tree species, deepen the multitemporal variability, and improve the accuracy of the assessment.

1. INTRODUCTION

Forests play a fundamental role in tackling climate change, as they directly absorb large quantities of carbon dioxide through the photosynthesis of chlorophyll. It is estimated that about a third of the carbon dioxide produced by activities of human origin is absorbed by forests (*Global Forest Resources Assessment 2020*, 2020). Traditionally, carbon stock capacity measures are based on the relationship between aboveground biomass (AGB) and stored carbon. To deal with climate change is necessary to safeguard biodiversity and protect the Earth's natural ability to absorb CO_2 from the atmosphere. For this reason, the scientific community is trying to propose a well-defined methodology to accurately monitor the forest biodiversity and estimate the biomass at the single tree level, mainly focusing on automatic, efficient and time-saving alternatives (Xu et al., 2021).

Usually, biomass can be measured directly on a tree after harvesting it or by using allometric equations, which estimate the biomass as a function of more easily measurable parameters such as the tree's height and density and the trunk's diameter (Diameter at Breast Height). The former method is the most rigorous, but simultaneously, it is highly time-consuming, destructive, irreversible and with high ecosystem impact. Alternatively, methodologies involving allometric equations and non-destructive measures are a faster and more ecological solution facilitating its repetition over time to monitor changes. Non-destructive measures of the main tree parameters (e.g. DBH and height) can be taken through direct in-situ measurements (e.g. calipers) or, even more detailed, using remote sensing techniques such as MLS (Mobile Laser Scanning) or TLS (Terrestrial Laser Scanning). However, allometric equations are not helpful when the tree suffers minor damage (i.e., the loss of some smaller branches), as these usually do not change the DBH or the height (typically, allometric equations are expressed in functions of these quantities). Because of this, using an allometric equation to estimate the biomass and evaluate its evolution over time is not the optimal solution.

Laser scanner technology enables a full 3D representation of trees structure, providing an alternative to allometric equations for accurately estimating AGB in stand trees. Since AGB estimation needs to be performed at the tree level, in literature, several algorithms that perform Individual Tree Detection (ITD) are proposed (Luo et al., 2021; Shendryk et al., 2016; Xu et al., 2023), differing according to the support (e.g., terrestrial, aerial), the type of data (e.g., photogrammetric, LiDAR point cloud) and the approach (e.g., voxel-based, raster-based, etc.). Furthermore, a correct calculation of the AGB requires that the points of the cloud are classified into woody and foliage classes to proceed with estimating the woody biomass only (Arseniou et al., 2023). In literature, several studies propose methodologies for wood-leaf classification (Sun et al., 2021; Vicari et al., 2019; Wang et al., 2020).

This study proposes a complete methodology for estimating tree biomass in forest environments. It starts from the acquisition and preprocessing of raw LiDAR point cloud acquired from a permanent laser scanning setup (Campos et al., 2021); it then continues with the segmentation at the single-tree level, the wood-leaf classification, and the biomass estimation based on a Quantitative Structure Model (QSM). More in detail, we tested three different approaches to classify wood and foliage points. We considered the point cloud's geometric and scanner-measured

radiometric features to perform the wood and leaf point cloud classification.

The assessment was validated with reference values calculated with an allometric equation precisely calibrated for Finnish trees (Kangas et al., 2023), assuming that the area under investigation was not subject to major disastrous events that could have seriously affected the main structure of the trees.

2. MATERIALS AND METHODS

2.1 Dataset

The study area is located in the Hyytiälä forest field station, a recognized forest educational and research center focused on long-term boreal forest zone monitoring in southern Finland. Hyytiälä forest is characterized by a predominance of coniferous boreal species, including Scots pine (Pinus sylvestris L.) and Norway spruce (Picea abies H. Karst), as well as deciduous trees, such as silver birch (*Betula pentula Roth*). Among the several data collections conducted in this research forest, long-term field measurement sites, soil and atmospheric sampling and continuous LiDAR time-series data collection can be highlighted. LiDAR data is acquired by a permanent laser scanning setup named LiDAR phenological station (LiPhe) (Campos et al., 2021). LiPhe consists of a RIEGL VZ-2000i laser scanner installed on the top of a 35-meter-high tower and tilted 60 degrees. This setup provides an oblique and above-canopy view of the forest, enabling a high spatial-resolution reconstruction of the scanned trees. The point cloud resolution is at least 0.01 m spacing between two neighboring points at a 100 m range. LiPhe point clouds attributes are point return number (scalar from 1 to 15), the number of returns (scalar from 1 to 15), the intensity (Amplitude in dB), the scan angles (theta and phi in degrees), the reflectance (dB), the return pulse deviation (a measure of pulse shape distortion) and the range (m).

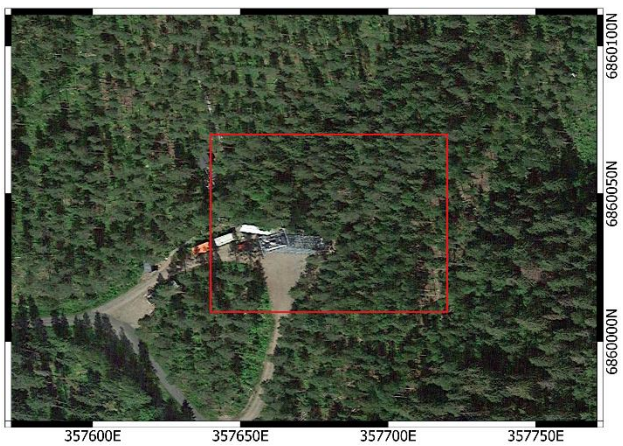

Figure 1. Study area (EPSG: 3067).

Figure 1 shows the scanned area covered by LiPhe at Hyytiälä forest, along with the tower where LiPhe is installed and the position of the studied trees. This study focuses on 20 silver birch trees within 60 m from the laser scanner system monitored at a one-time point in April 2020. In this period, the presence of snow occurring in the winter period was avoided, but at the same time the leaf-off season is still ongoing, resulting in a limited number of leaves; the wood detection is therefore facilitated. The data selection was performed considering only spatially scattered trees that were correctly and automatically identified. Silver birches were chosen among all the species since it is one of the most widespread species in the Finnish national territory and northern Europe; focusing on a single tree species allowed the validation of the workflow and the evaluation of its accuracy while avoiding species-specific discrepancies.

2.2 Data processing

Figure 2 shows a summary of the workflow performed in this work, which consists of data collection, pre-processing, data analysis and validation of the results.

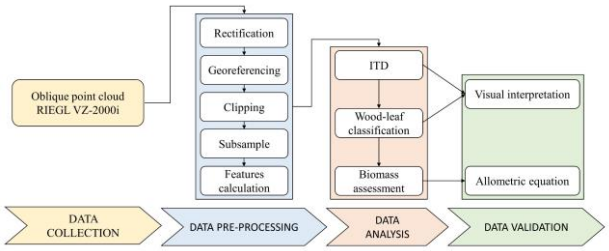

Figure 2. Workflow applied in this study.

Figure 2 graphs the workflow of this study described in the following subsections.

2.2.1 Pre-processing

The initial step involved in the data pre-processing workflow was rectifying and georeferencing the raw oblique point cloud, which comprised approximately 608 million points within an area of 41,500 m^2. LiPhe point clouds have the origin at the scanner position, resulting in the original point cloud tilted with respect to the ground. Therefore, an initial rectification is required to normalize the point cloud local reference system to the ground. Subsequentially, a Helmert 3D transformation was applied to transform the point clouds from a local to a georeferenced reference system, ensuring the correspondence of the point cloud data with the real-world coordinates and other reference data. Here, we focus on the LiPhe monitoring area with higher spatial resolution comprising 60 m from the scanner. The interested area is a subset of LiPhe total field of view and it covers 4800 m^2, 256 trees and 86 million points (**Figure 3**).

To normalize the point cloud resolution and reduce the computational cost, the point cloud was subsampled by a voxelization approach that selected the closest point to the voxel centroid, considering one cm-sized voxel. The selection of the voxel size was carefully performed to balance computational efficiency while simultaneously preserving a high level of detail necessary for the in-depth analysis of tree biomass. Finally, the commercial Cloud Compare software was used to add geometric features (density, verticality, anisotropy, linearity) to the point cloud.

2.2.2 Individual Tree Detection

ITD was performed with the PyCrown algorithm (Zörner et al., 2018), which was assessed in previous studies (Spadavecchia et al., 2022; Spiekermann et al., 2021). This method is based on the location of local maxima which is labelled as an 'initial region' around which a tree crown can grow; the heights of the four neighboring pixels are extracted from the CHM and these pixels are added to the region if (1) their vertical distance from the local maximum is less than some user-defined percentage of the local maximum height, and (2) less than some user-defined maximum

difference (Dalponte and Coomes, 2016). The tuning of the user-defined parameters was done iteratively according to the tree species and their dimensions; the threshold (1) was set equal to 0.45 and the threshold (2) equal to 0.55. The algorithm also requires the setting of a threshold (3) below which a pixel cannot be a tree (set equal to 2) and a threshold (4) which defines the maximum value of the crown diameter of a detected tree (set to 15 pixels). The main output of the algorithm is a segmented point cloud in which a unique label identifies every single tree. A single point cloud for each tree has been generated to simplify subsequent analyses.

After segmenting the point cloud at a single tree level, its accuracy was evaluated according to the procedure described in the validation process subsection. Among the trees that were correctly identified, for the study, it was decided to focus on estimating the biomass of 20 silver birches. These were randomly selected but in such a way as to be spatially scattered within the study area.

2.2.3 Wood-leaf classification

Table 1 summarizes the experiments performed testing two unsupervised and one semi-supervised machine learning classification approaches for wood and foliage separation with different point cloud features.

Experiment	Method	Features
1	(1) K-means	Intensity, reflectance, density
2	(1) K-means	Intensity, reflectance, density, deviation
3	(2) Threshold+K-means	Intensity, reflectance, density
4	(2) Threshold+K-means	Intensity, reflectance, density, deviation
5	(3) kNN+ADASYN	Anisotropy, linearity, verticality

Table 1. Summary of the experiments carried on in this study.

The first method relies on the K-means clustering algorithm. It starts with a random selection of *k* clusters and calculates their centroids; then, it performs iterative calculations to optimize the clusters concerning the positions of the centroids.

The second method is based on two steps. Initially, a classification is made according to a threshold defined as the median of the best-fitting distribution of the point cloud selected features. The Logistic distribution, the Normal distribution, and the Laplacian distribution were considered, and the features on which these distributions have been fitted are the radiometric ones (intensity and reflectance). The second step coincides with the first method, and a K-means clustering algorithm has been applied.

The third method is kNN-based, for which two features were calculated: (1) anisotropy and (2) linearity at different scales. At each scale corresponding to different kNN values, points with anisotropy greater than 0.95 and linearity exceeding 0.75 as wooden parts were labelled as wooden parts. Simultaneously, using the radius-based neighborhood approach, also all points with verticality over 0.99 were labelled as wooden parts. Regarding kNN-based linearity, the main branches of various sizes and some secondary branches were separated. In the case of anisotropy and verticality, the focus was on separating stem parts, fragments of upper branches, and branch segments tightly attached to the stem. It is worth noting that verticality proved to be a powerful feature for efficiently detecting the bottom parts of the stems (Wang et al., 2020). After the initial separation, the wooden component typically contained more points than the foliage component. This could pose a class imbalance issue in supervised classification based on automatically generated training data, often resulting in poor model predictions. To address this, the adaptive minority oversampling method ADASYN (Haibo He et al., 2008) was applied. It is commonly used to handle class imbalance in binary classification problems. This method artificially generates more wooden parts in 3D space (x,y, and z coordinates) from the point cloud, based on a kNN classifier and the relative positions of the wooden and foliage clusters. Essentially, ADASYN retrieves additional parts of branches and stems misclassified by the kNN-based method. It aims to complement instances of the minority class located close to the boundaries separating the two classes, where the minority class is more mixed with the majority class and frequently misclassified. The artificially generated samples are used in a label transfer procedure, where their labels are assigned to the original points in the point cloud that are neighbors with them. ADASYN is applied over multiple iterations, incrementally increasing the wooden component. However, it should be noted that there is no predefined stopping criterion for this process. The method was tested with manually labelled TLS data to determine the optimal number of iterations. It was found that 7-10 iterations yielded the best results, with maximum producer accuracy (the percentage of successfully retrieved wooden components) over 85%, and user accuracy (the percentage of correctly detected wooden parts) over 80%. 10 iterations were applied to the present dataset of 20 trees.

Following these elaborations, five classified clouds were obtained for each tree.

2.2.4 Aboveground biomass assessment and validation

The points describing the woody part of the tree have been transformed into local coordinates. This procedure was necessary to lighten the subsequent processing. The biomass calculation was performed using the TreeQSM algorithm (Raumonen and Åkerblom, 2022) written in MATLAB environment, which reconstructs Quantitative Structure Models for trees from point clouds. Specifically, QSM is a hierarchical geometric primitive model accurately approximating the tree branching structure, geometry, and volume (Åkerblom et al., 2017). As described in the algorithm help documentation, different results can be obtained by running the model several times due to random elements in the reconstructive process. Therefore, it was decided to consider the average of 10 results thus obtained. For more information about the model, please refer to the reference article (Raumonen and Åkerblom, 2022). Although the algorithm allows getting several outputs, we only considered the wood volume and biomass in this study.

To validate the results obtained exhaustively, it was decided to carry out a validation that considers both the segmentation procedure and the biomass estimation.

The segmentation procedure was validated using the F1 score parameter, widely considered in this kind of analysis (Belcore et al., 2020; Spadavecchia et al., 2022). This parameter relates the

number of correctly segmented trees (Number of Matches, NM) to the total number of segmented trees (Defined Trees, DT) and the total number of trees actually present in the study area (Reference Trees, RT) as follows:

$$Producer's\ Accuracy\ (PA) = \frac{NM}{RT}, \quad (1)$$

$$User's\ Accuracy\ (UA) = \frac{NM}{DT}, \quad (2)$$

$$F1\ score = \frac{2 * PA * UA}{PA + UA}, \quad (3)$$

The F1 score can have values between 0 and 1, and higher values correspond to better segmentation.

The reference values through which the accuracy of the biomass estimate was evaluated were calculated using an allometric equation. Equation (4) has been fine-tuned to best fit the main tree species in Finland (Kangas et al., 2023):

$$AGB = logit^{-1}\left(a + b * DBH + c * h + d * \frac{1}{h} + e * DBH * h + f * \frac{1}{DHB * h} + \varepsilon_{\text{cluster}} + \varepsilon_{\text{plot}}\right) * \frac{\pi * DSH^2 * h}{40} + \varepsilon_{\text{trees}} \quad (4)$$

Where DBH [cm] is the Diameter at the Breast Height (1.30 m), the DSH [cm] is the Diameter at the Stump Height, and h is the tree height [m]. Parameters a-f were tuned from a Finland-wide tree dataset; parameters a-c depend on the temperature sum, while d-f differ for each tree species. $\varepsilon_{\text{cluster}}$, $\varepsilon_{\text{plot}}$, and $\varepsilon_{\text{tree}}$ are, respectively, the zero-mean cluster-, plot- and tree-level random effects that follow the standard assumption of mixed-effects models (Mehtätalo and Lappi, 2020). More details about the parameters are described in (Kangas et al., 2023). The tree height has been manually measured from the CHM, while the DBH has been calculated as the diameter of a cylinder that best fits the points of the trunk between 1.1 m and 1.5 m above the ground. Finally, the DSH has been estimated as follows:

$$DSH = \frac{h}{h - 1.3} * DBH \quad (5)$$

3. RESULTS

3.1 Individual Tree Detection

The segmentation procedure identified a total of 240 trees. Of these, 150 were identified correctly, i.e., free of under-segmentation or over-segmentation errors. Under-segmentation errors occur when two or more trees are segmented as a single tree; on the opposite, over-segmentation is when a single tree is segmented as two or more trees. In this case, 47 trees were under-segmented, while 40 trees were over-segmented. **Figure 3** shows the resulting point cloud. The resulting F1 score value is equal to 0.60.

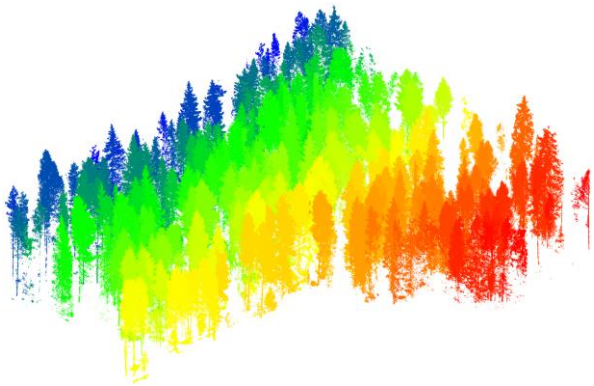

Figure 3. Individual trees.

The trees selected for subsequent analysis are shown in **Figure 4**. They are uniquely silver birches and are spatially distributed within the study area.

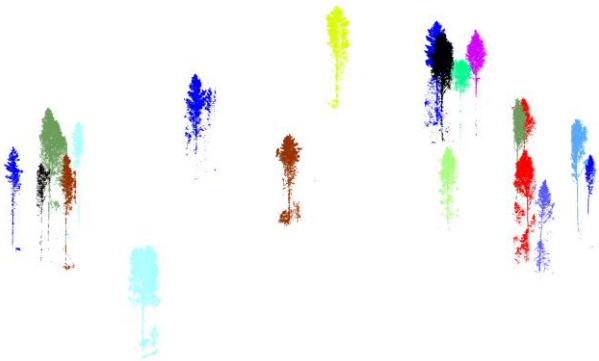

Figure 4. 20 individual silver birches selected.

3.2 Wood-leaf classification

Figure 5 shows the result of the classification of a tree which has been taken as an example for visualization. Based on visual inspection, all the experiments return an excellent classification. The radiometric feature of the deviation only partially affects the final result. Experiments 1-2, and experiments 3-4 do not show significant differences. Among them, it can be observed that the woody points are correctly identified in greater numbers in the second approach. Instead, with the third method, almost all the woody points are identified in the lower part of the tree. All methods show greater difficulty locating and correctly detecting the smaller trunks at the top of the tree. However, even if there is noise, with the first and second approaches, it is possible to identify some branches that are not detected with the third approach.

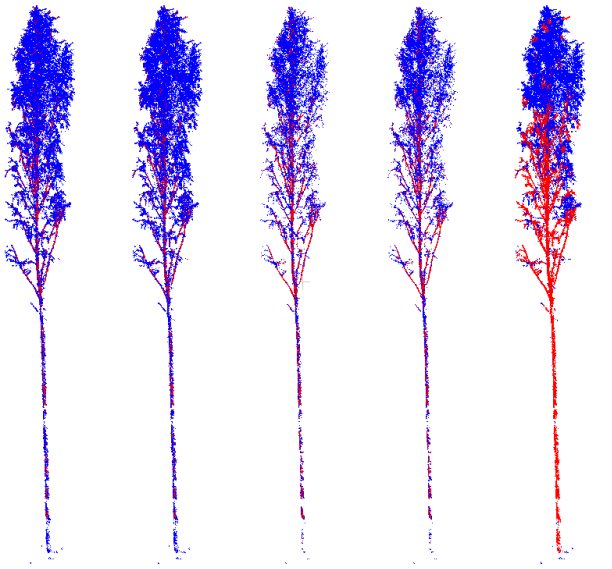

Figure 5. Classified tree: Experiment 1-5 from left to right.

In **Figure 6**, only the wood points of the cloud are shown. It is seen more clearly that more filtering is performed with the second and third method; in particular, with the third one, some woody parts disconnected from the rest of the tree structure are identified. However, experiment 5 identifies a denser wood structure: this helps the tree modelling algorithm to obtain more reliable biomass estimates.

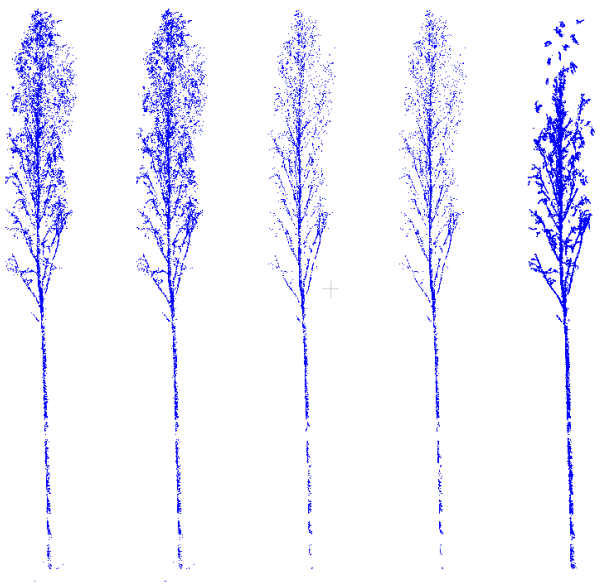

Figure 6. Wood points: Experiment 1-5 from left to right.

3.3 Aboveground biomass assessment

Table 2 summarizes each tree's measured (DBH and height) and estimated (DSH, AGB) allometric parameters.

The biomass estimation through the quantitative structure model algorithm shows promising results for about half of the trees considered. In contrast, in the case of the remaining half, the values are entirely different from the reference ones (**Table 3**).

Tree	DBH [cm]	h [m]	DSH [cm]	AGB [L]
1	17.7	22.95	18.7	280
2	27.7	21.21	29.5	614
3	29.4	26.11	30.9	808
4	28.8	20.59	30.8	646
5	25.8	21.03	27.6	536
6	22.0	21.85	23.4	408
7	24.1	23.07	25.6	508
8	22.8	20.79	24.3	418
9	19.9	18.38	21.5	293
10	21.5	20.93	22.9	376
11	15.2	19.71	16.3	184
12	28.7	16.5	31.2	538
13	32.7	17.5	35.3	720
14	19.0	17	20.6	249
15	25.3	16.5	27.4	421
16	18.8	19	20.2	269
17	20.2	20	21.6	321
18	21.0	20.35	22.4	350
19	15.7	16.52	17.0	168
20	15.7	16.63	17.0	169

Table 2. Reference each tree's allometric parameters (DBH, h, DSH, AGB).

Tree	Exp 1 [L]	Exp 2 [L]	Exp 3 [L]	Exp 4 [L]	Exp 5 [L]	Allometric model [L]
1	25352	22907			170	280
2	810	789	780	750	111	614
3	885	890	921	870	897	808
4	31102	37831	1663	1791	926	646
5	8938	10036	2255	1976	344	536
6	17105	5212	1215	1878	292	408
7	1295	1467	2120	2016	564	508
8	14433	12733	2459	3124	604	418
9	1994	1026	847	756	716	293
10	1696	1718	541	581	440	376
11	243	240	230	250	101	184
12	16902	18726	32805	6719	66	538
13	1160	1110	1093	1112		720
14	2034	1560	9692	10645		249
15	1509	1151			116	421
16	318	319	228	212	217	269
17	1119	1043	24105	8345	455	321
18	369	384	296	295	248	350
19			120	140	30	168
20	258	230	190	201	180	169

Table 3. Aboveground biomass [litres] for each experiment compared with the reference values.

As can be seen from **Table 4**, for trees whose biomass has been estimated appropriately, the percentage errors obtained are

consistent with the allometric equation results. Most estimates differ from the AGB calculated with allometric equations by about ± 25%, while some outliers reach differences equal to a maximum 60%. Furthermore, similar results are observed between methods 1 and 2; this is also true when comparing experiments 1-2 and 3-4. Unfortunately, in some cases the algorithm cannot reconstruct a model at all. Therefore, a biomass estimate is not available in these cases. Trees whose biomass was not correctly estimated are always greatly overestimated due to incorrect tree modelling due to undergrowth or other noisy points.

Tree	Exp 1 [%]	Exp 2 [%]	Exp 3 [%]	Exp 4 [%]	Exp 5 [%]
1	8954	8081			-39
2	32	28	27	22	-82
3	10	10	14	8	11
4	157	177	4712	5753	43
5	1568	1773	321	269	-36
6	197	360	4089	1176	-28
7	155	189	317	297	11
8	489	648	3357	2950	45
9	581	251	189	159	145
10	351	357	44	55	17
11	32	30	25	36	-45
12	3041	3380	5996	1148	-88
13	61	54	52	55	
14	716	526	3788	4171	
15	258	173			-72
16	18	19	-15	-21	-19
17	249	225	7416	2502	42
18	5	10	-16	-16	-29
19			-28	-16	-82
20	52	36	12	19	7

Table 4. Error percentage of the estimated aboveground biomass compared to the reference values. The cells highlighted in green refer to correct estimations.

4. DISCUSSION

The automatic biomass estimation over a large area conducted in this study highlights the potential and weaknesses of this approach. The innovative methodology of acquisition allows monitoring an extended portion of the forest over time, in a time-efficient manner and without the need for the presence of an operator during data collection. The need to use artificial intelligence algorithms for data processing is primary; in this way, it is possible to fully exploit the acquired data and extract additional information effectively and automatically. On the other hand, this approach requires a suitable structure; it is essential to use or possibly build adequate support positioned in a study area of particular interest. Moreover, by its nature, this methodology only allows a partial description of the scenario; due to a single observation point, the number of obstacles (trunks and branches) blocking the laser increases as the distance from the station point increases. This aspect is partially limited by the high density of the point cloud, facilitating the interpolation of missing data.

The automatic segmentation method chosen allowed individual trees to be identified with a percentage of success consistent with those described in previous studies (Ma et al., 2022; Qiu et al., 2020). Since the oblique acquisition of point cloud is an innovative methodology, to the best of our knowledge, no single tree-level segmentation algorithm in the literature was developed specifically for this type of data. The PyCrown algorithm was developed for segmenting aerial point clouds; however, since it associates the local maxima of the CHM (Canopy Height Model, calculated as the difference of the DSM with the DTM) with the treetops, it was supposed that it could be equally effective in this case study as well. As a matter of fact, the position of the laser scanner allows a view of the treetops, but at the same time, it can acquire points of the terrain for the generation of the DTM. It has also been observed that the segmentation errors are mainly localized in higher-density areas (in which trees are closest to each other, whose crown intersects each other, and in which shorter trees are dominated by surrounding trees) and distant from the LiDAR. In fact, the more you move away from the laser scanner, the more the angle of incidence of the beam decreases, and the density of points that describes the upper part of the trees decreases as well.

The classification of forest point clouds into wood and is an open topic. In this study it was decided to apply different approaches to evaluate their efficiency and identify the most suitable one to allow an accurate biomass estimation. The features were selected iteratively, identifying the best ones in density, intensity, reflectance, and deviation. Contrary to experiment 5, based on approach 3, deviation does not play a significant role in the first two approaches. With all the methodologies, the results were validated according to visual interpretation. Experiment 5 returns the most stringent classification, so a more accurate biomass estimation is expected when considering these point clouds. Very similar results are obtained through the first two methodologies; in experiments 3-4 the woody element receives fewer points at the expense of the foliage; however, it is still possible to visually locate the trunks and the larger branches.

The estimation of biomass through non-invasive techniques is currently intensely debated. One of the most common techniques used in this study is based on modelling the tree through a Quantitative Structure Model, which allows calculating its volume with a good approximation. Its limitation lies in providing an optimally segmented point cloud at the single tree level as input data. When the point cloud has too many points considered noise, such as dense undergrowth, this causes errors in the generation of the model, which very often leads to a significant overestimation of the volume. In fact, the undergrowth is seen as part of the trunk, which in this way is modelled with a diameter that is more incorrect the larger and more spatially distributed the undergrowth is. Nevertheless, when this phenomenon does not occur the biomass is estimated as consistent with the values calculated through the allometric equation. Through approaches 1 and 2, the biomass estimation is correctly performed for 8 trees out of 20; when approach 3 (experiment 5) is considered, there are 14 trees for which an estimate consistent with the reference is obtained. The first two approaches tend to overestimate the biomass, probably due to some points related to the foliage that have not been classified as such; in experiment 5, on the other hand, both overestimation (up to 45%) and underestimation (up to -45%) occur. In some cases (4 trees in different experiments) it was not possible to estimate the biomass due to errors in the generation of the tree structural model; this is probably due both to the presence of noise and to a wood point cloud which sometimes (particularly in the upper part of the tree) lacks the necessary continuity. The biomass of the

remaining trees has been misestimated, especially in experiments 1 and 2 it needs to be more accurately estimated. In these cases the algorithms failed to exclude the undergrowth, which in this way was modelled causing the generation of trees whose main trunk had an unrealistic diameter. In general, since the classification carried out in experiment 5 is the most rigorous, better results are obtained. Additionally, experiment 5 could be potentially improved by the initial separation of the point clouds based on flexible threshold values selected as infliction points of the fitted geometric feature distributions (such as linearity, anisotropy, and verticality at multiple spatial scales). Also the choice of the relevant neighborhood scales can be optimized more efficiently to reflect the point cloud geometry. Finally, it must also be remembered that allometric equations are not error-free. Even if an equation developed explicitly for the forest typology under examination is considered (as in this study), these do not consider specific conditions (e.g., falling branches due to weather conditions), which are fundamental for this analysis. For this reason the reference values only provide us with reference values, but more precise values are expected to be obtained through proper tree modelling.

This first application of a complete workflow for individual tree detection, wood and foliage classification, and biomass estimation with LiPhe shows that it performs outstandingly. This analysis is a first step of a more extensive study. More rigorous and in-depth investigations are needed to compare different segmentation and parameter estimation methods in other experiments based on various data. However, further investigations are required and are currently ongoing. Suggested future works include the improvement of the segmentation procedure to obtain a point cloud such as to allow plausible modeling of the tree. Subsequently, it is necessary to extend the study to different tree species to generalize the model and apply it in every scenario; finally, it is essential to analyze the multitemporal variation, to evaluate the interaction of the forest ecosystem concerning climate change.

5. CONCLUSION

Forest monitoring and biomass assessment are heatedly debated topics by the scientific community and public opinion. Traditional methodologies are often inadequate due to the difficulty of acquiring information about biomass in an extensive but at the same time efficient way. However, modern instruments such as the laser scanner and the implementation of automatic processing algorithms make these analyses easier. This study addresses using oblique point clouds in forestry environments for precision forestry purposes. In particular, a methodology has been developed that, starting from the raw point cloud, makes it possible to identify single trees, perform a wood-leaf classification, and estimate the biomass of each tree. Segmentation and classification results were validated through visual interpretation, while biomass was validated with reference values calculated with allometric equations. The results are promising, and when the workflow is efficient in all its parts, limited percentage errors are obtained on the biomass estimation of silver birches.

Nevertheless, further tests are needed to improve the entire procedure, generalize it to all tree species and different scenarios, and implement a multitemporal methodology; finally, it is essential to relate biomass to the amount of carbon it captures and stores.

ACKNOWLEDGEMENTS

The corresponding author would like to acknowledge Prof. Antero Kukko from the Department of Remote Sensing and Photogrammetry at the Finnish Geospatial Research Institute for allowing him to collaborate in person with his research group, and Dr. Eetu Puttonen for sharing the database used in this study.

Tower project information:

316096/320075/**345393** – "Upscaling of carbon intake and water balance models of individual trees to wider areas with short interval laser scanning time series"

REFERENCES

Åkerblom, M., Raumonen, P., Mäkipää, R., Kaasalainen, M., 2017. Automatic tree species recognition with quantitative structure models. Remote Sensing of Environment 191, 1–12. https://doi.org/10.1016/j.rse.2016.12.002

Arseniou, G., MacFarlane, D.W., Calders, K., Baker, M., 2023. Accuracy differences in aboveground woody biomass estimation with terrestrial laser scanning for trees in urban and rural forests and different leaf conditions. Trees 37, 761–779. https://doi.org/10.1007/s00468-022-02382-1

Belcore, E., Wawrzaszek, A., Wozniak, E., Grasso, N., Piras, M., 2020. Individual Tree Detection from UAV Imagery Using Hölder Exponent. Remote Sensing 12, 2407. https://doi.org/10.3390/rs12152407

Campos, M.B., Litkey, P., Wang, Y., Chen, Y., Hyyti, H., Hyyppä, J., Puttonen, E., 2021. A Long-Term Terrestrial Laser Scanning Measurement Station to Continuously Monitor Structural and Phenological Dynamics of Boreal Forest Canopy. Front. Plant Sci. 11, 606752. https://doi.org/10.3389/fpls.2020.606752

Dalponte, M., Coomes, D.A., 2016. Tree-centric mapping of forest carbon density from airborne laser scanning and hyperspectral data. Methods Ecol Evol 7, 1236–1245. https://doi.org/10.1111/2041-210X.12575

Global Forest Resources Assessment 2020, 2020. . FAO. https://doi.org/10.4060/ca9825en

Haibo He, Yang Bai, Garcia, E.A., Shutao Li, 2008. ADASYN: Adaptive synthetic sampling approach for imbalanced learning, in: 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence). Presented at the 2008 IEEE International Joint Conference on Neural Networks (IJCNN 2008 - Hong Kong), IEEE, Hong Kong, China, pp. 1322–1328. https://doi.org/10.1109/IJCNN.2008.4633969

Kangas, A., Pitkänen, T.P., Mehtätalo, L., Heikkinen, J., 2023. Mixed linear and non-linear tree volume models with regional parameters to main tree species in Finland. Forestry: An International Journal of Forest Research 96, 188–206. https://doi.org/10.1093/forestry/cpac038

Luo, H., Khoshelham, K., Chen, C., He, H., 2021. Individual tree extraction from urban mobile laser scanning point clouds using deep pointwise direction embedding. ISPRS Journal of Photogrammetry and Remote Sensing 175, 326–339. https://doi.org/10.1016/j.isprsjprs.2021.03.002

Ma, K., Chen, Z., Fu, L., Tian, W., Jiang, F., Yi, J., Du, Z., Sun, H., 2022. Performance and Sensitivity of Individual Tree Segmentation Methods for UAV-LiDAR in Multiple Forest Types. Remote Sensing 14, 298. https://doi.org/10.3390/rs14020298

Mehtätalo, L., Lappi, J., 2020. Biometry for Forestry and Environmental Data: with Examples in R, 1st ed. Chapman and Hall/CRC, Boca Raton, FL : CRC Press, 2020. | Series: Chapman & Hall/CRC applied environmental statistics. https://doi.org/10.1201/9780429173462

Qiu, L., Jing, L., Hu, B., Li, H., Tang, Y., 2020. A New Individual Tree Crown Delineation Method for High Resolution Multispectral Imagery. Remote Sensing 12, 585. https://doi.org/10.3390/rs12030585

Raumonen, P., Åkerblom, M., 2022. InverseTampere/TreeQSM: Version 2.4.1. https://doi.org/10.5281/ZENODO.6539580

Shendryk, I., Broich, M., Tulbure, M.G., Alexandrov, S.V., 2016. Bottom-up delineation of individual trees from full-waveform airborne laser scans in a structurally complex eucalypt forest. Remote Sensing of Environment 173, 69–83. https://doi.org/10.1016/j.rse.2015.11.008

Spadavecchia, C., Belcore, E., Piras, M., Kobal, M., 2022. An Automatic Individual Tree 3D Change Detection Method for Allometric Parameters Estimation in Mixed Uneven-Aged Forest Stands from ALS Data. Remote Sensing 14, 4666. https://doi.org/10.3390/rs14184666

Spiekermann, R.I., McColl, S., Fuller, I., Dymond, J., Burkitt, L., Smith, H.G., 2021. Quantifying the influence of individual trees on slope stability at landscape scale. Journal of Environmental Management 286, 112194. https://doi.org/10.1016/j.jenvman.2021.112194

Sun, J., Wang, P., Gao, Z., Liu, Zichu, Li, Y., Gan, X., Liu, Zhongnan, 2021. Wood–Leaf Classification of Tree Point Cloud Based on Intensity and Geometric Information. Remote Sensing 13, 4050. https://doi.org/10.3390/rs13204050

Vicari, M.B., Disney, M., Wilkes, P., Burt, A., Calders, K., Woodgate, W., 2019. Leaf and wood classification framework for terrestrial LiDAR point clouds. Methods Ecol Evol 10, 680–694. https://doi.org/10.1111/2041-210X.13144

Wang, D., Momo Takoudjou, S., Casella, E., 2020. LeWoS: A universal leaf-wood classification method to facilitate the 3D modelling of large tropical trees using terrestrial LiDAR. Methods Ecol Evol 11, 376–389. https://doi.org/10.1111/2041-210X.13342

Xu, D., Wang, H., Xu, W., Luan, Z., Xu, X., 2021. LiDAR Applications to Estimate Forest Biomass at Individual Tree Scale: Opportunities, Challenges and Future Perspectives. Forests 12, 550. https://doi.org/10.3390/f12050550

Xu, X., Iuricich, F., Calders, K., Armston, J., De Floriani, L., 2023. Topology-based individual tree segmentation for automated processing of terrestrial laser scanning point clouds. International Journal of Applied Earth Observation and Geoinformation 116, 103145. https://doi.org/10.1016/j.jag.2022.103145

Zörner, J., Dymond, J., Shepherd, J., Wiser, S., Jolly, B., 2018. LiDAR-Based Regional Inventory of Tall Trees—Wellington, New Zealand. Forests 9, 702. https://doi.org/10.3390/f9110702